

# Local Scales in Imaging

---

A Thesis  
Presented to the  
Faculty of  
San Diego State University

---

In Partial Fulfillment  
of the Requirements for the Degree  
Master of Science in Applied Mathematics  
with a Concentration in  
Dynamical Systems

---

by  
Jesus Rodolfo Perez Cuarenta  
Spring 2021

# SAN DIEGO STATE UNIVERSITY

The Undersigned Faculty Committee Approves the  
Thesis of Jesus Rodolfo Perez Cuarenta:

Local Scales in Imaging

---

Jérôme Gilles, Chair  
Department of Mathematics and Statistics

---

Christopher W. Curtis  
Department of Mathematics and Statistics

---

Ashkan Ashrafi  
Department of Electrical and Computer Engineering

---

Approval Date

Copyright © 2021  
by  
Jesus Rodolfo Perez Cuarenta

## DEDICATION

This work is dedicated to my parents for their unconditional love as well as my godparents for opening their doors which allowed me to pursue mathematical studies outside of my native country.

-

Dicen que lo que se ve no se pregunta.

– Alberto Aguilera Valadez

# ABSTRACT OF THE THESIS

Local Scales in Imaging

by

Jesus Rodolfo Perez Cuarenta

Master of Science in Applied Mathematics with a Concentration in Dynamical Systems  
San Diego State University, 2021

The field of functional analysis has provided insight into function spaces not obtained through the lens of linear algebra. These spaces are characterized by defining a norm which allows for measuring distinct degrees of smoothness (scale)  $t$ , say for an image  $f$ . Historically, we have seen the development of Hölder, Sobolev, and Hardy spaces culminate as special cases of Besov and Triebel-Lizorkin spaces. This work focuses on the latter space,  $\dot{F}_{p,\infty}^\alpha$ , with a connection to the fields of neuroscience and computer vision. A definition for local scale selection of  $f$  is provided using both isotropic and anisotropic Gaussian kernels. Numerical results include applications related to image classification and kernel estimation for deblurring processes in the isotropic case. In the anisotropic case results are in accordance with image features related to scale and orientation.

# TABLE OF CONTENTS

	PAGE
ABSTRACT .....	vi
LIST OF FIGURES .....	viii
ACKNOWLEDGMENTS .....	x
CHAPTER	
1 INTRODUCTION .....	1
1.1 History .....	2
1.2 The Search for an Adequate Convolution Kernel .....	4
2 FEATURE DETECTION IN IMAGES .....	8
2.1 Background .....	8
2.2 Local Scales .....	9
2.3 Numerics Involving $Sf(x, t)$ .....	14
3 LOCAL SCALES IN APPLICATIONS .....	18
3.1 Clustering Image Regions Based on $ Sf $ Similarity .....	18
3.2 On the Scale-Space Representation of $K_t * f$ .....	22
4 GAUSSIAN ANISOTROPIC KERNEL OF INCREASING SCALE .....	28
4.1 Preservation of Scale-Space Properties .....	28
4.2 Numerics Involving $Sf(x, t, \theta)$ .....	29
5 FINAL REMARKS .....	34
BIBLIOGRAPHY .....	35

# LIST OF FIGURES

	PAGE
1.1 Applying the Canny edge detector (bottom) to our input image (top) using Matlab's <i>edge</i> function. ....	7
2.1 Images of Barbara with decreasing detail from top to bottom, left to right. Image blur implemented in Matlab using the <i>fspecial</i> function with a filter size of 20 and the specified standard deviation, $\sqrt{t}$ . ....	10
2.2 Top: input signal, middle: Laplacian of Gaussian kernel, bottom: convolution between the input signal and the Laplacian of Gaussian kernel. Note that these graphs correspond to $t = 0.3$ . ....	11
2.3 Laplacian response for increasing values of scale, $t$ . From top to bottom $t$ attains the values 0.1, 0.6, 1.1, and 1.6. ....	12
2.4 Normalized Laplacian response for increasing values of scale, $t$ . From top to bottom $t$ attains the values 0.1, 0.6, 1.1, and 1.6. Note the consistent range of $t\partial_t K_t * \chi(x)$ compared to curves in Figure 2.3. .	14
2.5 Binary image consisting of large, medium, and small squares with highlighted regions where $ Sf $ will be computed. Red dots are labeled $P_1, P_2$ and $P_3$ from top to bottom, left to right. ....	16
2.6 The oscillatory level, $ Sf $ , of $f$ at large square region $P_1$ (leftmost column), medium square region $P_3$ (middle column), small square region $P_2$ (rightmost column) for increasing values of $\alpha$ (top to bottom). The vertical axis corresponds to $ Sf $ and the horizontal axis to the discrete variable $\tau$ . See Figure 2.5 to identify regions. ....	17
3.1 Average of all $ Sf $ curves with respect to position $x$ . We separate three regions where $ Sf $ maxima is prevalent by placing a vertical line near local minima. ....	19
3.2 Classifying all pixels (bottom) from Figure 2.5 (top) dependant on absolute maximum values of $ Sf $ with respect to regions $R_1$ (dark blue), $R_2$ (cyan), and $R_3$ (yellow). We denote this technique the $R_i$ Method. ....	20
3.3 Side-by-side comparison of $R_i$ (top) and $k$ -means (bottom) clustering methods. ....	21
3.4 Side-by-side comparison of $k$ -means with three (top) and ten (bottom) clusters. ....	22
3.5 Graphing $f$ using the <i>surf</i> command in Matlab. Note the image rotation to highlight the oscillatory region. ....	24
3.6 Influence of blurring evident in the small-sized squares region. ....	25



3.7	Influence of blurring evident in the medium-sized squares region. ....	26
3.8	Influence of blurring evident in the large-sized squares region. ....	27
4.1	Image $f$ with three highlighted pixels where $ Sf(x, t, \theta) $ is computed. From top to bottom we have assigned the labels Pixel 2, 1, and 3 to the locations of colored dots respectively. ....	30
4.2	Computations of $ Sf $ at Pixel 1 with various values of $\alpha$ showcasing extrema bias. ....	31
4.3	Computations of $ Sf $ at Pixel 2 with various values of $\alpha$ showcasing extrema bias. ....	32
4.4	Computations of $ Sf $ at Pixel 3 with various values of $\alpha$ showcasing extrema bias. ....	33

## ACKNOWLEDGMENTS

I would like to sincerely give gratitude to everyone in my committee for accepting the invitation and taking the time from their busy schedules to examine this thesis work. I have the utmost respect and admiration for your research and teaching, guiding people like myself in understanding nature and the secrets she keeps.

Specifically, thank you Dr. Gilles for your insight into the wonderful world of harmonic analysis. I am in awe of your ability to pass along complicated ideas in intuitive fashion. I will miss our mathematical conversations as well as your stories.

On a similar note thank you Dr. Curtis for your enlightening lectures and problem sets. I always looked forward to hearing your perspective on mathematical topics, and even philosophical ones on few occasions. I am still upset for not having the opportunity to attend your lectures on complex analysis.

Additionally, thank you Dr. Ashrafi for your promptness in responding to my invitation. I recall the myriad of students waiting outside your office before our first meeting. Regardless, your welcoming was kind and calm. I find your dedication inspirational.

Outside of my committee I am grateful to Drs. Mahaffy and Salamon for providing me with rewarding *challenges* and opportunities. I am also blessed to have shared time with Teri Luque, Burak Cebecioglu, A. Minerva García, and Alex Mancinelli at Southwestern College. You have all in some way or another shaped who I am today.

# CHAPTER 1

## INTRODUCTION

The study of properties encoded in digital images has cemented its importance within different scientific communities in the past few decades. As a consequence, the motivation to further our understanding of image processing tools is fueled by a vast amount of applications in engineering, biology, computational sciences, as well as theoretical results in mathematics. It is difficult to underestimate the importance of this field of study once we consider the technological advances we make use of in our day to day. From computed tomography to image and video compression amidst a pandemic, image processing is the cornerstone of several technological applications. Among the myriad of techniques related to image processing, there exists the notion of scale-space theory.

We further elaborate on the aforementioned theory by considering a task shared by the human eye and a digital camera: focusing an image. The camera lens can be adjusted in order to focus on regions with mainly geometric components, such as a picture of a landscape, or to capture fine details, such as an image of a fingerprint. On the other hand there is the human eye; here the cornea and lens, in the anatomical sense, are responsible for making a distinction between coarse regions and fine detail. Additionally, it is an intuitive process for humans to describe objects with an appropriate unit, e.g., it is natural to measure the distance between countries and the size of an ant in kilometers and millimeters correspondingly. The question at hand now is how can a computer achieve this task without *a priori* knowledge of which regions or pixels, in a given image, correspond to faces, edges, roads, or tumor tissue (in the case of MRI based brain tumor detection) in a consistent manner? This is our motivation to represent an image in a space distinct to  $\mathbb{R}^n$  where characteristics we are interested in can be identified.

In this work we first layout a brief history of feature detection in images. Next we explore the Gaussian *scale-space* representation of images followed by a localized scale selection method motivated by the Triebel-Lizorkin function space,  $\dot{F}_{p,\infty}^\alpha(\mathbb{R}^n)$ . This is accompanied by a discussion on the implemented algorithm. In the third chapter, we turn our attention to image segmentation and Gaussian blurring. Here we apply the extracted scales to classify image regions via a *k-means* clustering approach, and mention the possibility of identifying blurring Gaussian kernels without *a priori*

knowledge of such kernel. In the fourth chapter, we extend our algorithm to include an additional parameter in hopes of detecting structures in images, such as texture orientation, and demonstrate the obtained results. We conclude our work in the fifth chapter with final remarks and possible future implementations.

## 1.1 History

Let us begin with a curious question: How do humans process features such as edges, blobs, and orientation of objects, portrayed in the field of vision? In other words, which mechanism in our brain is responsible for making this interpretation a trivial task? The seminal work in [10], published in 1959, gave us some insight on this question. In the aforementioned paper, Hubel and Wiesel observed the brain activity in cats, a process facilitated by placing an electrode into the cat's visual cortex in order to record neuronal activity. The motivation for such a task was to identify the relationship between specific visual stimuli and neuronal activity. Initially, the neural response to image slides was somewhat cryptic and this raised doubts as to whether the experiment would be successful. The experiment did not produce the originally intended results since the neural response was too complex to classify. Instead, a response was observed to be consistent, not with the content within the slides, but with the act of changing them. This event was linked to the movement of edges across the projected space. Later we will realize that edge detection in images plays an important role in scale selection hence the relevance of such work, regardless of it being done outside of mathematics.

In the early 60's, Larry Roberts, recognized as father of the ARPANET <sup>1</sup>, published his PhD thesis [26] outlining a novel automated method for edge detection. In Roberts' words, the relevance of feature detection in images is described as follows:

It is assumed that a photograph is a perspective projection of a set of objects which can be constructed from transformations of known three-dimensional models, and that the objects are supported by other visible objects or by a ground plane. These assumptions enable a computer to obtain a reasonable, three-dimensional description from the edge information in a photograph by means of a topological, mathematical process.

This is early evidence of the need for edge detection if one aims to identify desired characteristics in an image. We may take the interpretation that the intersection of Huebel and Wiesel's work with that of Roberts hinges on edge detection. In the following years, the Summer Vision Project [24] at Massachusetts Institute of Technology set out to construct a system capable of segmenting regions in images

---

<sup>1</sup>Advanced Research Projects Agency Network, precursor to the Internet.

corresponding to objects, background areas, and chaos. Image segmentation is a subject which has active research to date [1, 22, 27, 28]. The culmination of the project did not fully produce a system capable of the above, demonstrating the difficulty level of such a task. Notable work in the next decade is that of David Marr and his description of the stages of visual representation [21], a successor to the Summer Vision Project.

Moving forward in time, the 80's brought better performance on digital cameras. This opened the door for researchers to identify and extract features in more realistic images. Noteworthy research during this decade is that of John F. Canny where the *Canny edge detection* algorithm was first proposed [6]. We briefly motivate Canny's work, and the desire to extract edges in a given image, from the point of view of two research fields, namely neuroscience and computer vision. The former field of study witnessed early success in recording consistent neural response in the visual cortex of cats dependant on edge detection, as evidenced in Hubel and Wiesel's work. More recent work can be linked with the latter field of study and its relevance to our modern computer era. For example, training a machine learning model can quickly become a time-consuming task. In 1986 Canny wrote the following passage related to Figure 1.1 demonstrated at the end of Chapter 1.

The edge detection process serves to simplify the analysis of images by drastically reducing the amount of data to be processed, while at the same time preserving useful structural information about object boundaries.

Indeed, the reduction of processed data is valuable if one aims to carry out problems in a more efficient fashion, as in the case of the machine learning community. Although there are other popular edge detection operators, e.g., Sobel, Kayyali, and SUSAN; the Canny operator is presently relevant. Efficient edge detection in images is valuable for a wide range of scientists and researchers. Examples include fingerprint reading in the case of smartphone user validation, object detection for automated vehicles, error detection in printed circuit boards, and road detection in satellite images. A mathematical explanation of edge detection will be introduced in the next chapter prior to the discussion of scale-space representation of images.

We have discussed the value of edge detection for several research fields. Other sought-out features in images can be found by building on an edge detection framework, namely blobs and ridges. We will refer to the three characteristics consisting of edges, blobs, and ridges, as *low level descriptors*. Later in this work we introduce the notion of blob detection from a functional analysis point of view.

## 1.2 The Search for an Adequate Convolution Kernel

So far we have mentioned attempts at constructing a model for a visual system and the complexity of this task but we will not proceed with such a framework. Instead, we choose to identify and characterize those features and transformations which directly influence our vision process. We hope to narrow down our options to elementary properties and mathematical operations which can extract them.

The notion of invariance is of great interest for mathematicians and physicists. In these fields, undergraduate students spend considerable time working with eigenvalues, eigenvectors, and eigenspaces, all of which are related to invariant vector subspaces under some operator. At the research level there exists open problems such as the Invariant Subspace Problem, popularized by Paul Halmos, from which interesting mathematics<sup>2</sup> have come to fruition. From a more practical point of view we have visual processes which possess invariant properties for which we give the following example.

At stadiums people pay considerable amounts of money for priority seating during concerts and sporting events. Regardless, most attendees are able to identify specific performers or athletes at varying distance and under different lighting conditions. We have that the recognition of people or objects is not influenced by distance up to some scale. The two processed features at play in our visual cortex pertain to scale invariance and intensity transformations in the sense of measured luminosity. Hence if we aim to describe the human vision process we must consider a system which acknowledges invariant features.

From biology and neuroscience we have a better understanding of the human brain, specifically the visual cortex which is categorized in distinct visual areas. Relevant areas to our work are the primary visual cortex, secondary visual cortex, and middle temporal motion area termed V1, V2, and MT respectively. Hubel and Wiesel's work showed the functionality of cells in V1. Cells which respond to lines of a specific orientation are called simple cells. Moreover, similar cells which respond to lines over larger regions are called complex cells. It is important to appreciate the relationship between V1 cells and low level descriptors mentioned in the previous section at this point. The remaining areas, V2 and MT, can be interpreted as visual models receiving information from V1 as input data for tasks such as depth and motion perception.

For now, we will not be careful specifying which class of functions our image  $f$  belongs to. We acknowledge the importance of placing  $f$  in an adequate space in order

---

<sup>2</sup>For example Quasi-diagonal  $C^*$ -Algebras.

to properly measure regularity locally, i.e., we do not yet impose that  $f$  belongs to functional spaces such as Besov or Sobolev spaces. In mathematical terms the visual processes occurring in the primary visual cortex can be described as an operator  $K_t$  acting on image data  $f$  to produce an element  $u$  belonging to a family of visual representations  $U$

$$K_t f = u_t.$$

The purpose of an additional parameter  $t$  at this stage is to impose an ordering on elements of  $U$  since each  $u_t$  is produced by a specific  $t$ . Note that  $u_t$  is the potential input data for some visual model corresponding to V2 or MT. We will now discover a prototype for  $K_t$  and how to generate the set of visual representations corresponding to  $f, U$ .

One key property of our promised operator is that of shift-invariance. If we were to identify features corresponding to, say, a basketball at the coordinates  $(x_1, x_2)$  in an image, then we wish to extract the same features for such an object if we were to relocate the basketball at coordinates  $(x'_1, y'_1)$  such that

$$(x'_1, y'_1) = (x_1 - h, y_1 - k)$$

where the variables  $h$  and  $k$  represent the amount shifted by in each component. Mathematically we have that  $K_t$  commutes with the shift operator,  $\delta_{h,k} f(x_1, x_2) = f(x_1 - h, x_2 - k)$ , so that

$$K_t (\delta_{h,k} f) = \delta_{h,k} (K_t f).$$

By imposing linearity we arrive at the convolution operator,  $*$ , as a potential candidate to obtain

$$K_t * f = u_t.$$

Other properties we require but will not discuss in depth are a semi-group structure over  $t$ , self-similarity for  $K_t$ , rotational symmetry, and sufficient spatial and scale regularity. We give a brief description on the necessity for each item. The semi-group structure allows for  $U$  to be closed under convolution. A self-similar convolution kernel will identify repeating patterns at different scales. By considering rotational symmetry we avoid discriminating a specific orientation in space. Admitting spaces with the proper regularity allows for the use of partial differential equations to describe the evolution of elements of  $U$  with respect to a continuous variable  $t$ .

The last property to make our selection for  $K_t$  is easily motivated by an example where we consider a photograph corresponding to landscape and trees. An important

aspect for photographers is proper focus which the general public might interpret as being related to sharpness. Taking a correctly focused image would allow our eyes to appreciate fine details in the picture such as tree leaves. Now suppose we have a knob which controls the intensity of a blurring effect on the image by the parameter  $t$ . As  $t$  increases we would witness the dissipation of sharp details without losing coarse characteristics. At no point in the blurring effect would we see an object become sharper. This property is known as non-enhancement of local extrema and is the last characteristic we need to arrive at  $K_t$  being none other than the Gaussian kernel. Readers are referred to [18] for a detailed analysis of the aforementioned properties. We proceed with the definition for  $K_t$  and its remarkable properties in the subsequent chapter.



## Original Image



## Canny Edge Detector



Figure 1.1. Applying the Canny edge detector (bottom) to our input image (top) using Matlab's *edge* function.

## CHAPTER 2

### FEATURE DETECTION IN IMAGES

An elementary approach to study information encoded in images is via a Gaussian *scale-space*. Let an image  $f(x)$  be defined on  $\mathbb{R}^2$ . In this chapter we seek a transformation of  $f$  such that spatial measurements can be approximated by automatic methods.

#### 2.1 Background

Throughout this work we admit the Euclidean norm in  $\mathbb{R}^2$ , i.e., for  $x = (x_1, x_2)$  we have

$$\|x\| = \sqrt{x_1^2 + x_2^2}$$

unless stated otherwise. We also consider the Gaussian kernel,  $K_t(x)$ , defined by

$$K_t(x) = \frac{1}{\sqrt{2t}} \exp\left(-\pi \frac{\|x\|^2}{2t}\right), \quad t > 0.$$

From the point of view of a normal distribution we have that  $t$  corresponds to the variance of the distribution. We will soon realize  $t$  is directly related to the notion of scale. From classical results of Fourier analysis we know that the Fourier transform of  $K_t$ ,  $\mathcal{F}\{K_t\}$ , corresponds to another Gaussian as a function of  $\xi = (\xi_1, \xi_2)$ ,

$$\mathcal{F}\{K_t\}(\xi) = \exp(-2\pi t \|\xi\|^2).$$

Now since the Fourier transform of the convolution,  $u(x, t) = K_t * f$ , over the spatial domain corresponds to a point-wise multiplication in the Fourier domain, i.e.,

$$\mathcal{F}\{u\}(\xi) = \mathcal{F}\{K_t * f\}(\xi) = \mathcal{F}\{K_t\}(\xi) \cdot \mathcal{F}\{f\}(\xi),$$

we make the following observations. First, Gaussian kernels have fast decay in the Fourier domain which means we may allow for our image  $f$  to be highly oscillatory. Second, the Gaussian  $\mathcal{F}\{K_t\}$  is well-behaved in the sense that its smoothness is able to eliminate frequencies belonging to  $f$  without creating incoherent structures in the Fourier domain. In a sense, we are in a position to represent the input image,  $f$ , for different values of  $t$ , each corresponding to low-pass filters. We now define the Gaussian scale-space representation of an image  $f$ .

**Definition 2.1.** Let  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  be a given image and  $K_t : \mathbb{R}^2 \rightarrow \mathbb{R}$  be defined by

$$K_t(x) = \frac{1}{\sqrt{2t}} \exp\left(-\pi \frac{\|x\|^2}{2t}\right), \quad t > 0.$$

Then the Gaussian scale-space representation of  $f$ ,  $u(x, t) : \mathbb{R}^2 \times \mathbb{R}_+ \rightarrow \mathbb{R}$ , is defined by

$$u(x, t) = K_t(x) * f(x).$$

where  $x \in \mathbb{R}^2$  and  $*$  denotes the convolution operation over the spatial variable  $x$ . We refer to  $t$  as the scale parameter.

An example is given in Figure 2.1, where an image  $f$  corresponding to Barbara is represented with increasing values of  $t$  in accordance with the expression of  $u(x, t)$ . We notice that the details in the garments and boundaries dissipates as  $t$  increases. For values of  $t$  close to zero we have that  $K_t(x)$  is an approximation to the Dirac delta function and thus  $u(x, t) \approx f(x)$ . In contrast, the representation of  $f$  at larger scales shows a cartoon-like image lacking in details.

The evolution of the input image  $f$  with respect to  $t$  is intimately related to a diffusion process. Indeed,  $u(x, t)$  satisfies the 2D heat equation,

$$\left(\frac{\partial}{\partial t} - \sum_{i=1}^2 \frac{\partial^2}{\partial x_i^2}\right) u(x, t) = 0, \quad u(x, 0) = f(x). \quad (2.1)$$

Hence, an equivalent method of defining the scale-space representation of  $f$  is achieved by using solutions to the heat equation corresponding to isotropic diffusion produced by the Gaussian kernel. These solutions coincide with the dissipation of fine details such as lines, edges, and corners for large values of  $t$ . There is extensive literature on applications involving Gaussian scale-space representation. The reader is referred to [5, 16, 17, 20, 30]. These applications depend on the choice of a single local scale for a given pixel. In [11], P. Jones and T. M. Le take a different approach by assigning more than one scale to  $x$ . This is explored further in the next section.

## 2.2 Local Scales

It is sane to assume that  $f$  may have various scales of oscillations encoded locally as in the case of texture. This is motivation to study a vector of scales dependant on the location, i.e.,  $t(x) = (t_1, t_2, \dots, t_N), x \in \mathbb{R}^2$ . To proceed, we have to place  $f$  in an adequate space. We require a function space where it is possible to measure the local regularity of  $f$ . The following definition for the Triebel-Lizorkin space is taken from H. Triebel [32] and is characterized by the Gaussian kernel,  $K_t$ .

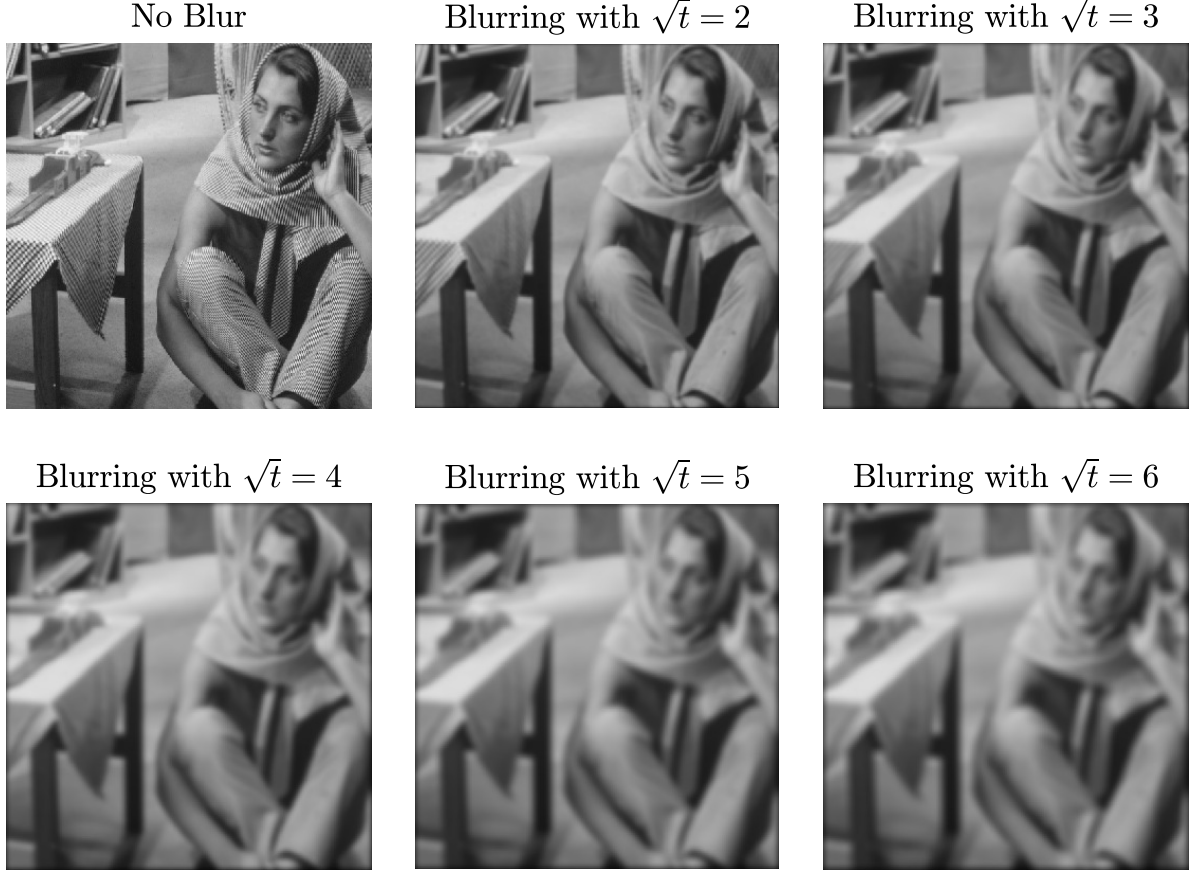


Figure 2.1. Images of Barbara with decreasing detail from top to bottom, left to right. Image blur implemented in Matlab using the *fspecial* function with a filter size of 20 and the specified standard deviation,  $\sqrt{t}$ .

**Definition 2.2.** For any  $\alpha \in \mathbb{R}$ , let  $m \in \mathbb{N}_0$ , such that  $m > \frac{\alpha}{2}$ , we say a function  $f$  belongs to the homogeneous Triebel-Lizorkin function space  $\dot{F}_{p,\infty}^\alpha(\mathbb{R}^2)$ ,  $1 \leq p \leq \infty$ , if

$$\|f\|_{\dot{F}_{p,\infty}^\alpha(\mathbb{R}^2)} = \left\| \sup_{0 < t < \infty} \left\{ t^{m-\frac{\alpha}{2}} \left| \frac{\partial^m}{\partial t^m} K_t * f \right| \right\} \right\|_{L^p(\mathbb{R}^2)} < \infty,$$

and we say  $f \in \dot{F}_{p,\infty}^\alpha(\Omega)$ ,  $1 \leq p \leq \infty$  if

$$\|f\|_{\dot{F}_{p,\infty}^\alpha(\Omega)} = \left\| \sup_{0 < t < 1} \left\{ t^{m-\frac{\alpha}{2}} \left| \frac{\partial^m}{\partial t^m} K_t * f \right| \right\} \right\|_{L^p(\Omega)} < \infty,$$

where  $\Omega = [0, 1]^2$  is the unit square.

It can be shown that taking any  $m \in \mathbb{N}$ , such that  $m > \frac{\alpha}{2}$ , yields equivalent norms for the Triebel-Lizorkin function space. Throughout this work we consider the case  $m = 1$ . To measure the oscillatory level of  $f$ , we first define

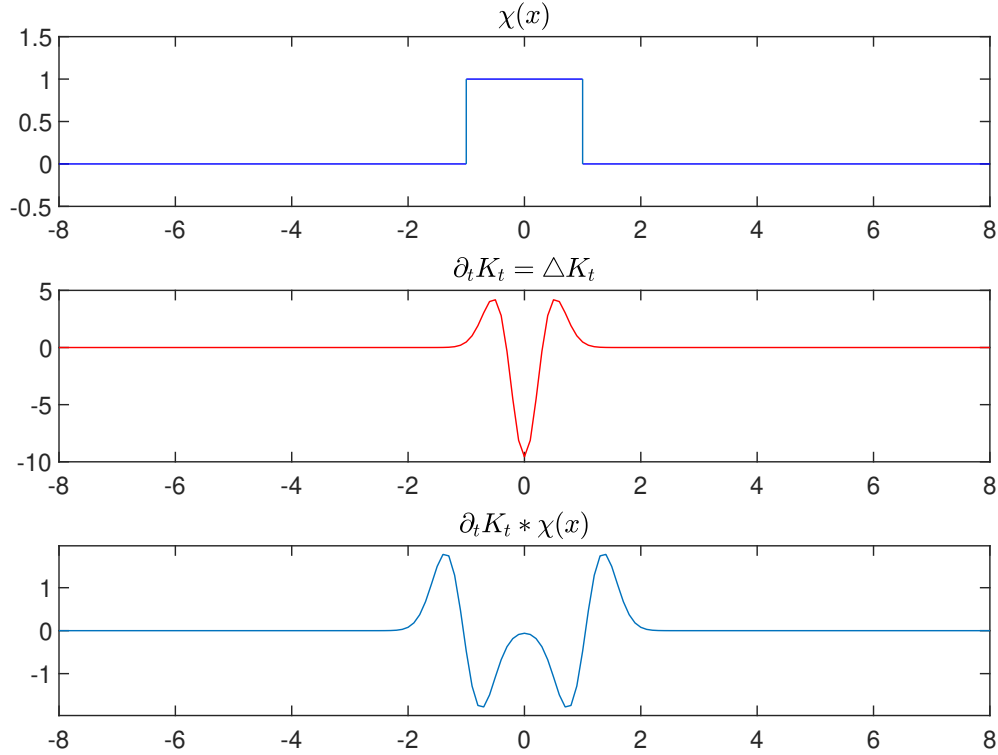
$$Sf(x, t) = t^{1-\frac{\alpha}{2}} \frac{\partial}{\partial t} K_t(x) * f(x). \quad (2.2)$$

Note that we require  $\alpha < 2$  based on our choice of  $m$ . We also have that if  $f \in \dot{F}_{p,\infty}^\alpha(\mathbb{R}^2)$  then  $Sf$  is defined for any  $t > 0$ . If we consider  $f \in \dot{F}_{p,\infty}^\alpha(\Omega)$  then  $Sf$  is defined for all  $0 < t < 1$ . We proceed to explore the role of the partial derivative with respect to  $t$  of  $u_t$ ,  $\partial_t K_t * f$ , in Equation (2.2) with the following example for a 1D signal. Take the characteristic function

$$\chi(x) = \begin{cases} 1 & \text{if } -1 \leq x \leq 1, \\ 0 & \text{else.} \end{cases}$$

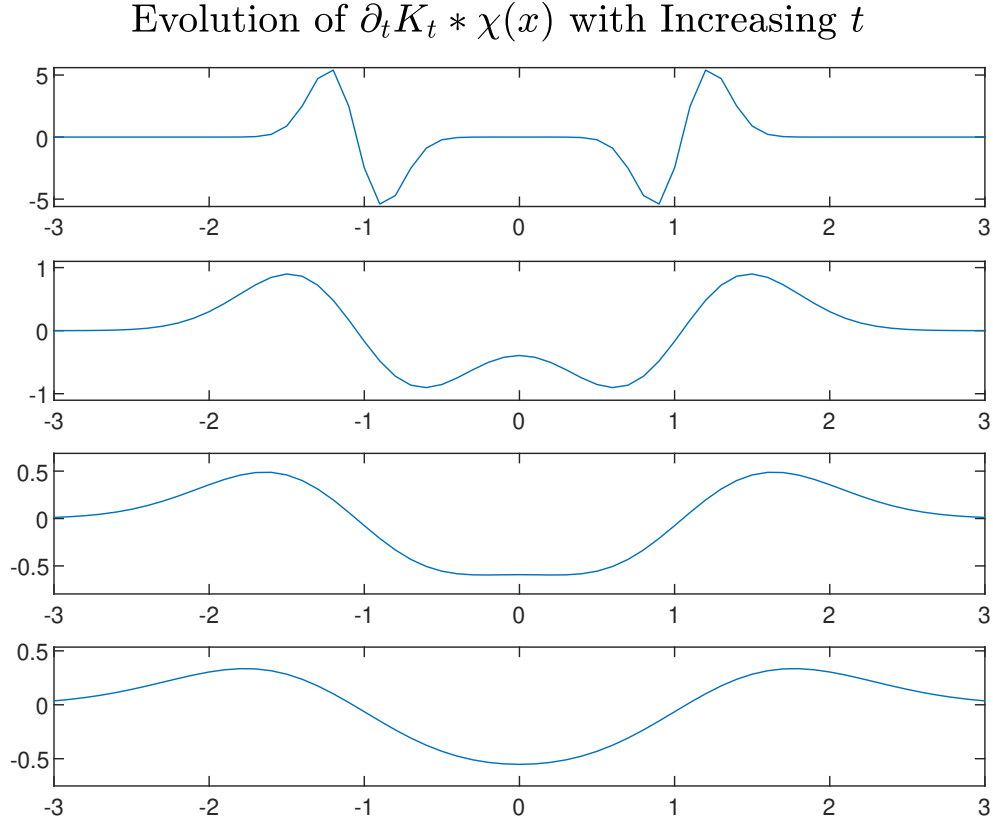
We now present Figures 2.2, 2.3. and 2.4. Starting with Figure 2.2, in the top we have the characteristic function,  $\chi(x)$ . This is the signal for which we wish to assign an appropriate scale. The middle graph shows the partial derivative of  $K_t$  with respect to  $t$ . Since Equation (2.1) is also solved by  $K_t$  we can relate  $\partial_t K_t$  with the Laplacian operator,  $\Delta$ . The remaining graph shows the convolution of  $f$  with  $\Delta K_t$ .

### Characteristic Function, Laplacian of Gaussian, and Convolution



**Figure 2.2.** Top: input signal, middle: Laplacian of Gaussian kernel, bottom: convolution between the input signal and the Laplacian of Gaussian kernel. Note that these graphs correspond to  $t = 0.3$ .

The last graph sheds some light on Equation (2.2). The edges of  $\chi(x)$  coincide with the zero crossings of the Laplacian response. We may appreciate how features pertaining to  $f$  are also apparent in the convolution. We remark that the behavior of  $\partial_t K_t * \chi(x)$  is characterized by the choice of  $t$ , the scale parameter. The characteristic we are mainly interested in for this example is the interval for which  $\chi(x) = 1$ , which in the 2D case corresponds to a blob. We proceed with a plot showing a family of convolution curves for various values of  $t$ .



**Figure 2.3.** Laplacian response for increasing values of scale,  $t$ . From top to bottom  $t$  attains the values 0.1, 0.6, 1.1, and 1.6.

In Figure 2.3 we notice the ripples in the last curve collide to create a global minimum. The zero crossings still coincide with the edges of  $\chi(x)$ , but the width of the valley of the Laplacian matches the interval for which  $\chi(x) = 1$ . It is noteworthy that measurements of the characteristic function can be made indirectly by studying  $\partial_t u(x, t)$ . We refer the reader to [14, 17] where it is shown the frequency of a sinusoidal function can be estimated by the scale parameter analytically. Another observation is

that the magnitude of  $\partial_t u(x, t)$  decays as  $t$  increases. The problem here is that the response is not scale-invariant. Indeed, the 1D Laplacian of  $K_t$

$$\partial_t K_t(x) = \frac{\partial^2}{\partial x^2} K_t(x) = -\frac{\pi \sqrt{2} \exp(-\frac{\pi x^2}{2t}) (t - \pi x^2)}{2 t^{5/2}}$$

has local extrema at the critical points  $x \in \{0, \pm \frac{\sqrt{3}\sqrt{t}}{\sqrt{\pi}}\}$ , which gives us

$$\max(\partial_t K) = \frac{\pi \sqrt{2} e^{-\frac{3}{2}}}{t^{3/2}}.$$

It becomes apparent  $\partial_t K_t$  tends to the zero function for sufficiently large  $t$ . To remedy the decay in the magnitude of the Laplacian response for increasing scales, we resort to a normalized derivative. Normalization for a Gaussian (spatial) derivative of order  $\gamma \in \mathbb{N}$  is achieved by multiplying by corresponding powers of the standard deviation,  $(\sqrt{t})^\gamma$ . For the Laplacian, we have  $\gamma = 2$  and thus multiply by  $(\sqrt{t})^2 = t$  which coincides with the equation

$$t \partial_t K_t * f(x) = t \frac{\partial^2}{\partial x^2} K_t * f(x).$$

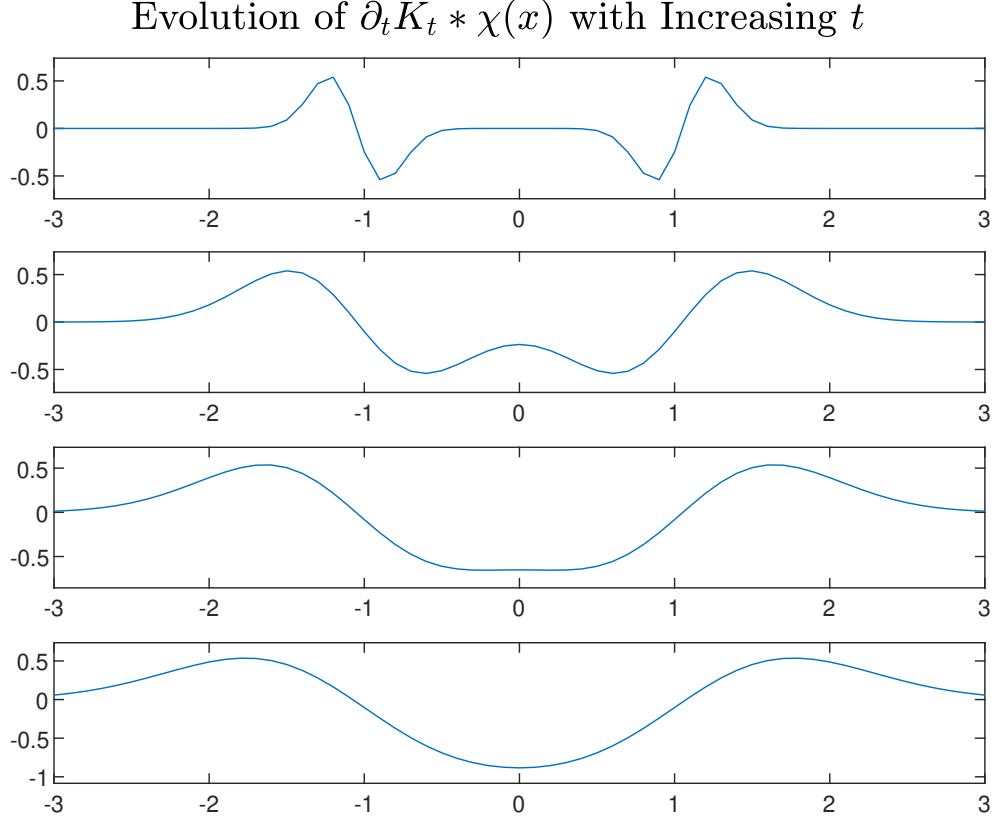
In Figure 2.4 we present the same functions from Figure 2.3 but now we normalize the derivative and observe the adjusted window for the magnitude of  $t \partial_t K_t$ .

We are now better equipped to understand Definition 2.1 which will soon be tied to the scale-selection method in this work. The last item to end the exploration of Equation (2.2) is the term  $t^{-\frac{\alpha}{2}}$ . In short, the mapping  $x \mapsto t^{-\frac{\alpha}{2}} x$  corresponds to dimensionless coordinates and the degree of freedom,  $\alpha$ , arises by necessity as evidenced by carefully chosen principles for scale selection. These principles include a normalized derivative operator to account for the decrease in magnitude of the Laplacian response and the preservation of extrema across scales. The effect of varying  $\alpha$  will be explored in the numerical results section. The reader is redirected to the Appendix in [17] for a comprehensive discussion on the dimensional analysis of the spatial variables and introduction of  $\alpha$ . We now define the local scales of  $f$  at  $x$  as given in [11].

**Definition 2.3.** Let  $f \in \dot{F}_{p,\infty}^\alpha$ ,  $\alpha < 2$ ,  $1 \leq p < \infty$ . The local scales of  $f$  at  $x$  are defined to be the set  $T_f(x)$  consisting of  $t \geq 0$ , such that  $|Sf(x, t)|$  is a local maximum. Moreover, for each  $t \in T_f(x)$ , if there exists an interval  $[a, b]$ ,  $a < b$ , containing  $t$  such that for all  $t \in [a, b]$ ,

$$\frac{\partial}{\partial t} Sf(x, t) = 0,$$

then we only select  $t = a$  to be in  $T_f(x)$  and discard all other  $t' \in (a, b]$ . Therefore, if  $t$  and  $t'$  belong to  $T_f(x)$ , then  $|t - t'| > 0$ . For each  $x$ , we refer to  $|Sf(x, t)|$  as the **oscillatory level** of  $f$  at scale  $t \in T_f(x)$ .



**Figure 2.4.** Normalized Laplacian response for increasing values of scale,  $t$ . From top to bottom  $t$  attains the values 0.1, 0.6, 1.1, and 1.6. Note the consistent range of  $t\partial_t K_t * \chi(x)$  compared to curves in Figure 2.3.

In the following section we present numerical results regarding the computation of  $Sf(x, t)$  for distinct values of  $\alpha$  at a given location  $x$ .

### 2.3 Numerics Involving $Sf(x, t)$

The algorithm implemented to extract local scales assumes the substitution  $\tau = \log_a(t)$ . The reasoning here is that for each  $f \in \dot{F}_{p,\infty}^\alpha$ , we have that

$$\left| \frac{\partial Sf(x, t)}{\partial t} \right| \leq Ct^{-1} \|f\|_{\dot{F}_{p,\infty}^\alpha},$$

as well as

$$\left| \frac{\partial Sf(x, t)}{\partial \tau} \right| \leq C \ln(a) \|f\|_{\dot{F}_{p,\infty}^\alpha},$$

result in a transition between scales of order  $\mathcal{O}(t)$  and  $\mathcal{O}(1)$  respectively for  $t$  and  $\tau$  as demonstrated in [11]. Since we consider images of size 256 by 256 we take  $a = (0.95)^{-1}$



and compute scales over the set

$$T_f(x) = \{t_0, t_1, \dots, t_N\} = \{a^{\tau_0}, a^{\tau_1}, \dots, a^{\tau_N}\}$$

where  $\forall i \in [0, 1, 2, \dots, N]$ ,  $t_i = a^{\tau_i}$ . The initial scale is selected to be  $\tau_0 = -290$  so that  $K_t$  is an approximation to the Dirac delta function in the spatial domain while  $\mathcal{F}\{K_t\}(\xi)$  mimics the indicator function

$$\forall \xi \in \mathbb{C}, \mathbb{1}(\xi) = 1$$

in frequency space. On the other hand, the major scale  $\tau_N = -35$  yields a kernel which approximates

$$\forall x \in \mathbb{R}^2, \mathbb{1}(x) = 1$$

in spatial coordinates and the unit impulse function in the Fourier domain. Note that the scale interval may be chosen depending on the application. In [17], the author accounts for an interval such that

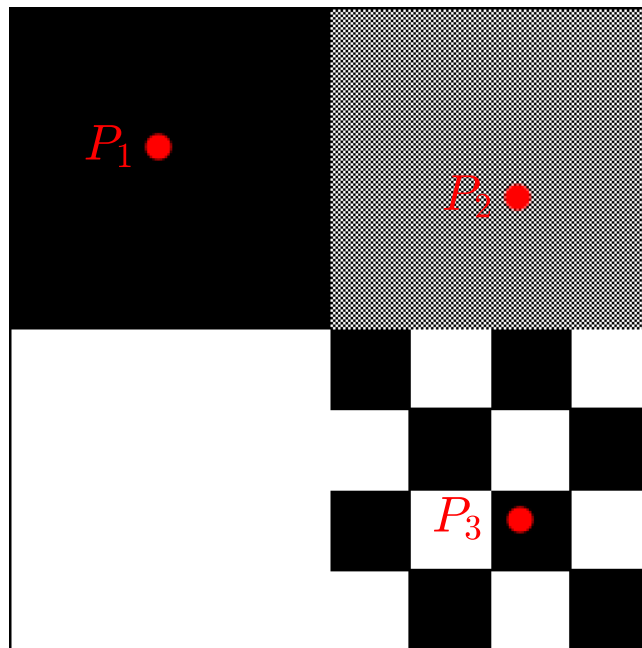
$$\frac{t_{i+1}}{t_i} \approx C, \quad C \in \mathbb{R}_+.$$

The first 2D example we study pertains to an image where we intuitively anticipate our algorithm to return three distinctive scales corresponding to regions with large-, medium-, and small-sized squares. We select three coordinates highlighted in Figure 2.5, each one located at sufficiently distinct regions. Here the image is of size 256 by 256 with coordinates (60, 56), (204, 204) and (203, 76) for the selected large, medium, and small square regions respectively. We take the convention that  $(i, j)$  identifies the  $i$ -th column and  $j$ -th row. For ease of reference, let us denote the large, small, and medium square regions as  $P_1$ ,  $P_2$ , and  $P_3$  accordingly.

In Figure 2.6 we plot the discrete interval  $[0, 1, \dots, 255]$  against  $|Sf|$  for various values of  $|\alpha| < 1$  at the aforementioned locations. The numerical results coincide with the one dimensional example. We observe that the location of maxima for  $|Sf|$  in Figure 2.6 are indeed related to spatial features from Figure 2.5. In other words, maxima of  $|Sf|$  found at larger values of  $\tau$  correspond to regions with mainly geometric components, and similarly, for smaller values of  $\tau$  we can identify oscillatory regions. The influence of  $\alpha$  on  $Sf$  is also relevant and is the next item to be discussed.

In the left and middle column from Figure 2.6 we observe a clear separation of the scales which correspond to a large sized square region (coordinates of  $P_1$ ) and one belonging to a medium sized region (coordinates of  $P_3$ ), but this is not evident in some graphs corresponding to the small square region for  $\alpha < 0$ . It is only for  $0 < \alpha < 1$  that

## Binary Input Image $f$

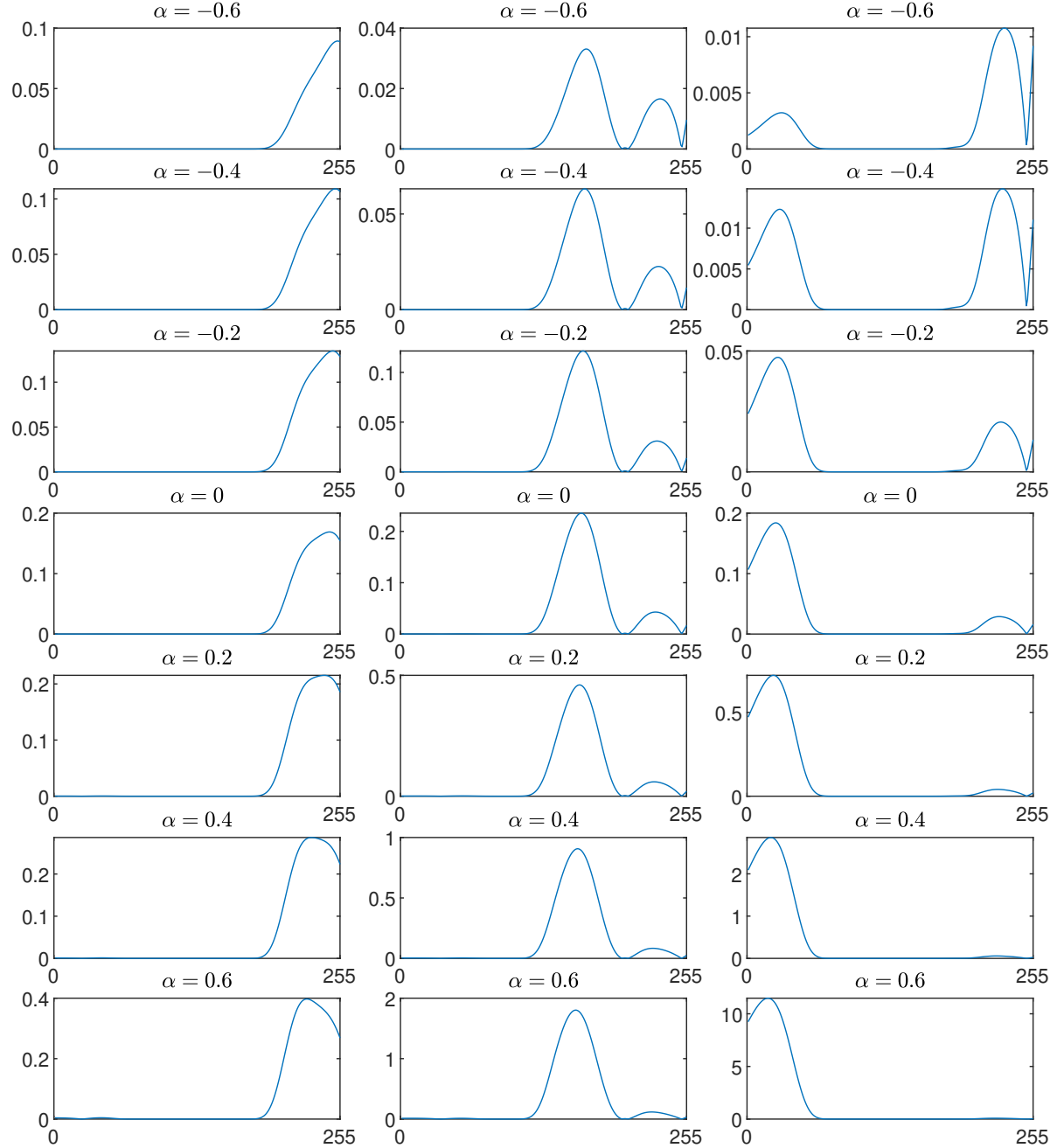


**Figure 2.5.** Binary image consisting of large, medium, and small squares with highlighted regions where  $|Sf|$  will be computed. Red dots are labeled  $P_1, P_2$  and  $P_3$  from top to bottom, left to right.

we detect a favorable scale for the small square region. Overall, we see numerical evidence that values of  $\alpha$  close to  $-1$  favor the detection of large scales. In contrast, choosing  $\alpha$  near  $1$  seems to support the detection of maxima for small scales. In the next chapter we turn our attention to applications dependant on appropriate choices of scales.

We conclude with the following key observations. Gaussian kernels with large variance in spatial coordinates capture features at coarse scales as demonstrated by the location of extrema for  $|Sf|$  pertaining to a large square region. Similarly, kernels with variance close to zero can identify scales corresponding to regions rich in texture, e.g.,  $|Sf|$  has extreme values at smaller scales when computed at coordinates for  $P_2$  selected from Figure 2.5. We have also presented curves showing local extrema at both small and large scales in Figure 2.5 and we note the existence of maxima can be emphasized or suppressed by modifying  $\alpha$  in accordance with different applications of scale-space tools.

### Illustration of $|Sf|$ for Various Values of $\alpha$



**Figure 2.6.** The oscillatory level,  $|Sf|$ , of  $f$  at large square region  $P_1$  (leftmost column), medium square region  $P_3$  (middle column), small square region  $P_2$  (rightmost column) for increasing values of  $\alpha$  (top to bottom). The vertical axis corresponds to  $|Sf|$  and the horizontal axis to the discrete variable  $\tau$ . See Figure 2.5 to identify regions.

## CHAPTER 3

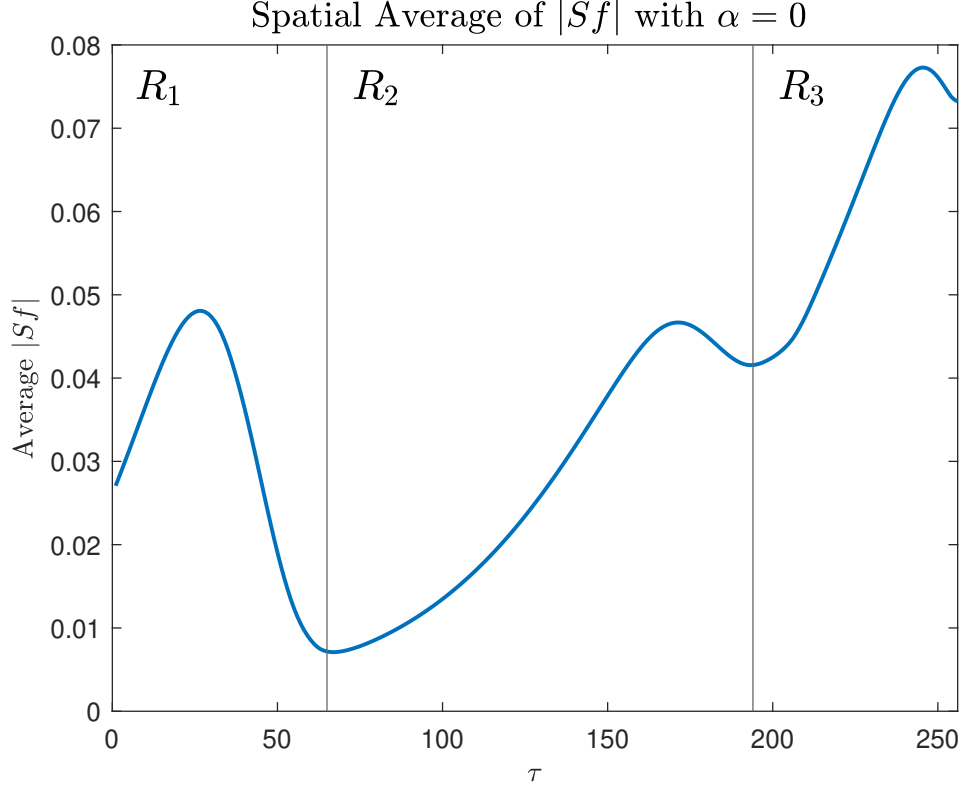
### LOCAL SCALES IN APPLICATIONS

In this chapter we pay closer attention to the practicality of scale-space representations. So far we have highlighted the importance of identifying key features in images for tasks such as data reduction and better understanding of visual processes in the human brain but it is no surprise there exist other contemporary applications which depend on feature extraction. To showcase the relevance of low level descriptors we state a collection of examples related to other research fields besides mathematics before presenting our own results. In marine biology one might be interested in coral reefs, wishing to construct a 3D model given only 2D images, say, to study coral growth patterns. In engineering a person responsible for developing autonomous vehicle navigation would desire a robot capable of detecting stop signs or to distinguish which light is active at a stoplight. Last, criminal justice practitioners can benefit from identifying felons based on facial features recognized by mobile devices in real-time. All the aforementioned applications have made use of feature detection as evidenced in [3, 7, 8]. With these examples in mind we hope to establish even more credibility in analyzing scale-space applications due to their interdisciplinary nature.

We keep progress with a focus on classification of image pixels by selecting global extrema of  $|Sf|$  curves, as well as a comparison of clustering against *k-means*. For a detailed survey of the latter method readers are referred to [15]. To conclude the chapter we demonstrate the influence of Gaussian blur on  $Sf$  as a hint for potential deblurring techniques.

#### 3.1 Clustering Image Regions Based on $|Sf|$ Similarity

Recall that the scale-space representation of an image yields insight into the regularity of  $f$  encoded locally, i.e., it is possible to obtain a sense of how oscillatory a region is by studying scales at a given coordinate  $x$ . In some cases, such as with Figure 2.5, it is intuitive which scale-space representations are appropriate. From Figure 2.6 we observe that indeed extrema of  $|Sf|$  are restricted to three regions, globally speaking. In other words by averaging a collection of  $|Sf|$  curves with respect to position  $x$ , denoted by  $|Sf|_{\text{avg}}$ , we notice three dominant bumps. To showcase this phenomenon we present Figure 3.1.

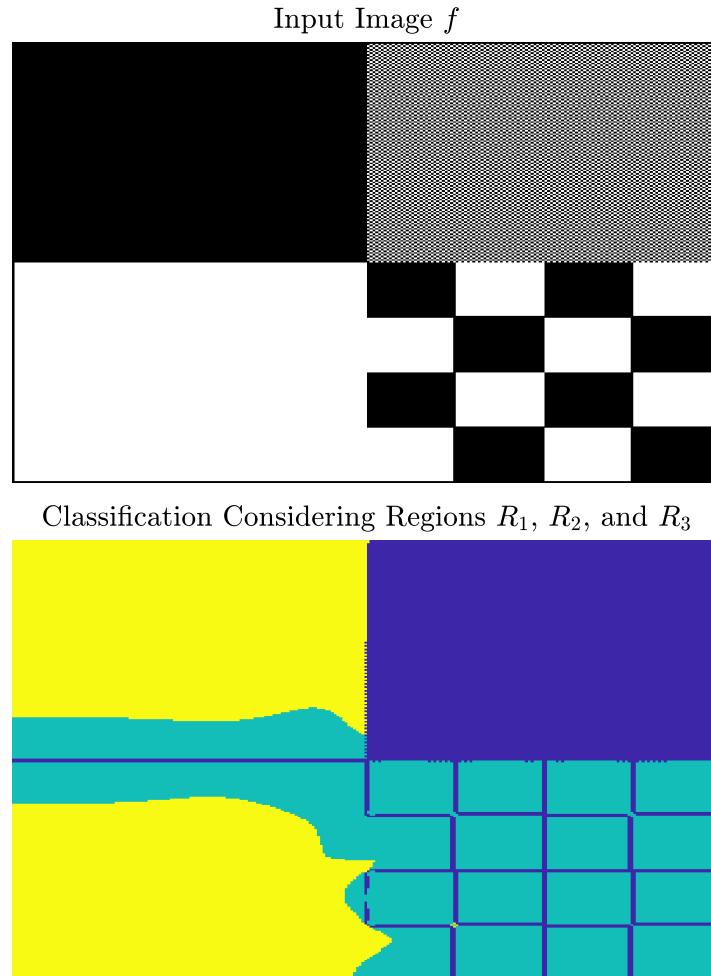


**Figure 3.1.** Average of all  $|Sf|$  curves with respect to position  $x$ . We separate three regions where  $|Sf|$  maxima is prevalent by placing a vertical line near local minima.

For reasons made clear soon we denote regions in Figure 3.1 by  $R_1$ ,  $R_2$  and  $R_3$  respectively from left to right. Since it is likely for extrema of  $|Sf|$  at a given coordinate to be found near maxima shown in Figure 3.1, we proceed by classifying pixels of  $f$  based on two methods. The goal here is to allow a machine to distinguish between coarse and oscillatory regions as well as a mix thereof. The first method, denoted  $R_i$  Method, involves a loop through all coordinates of  $f$  relating global maxima of  $|Sf|$  at a given location with a region  $R_1$ ,  $R_2$ , or  $R_3$ . For example say we take coordinates  $x^* = (100, 200)$  and compute  $|Sf(x^*, \tau)|$  with a global maximum at  $\tau = 35$ . Then we have  $x^*$  belongs to the class of  $R_1$ . If another pixel produces a global maximum of  $|Sf|$  at  $\tau = 240$  then that location is in the class of  $R_3$ , and so on. Secondly we turn our attention to the well-established algorithm, *k-means* clustering, an unsupervised learning technique allowing us to analyze similarity between  $|Sf|$  curves with greater amounts of clusters in contrast to only three from the first method. To initialize *k-means* we construct a matrix of  $256^2$  rows by 256 columns where rows represent coordinates of  $f$  whereas columns correspond to values of  $|Sf|$ . Such matrix

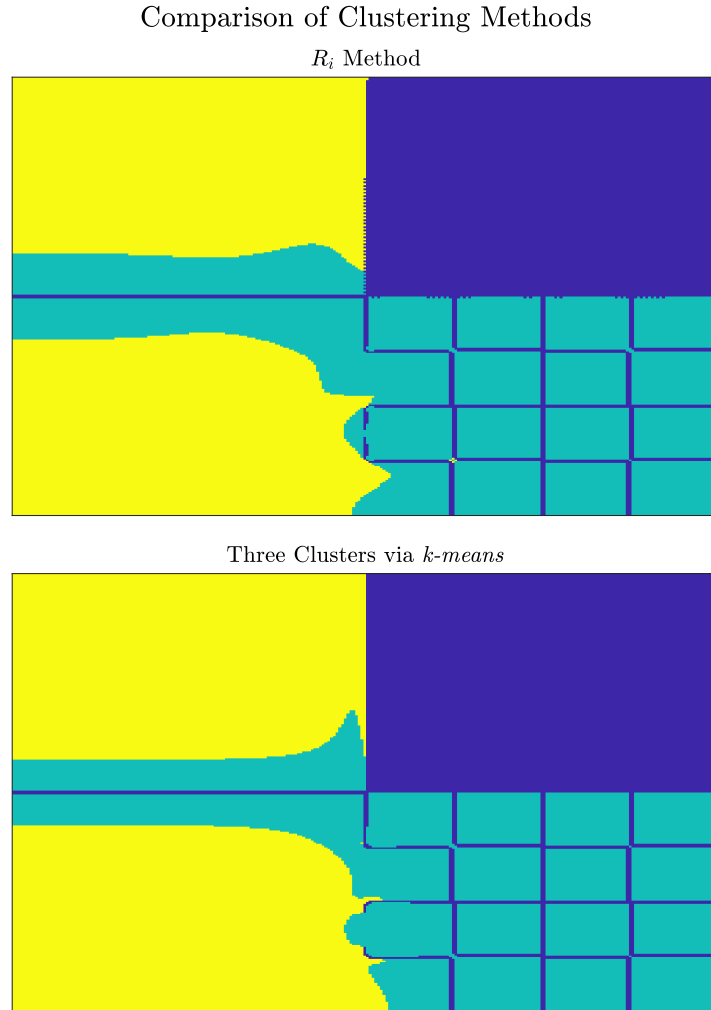
is our input data for the *kmeans* function in Matlab. For a detailed description of the algorithm see [13].

We make the following observations in accordance with Figure 3.2. In regard to assigning a class  $R_1$ ,  $R_2$  or  $R_3$  to the coordinates of  $f$  we have that, for the most part, neighboring pixels are equally categorized except those near a transition between  $|Sf|$  extrema. Since the amount of partitions is dependant on  $|Sf|_{\text{avg}}$ , this method might be better suited for those who wish to separate oscillatory and geometric components in an image up to a certain threshold for some value of  $\tau$ . In our example, such a separation would correspond to the act of combining the classes  $R_2$  and  $R_3$  to extract the geometric components in Figure 2.5. The remaining pixels of  $f$  would belong to  $R_1$  and can be considered as being in the oscillatory category.



**Figure 3.2.** Classifying all pixels (bottom) from Figure 2.5 (top) dependant on absolute maximum values of  $|Sf|$  with respect to regions  $R_1$  (dark blue),  $R_2$  (cyan), and  $R_3$  (yellow). We denote this technique the  $R_i$  Method.

The remaining clustering approach involving *k-means* yields comparable results to the  $R_i$  Method. There is one advantage evident when selecting a larger value of clusters: the transition between regions is smoother when clustering via *k-means*. In practice, choosing an appropriate value of clusters is non-trivial and can be task-dependant. A popular method of choosing the amount of clusters,  $k$ , is described in [31]. Figure 3.3 shows there is not a significant difference between the  $R_i$  Method and *k-means* for the case of three classes.



**Figure 3.3.** Side-by-side comparison of  $R_i$  (top) and *k-means* (bottom) clustering methods.

The motivation for a larger selection of clusters stems from graphs as shown in Figure 2.6 (right-most column, second plot from top to bottom) which have  $|Sf|$  extrema at both small and large scales simultaneously. By increasing the amount of

clusters in *k-means* we witness how  $|Sf|$  curves are classified dependant on transitions of scales as we move from homogeneous regions to edges evidenced in Figure 3.4.

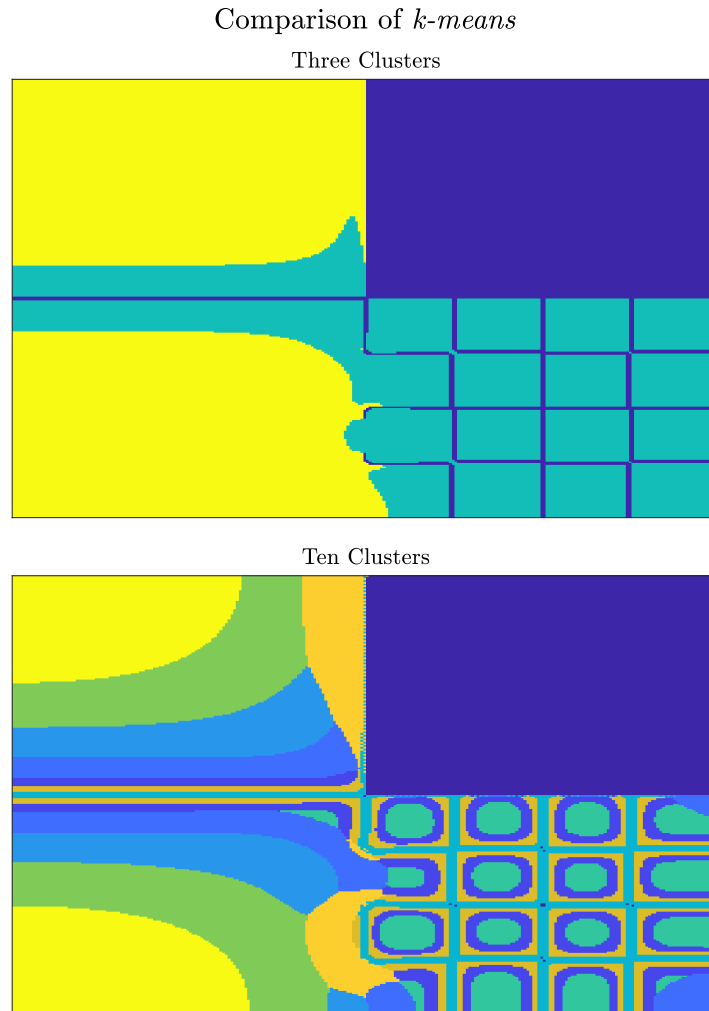


Figure 3.4. Side-by-side comparison of *k-means* with three (top) and ten (bottom) clusters.

### 3.2 On the Scale-Space Representation of $K_t * f$

In this section we present results that pertain to the Gaussian scale-space representation of a family of blurred images. An interesting aphorism taken from [23] helps us lay out the principles which we will build upon. Given two well-behaved functions  $f$  and  $g$ , in the sense of finite norms, we admit the following interpretation.

The convolution of  $g * f$  is at least as smooth a function as  $g$  and  $f$  are separately.



In other words, suppose we have a smooth function,  $g$ , and a non-smooth function,  $f$ , then we may construct smooth approximations of  $f$  via the convolution operation  $g * f$ . In our case we of course have  $f$  is an input image while  $g$  is the Gaussian kernel  $K_t$  with varying scale  $t > 0$ . Note that by increasing the scale parameter  $t$  we get smoother (more blurry) images  $K_t * f$  as shown in Figure 2.1. Given the information we have up to this point we can discuss the behavior of  $Sf$  if we replace  $f$  by  $K_t * f$ . Mathematically speaking, we wish to study elements of the sequence

$$\{S(f), S(K_{t_1} * f), S(K_{t_2} * f), \dots, S(K_{t_n} * f)\}, \quad t_1 < t_2 < \dots < t_n,$$

where the use of parenthesis emphasizes which image's scales we are examining.

We first give a 3D representation of the image we denote  $f$  from Figure 2.5 to help us visualize the smoothing effects of the convolution operation  $K_t * f$  for several  $t$ . Next, by increasing  $t$ , we showcase how  $K_t$  first influences the oscillatory region, then the region with medium-sized squares, and last the two regions with large squares. See Figures 3.6, 3.7, 3.8. The convolution between  $K_t$  and  $f$  was implemented in Matlab by directly constructing the kernel in the frequency domain using

$$\begin{aligned} \Gamma_1 &= \{-290, -280, -269, -258, -247\}, \\ \Gamma_2 &= \{-181, -166, -151, -136, -121\}, \\ \Gamma_3 &= \{-96, -81, -66, -51, -36\}, \end{aligned}$$

where  $\tau \in \Gamma_i$ ,  $i \in \{1, 2, 3\}$ , is related to the width of Gaussian kernel implemented to blur the small, medium, and large square regions respectively. Recall that  $\tau$  is useful for the discretization of  $Sf$  and the scale  $t$  is recovered via  $t = a^\tau$ ,  $a = (0.95)^{-1}$ . Within the context of probability, we have that  $t$  is the variance of the Gaussian kernel in question. In our numerical results we index  $\tau_j$  with  $1 \leq j \leq 5$  if  $\tau_j \in \Gamma_1$ ,  $6 \leq j \leq 10$  if  $\tau_j \in \Gamma_2$ , and  $11 \leq j \leq 15$  if  $\tau_j \in \Gamma_3$ . For example we have that  $\tau_9 = -96$  corresponds to  $t_9 = a^{\tau_9} \approx 0.007$ . We also take larger step sizes between values of  $\tau$  when blurring coarse regions so that the effect is noticeable.

Finally, we note that the problem of deblurring is ill-posed in a mathematical sense [25]. We may appreciate the sensitivity of noisy regions in  $f$  (small square region) to blurring since extrema of  $|Sf|$  at small scales can be suppressed by kernels which approximate the Dirac delta function. In Figure 3.6, we have that  $t_1 \approx 3 \cdot 10^{-7}$  and  $t_5 \approx 3 \cdot 10^{-6}$ . Then, in Figure 3.8 we choose  $t_{11} \approx 7 \cdot 10^{-3}$  and  $t_{15} \approx 1 \cdot 10^{-1}$ . These values provide evidence for our case: changes within one order of magnitude for  $t$  are enough to suppress small scales and notice a significant visual difference, while a change of two orders of magnitude are required for the case regarding coarse regions in  $f$ . This

numerical experiment showcases the challenges of reconstructing noisy data but also indicates a relationship between local scales and Gaussian blur which has potential to be exploited in localized deblurring algorithms. The value of  $\alpha = 0$  was considered for every computation of  $Sf$  in this section. The location at which  $Sf$  is calculated is unchanged from Figure 2.5.

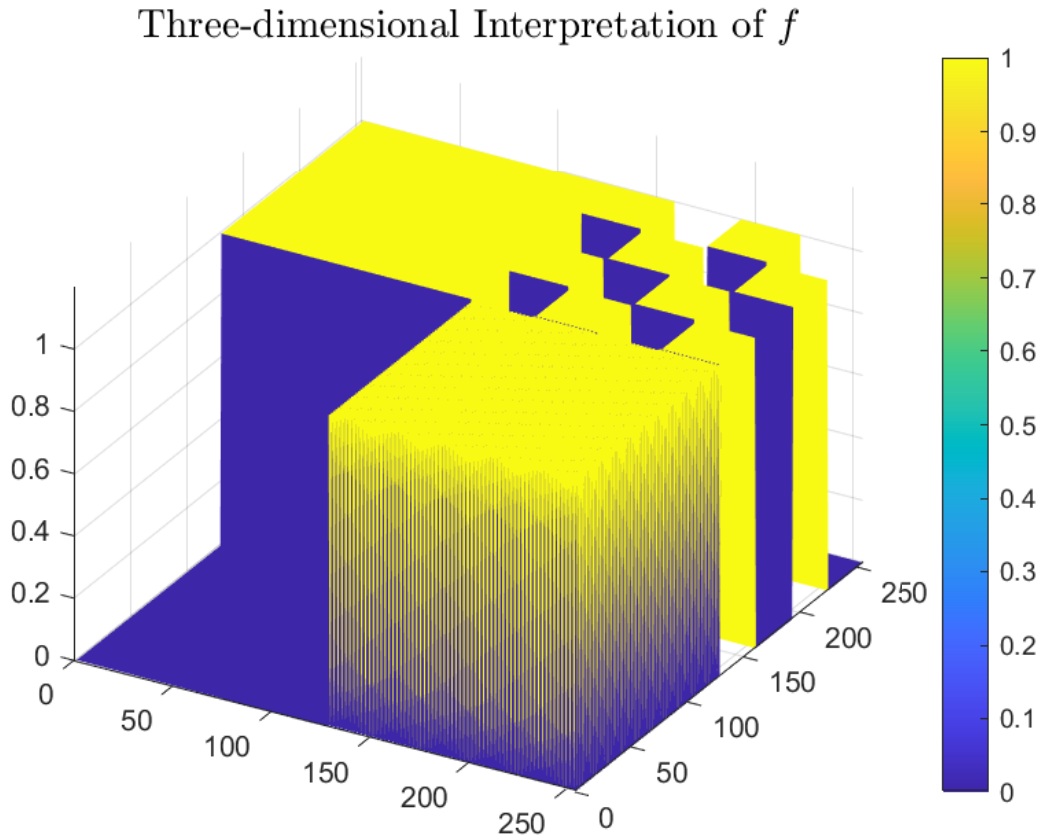


Figure 3.5. Graphing  $f$  using the *surf* command in Matlab. Note the image rotation to highlight the oscillatory region.

## Suppression of $|Sf|$ Extrema in Small Square Region

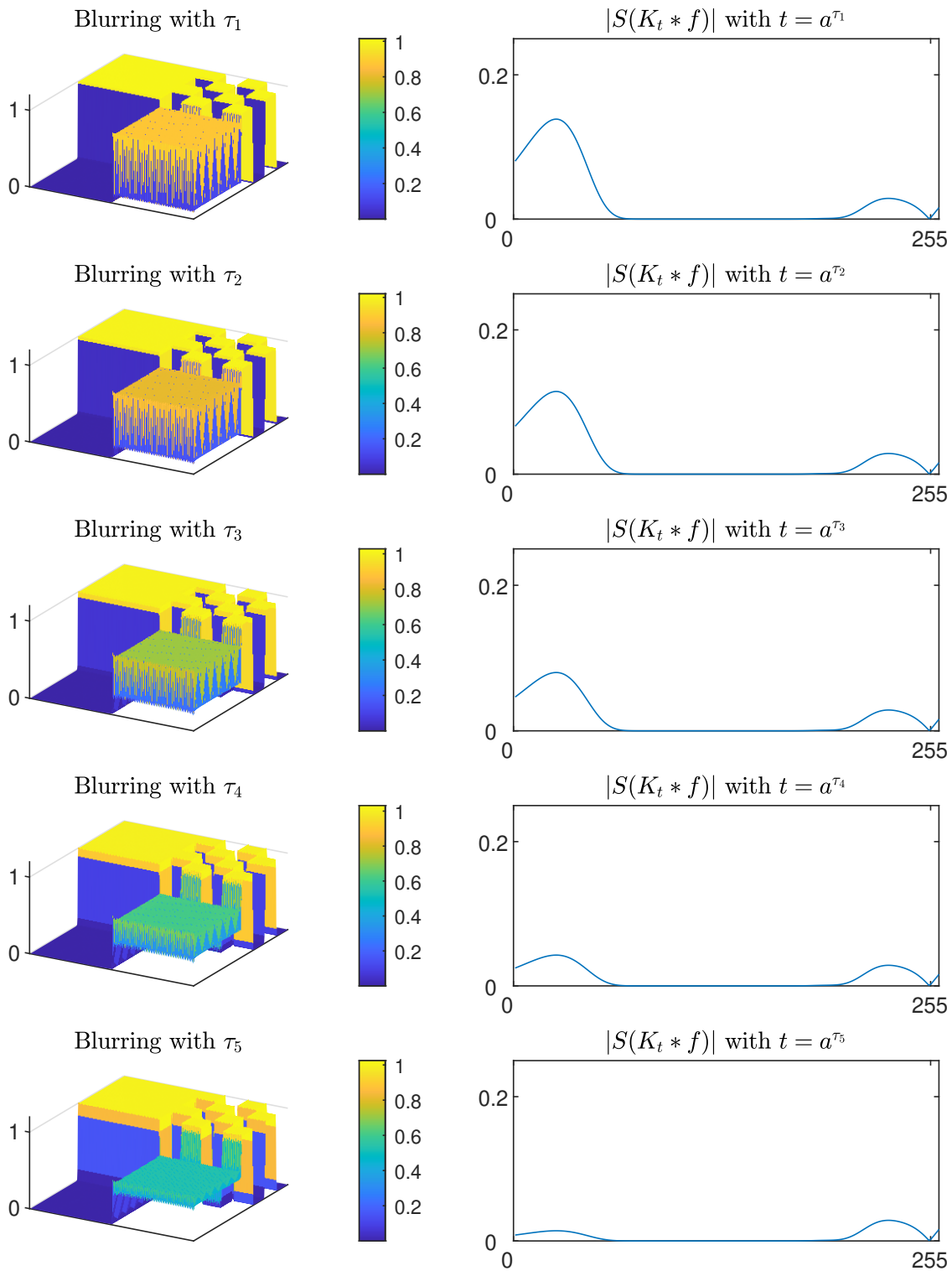


Figure 3.6. Influence of blurring evident in the small-sized squares region.

## Suppression of $|Sf|$ Extrema in Medium Square Region

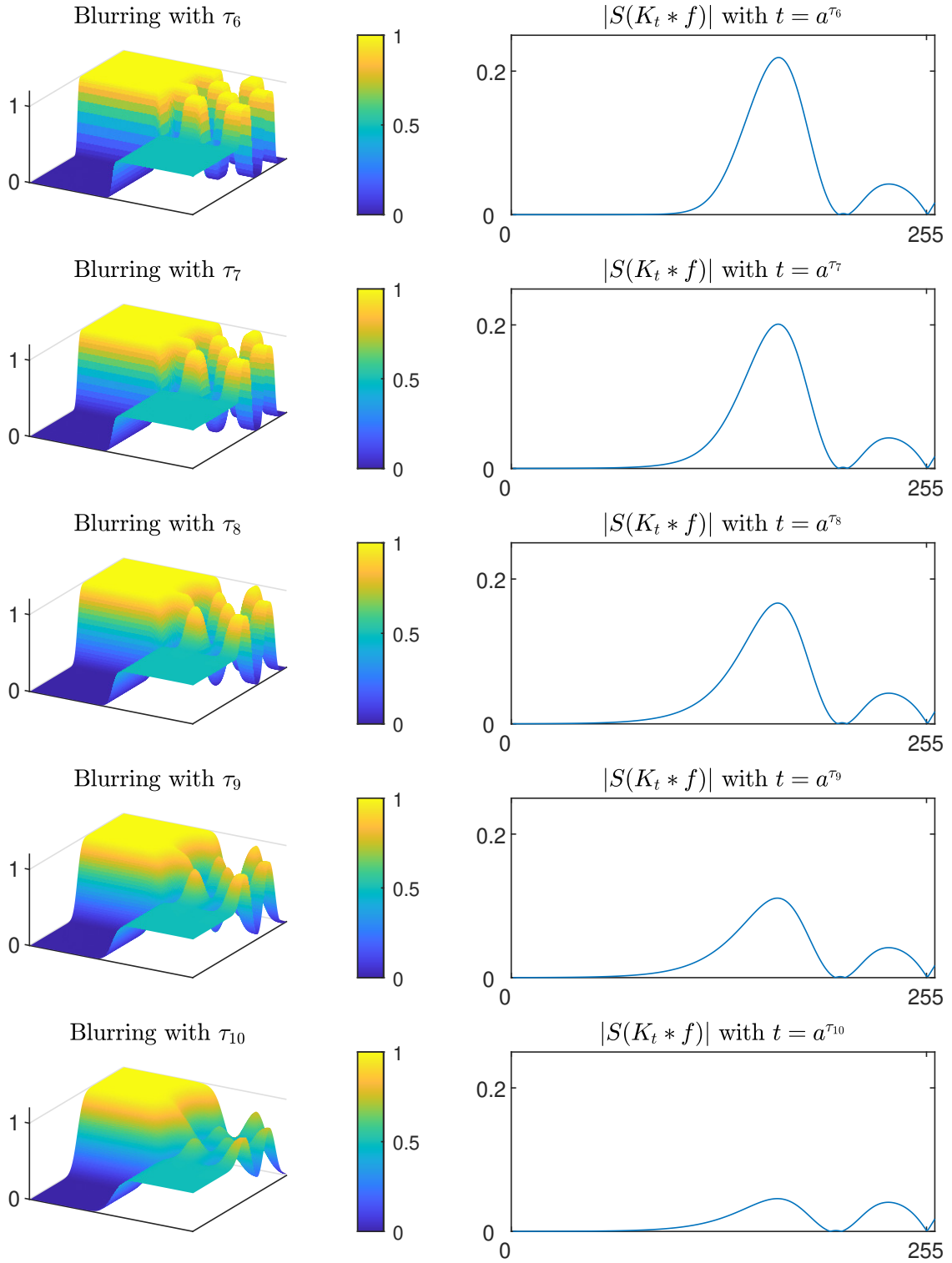


Figure 3.7. Influence of blurring evident in the medium-sized squares region.

## Suppression of $|Sf|$ Extrema in Large Square Region

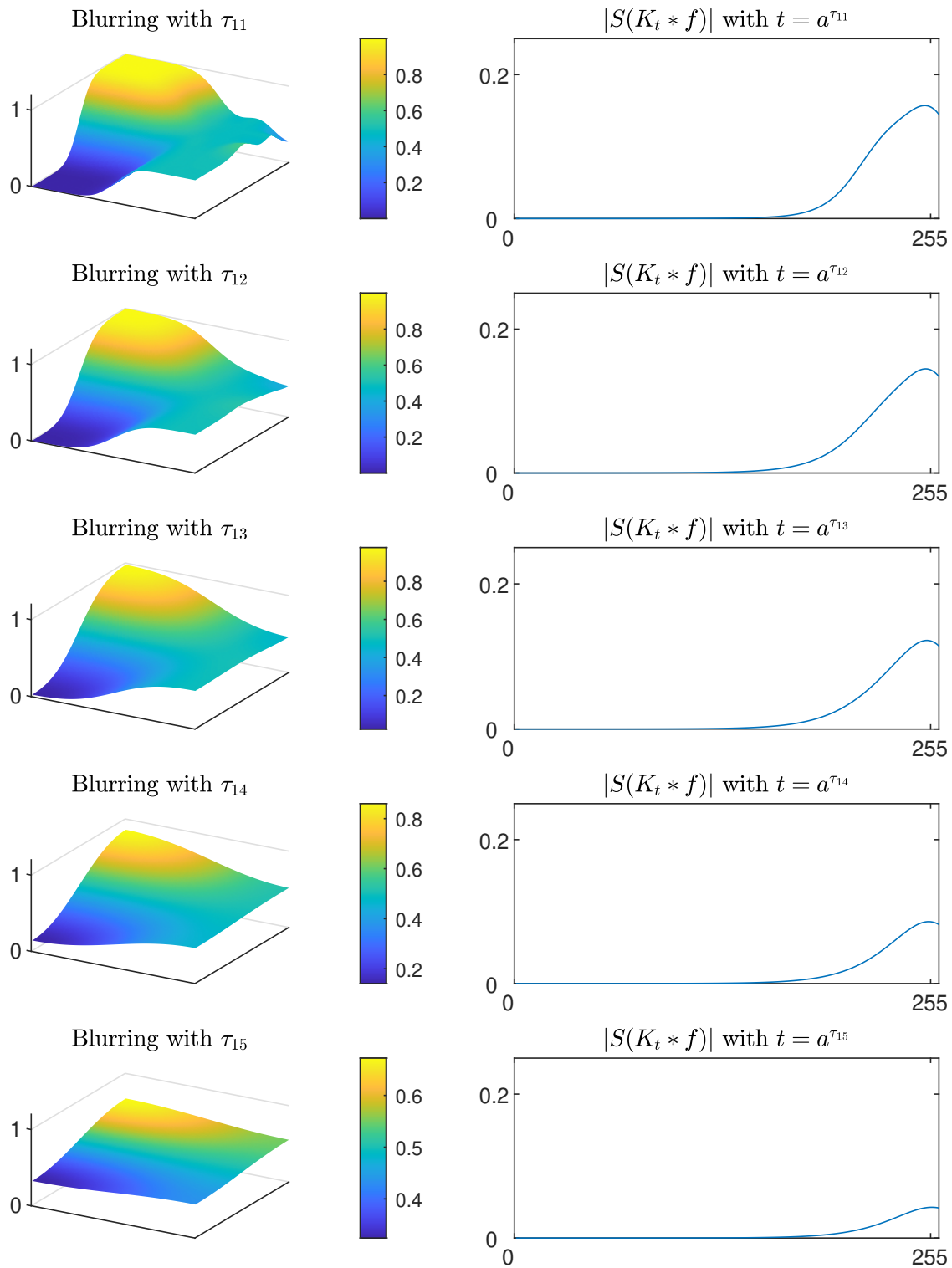


Figure 3.8. Influence of blurring evident in the large-sized squares region.

## CHAPTER 4

### GAUSSIAN ANISOTROPIC KERNEL OF INCREASING SCALE

We have previously established the scale-space representation of a given image,  $f$ , by implementing an isotropic Gaussian kernel,  $K_t$ , in Definition 2.1. The isotropic property is clear if we rewrite the expression for the chosen kernel in the Fourier domain

$$\mathcal{F}\{K_t\}(\xi) = \exp(-2\pi t \|\xi\|^2) = \exp(-2\pi (t\xi_1^2 + t\xi_2^2))$$

Again, using terminology from probability, the parameter  $t$  controls the variance of  $K_t$  in an isotropic manner along both dimensions  $\xi_1$  and  $\xi_2$ . Scenarios where drawbacks of the isotropic approach are evident can be studied in [4, 9, 12, 29]. In Chapter 4 we update our isotropic kernel to account for features dependant on scale as well as orientation. We also cite work where it is shown which scale-space properties are preserved by getting rid of the isotropic property belonging to  $K_t$  therefore justifying the search for a distinct kernel. For the numerical section, we present our selected test images and scale-selection results regarding the use of an anisotropic kernel.

#### 4.1 Preservation of Scale-Space Properties

We wish to break free from rotational invariance and are now motivated to decouple  $t$  into  $t'_1$  and  $t'_2$ , corresponding to the variance along  $\xi_1$  and  $\xi_2$  respectively, such that the ratio between scales is not unitary, i.e., without loss of generality

$$\frac{t'_2}{t'_1} = R \iff t'_2 = R t'_1, \quad R \in \mathbb{R}_+ \setminus \{1\}.$$

We then obtain our anisotropic kernel represented in the Fourier domain,  $\mathcal{F}\{\bar{K}_{t'_1}\}$ , where

$$\begin{aligned} \mathcal{F}\{\bar{K}_{t'_1}\}(\xi) &= \exp(-2\pi (t'_1 \xi_1^2 + t'_2 \xi_2^2)) \\ &= \exp(-2\pi (t'_1 \xi_1^2 + R t'_1 \xi_2^2)) \\ &= \exp(-2\pi t'_1 (\xi_1^2 + R \xi_2^2)). \end{aligned}$$

Interestingly, such a kernel is relevant in applied problems outside of scale-space theory, namely in the detection of corneal and lenticular astigmatism [33]. Here, blurred vision is a result of dominant curvature along a given direction in the cornea and lens analogous to a kernel with distinct variance values.

First, the selection of  $R$  corresponding to the anisotropy ratio is selected by considering the bound of  $3 + 2\sqrt{2}$  given in [19]. Secondly, we make use of a 2D rotation matrix allowing us to change coordinates from  $(\xi_1, \xi_2)$  to  $(\omega_1, \omega_2)$  and study local regularity with respect to a chosen angle,  $\theta$ . The updated coordinates expressed in terms of  $\omega(\theta) = (\omega_1(\theta), \omega_2(\theta))$  are given by

$$\omega(\theta) = \begin{pmatrix} \omega_1(\theta) \\ \omega_2(\theta) \end{pmatrix} = \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix}.$$

Since we will not discuss the isotropic kernel any further and will have control of  $t'_2$  by selecting  $t'_1$  as well as the orientation  $\theta$ , we relabel  $K_t$  for the rest of our work, having simplicity in mind. In other words we will implement the anisotropic kernel

$$\mathcal{F}\{K_{t,\theta}\}(\omega) = \exp(-2\pi t (\omega_1^2 + R\omega_2^2))$$

which leads us to an updated Equation 2.2 with an additional parameter  $\theta$ ,

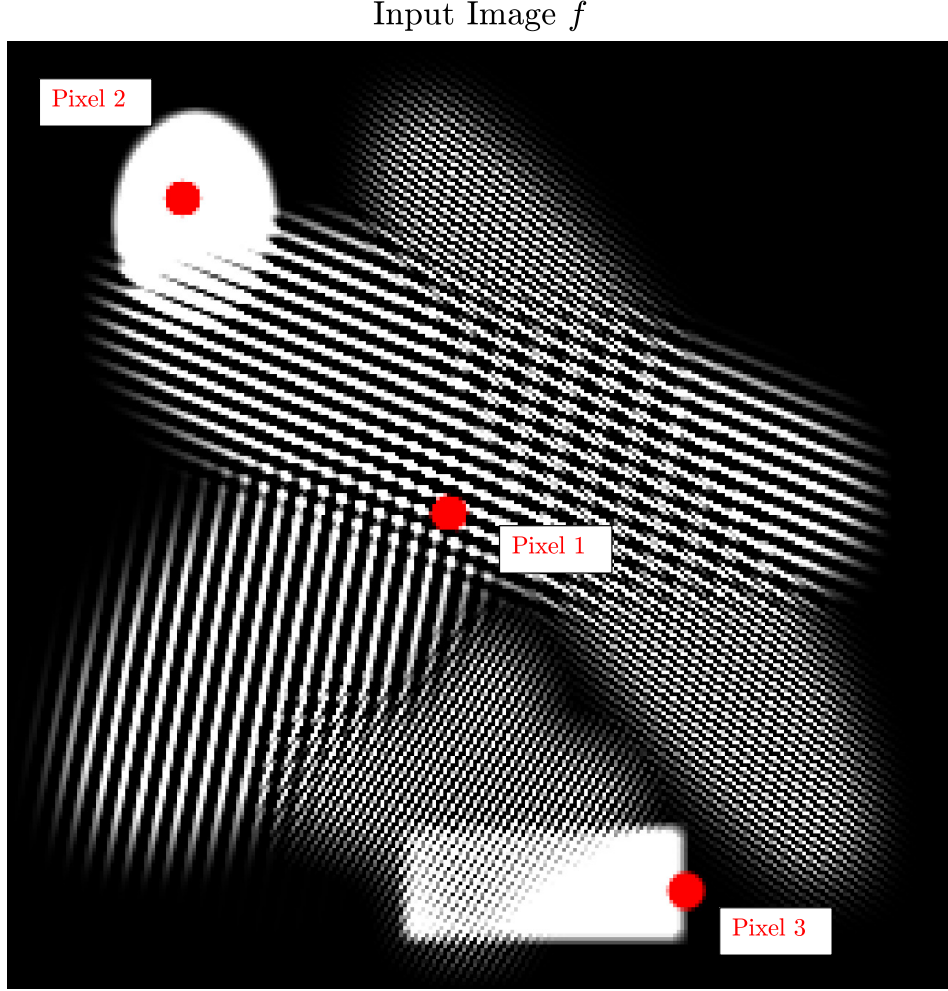
$$Sf(x, t, \theta) = t^{1-\frac{\alpha}{2}} \frac{\partial}{\partial t} K_{t,\theta}(x) * f(x). \quad (4.1)$$

## 4.2 Numerics Involving $Sf(x, t, \theta)$

To begin our numerical study we introduce the image  $f$  in Figure 4.1 where akin to Figure 2.5 we have highlighted the locations at which  $Sf(x, t, \theta)$  will be computed. We have previously computed  $Sf(x, t)$  in Figure 2.6 using an image in which square patterns repeat at distinct scales but such an example lacks features related to orientation which are evident in the general case of texture. Now, we present  $f$  constructed from several sinusoids of varying oscillation. Again,  $f$  is of size 256 by 256 with selected coordinates signaled by red circles at locations (128, 120), (43, 48) and (230, 184) which we denote Pixel 1, 2, and 3 respectively.

We now state our observations regarding Figures 4.2, 4.3, and 4.4. Globally speaking, the role of  $\alpha$  is unchanged from what was previously shown in Figure 2.6. We again have that negative values of  $\alpha$  emphasize scale detection for large values of  $\tau$  whilst positive values favor extrema in the neighborhood of  $\tau_0$ . We observe local behavior of  $Sf$  varies as expected in accordance with oscillatory and coarse regions. Our selected locations correspond to an oscillatory region (Pixel 1), another which is coarse (Pixel 2), and a position near an edge (Pixel 3).

Figure 4.2 shows the strongest responses are evident at two angles with values close to  $\frac{\pi}{4}$  and  $\frac{3\pi}{4}$  and scale  $\tau$  approximately 90. The computations of  $|Sf|$  with varying  $\alpha$  corresponds to the location Pixel 1 in Figure 4.1 where there is an intersection of ellipses of different orientation coinciding with the detection of extrema evident in  $|Sf|$ .



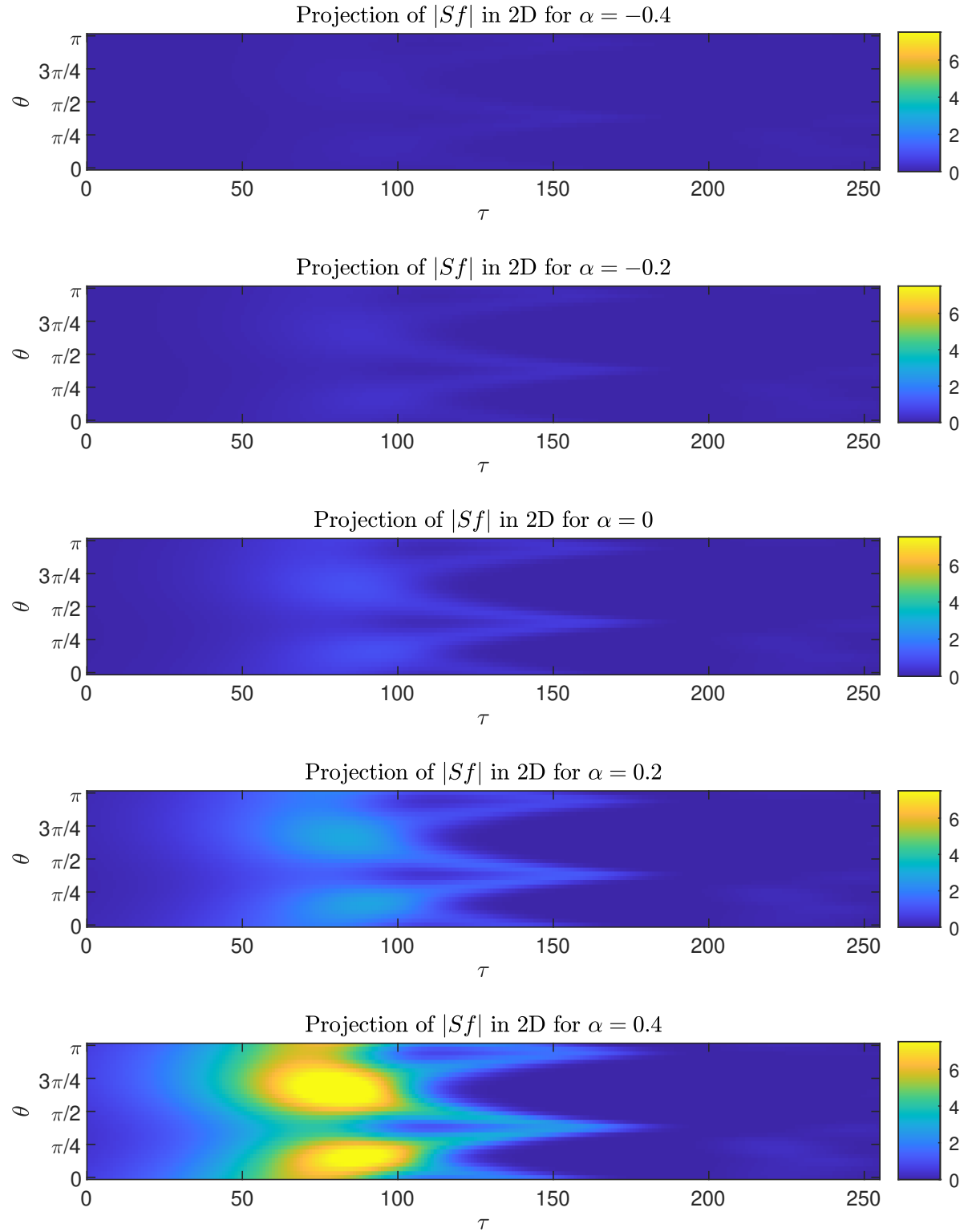
**Figure 4.1.** Image  $f$  with three highlighted pixels where  $|Sf(x, t, \theta)|$  is computed. From top to bottom we have assigned the labels Pixel 2, 1, and 3 to the locations of colored dots respectively.

Figure 4.3 corresponds to Pixel 2 where we observe  $|Sf|$  extrema is persistent at all angles in the interval  $[0, \pi)$  with scale  $\tau$  in the neighborhood of the value 200. Again, such results are to be expected in coarse regions akin to the leftmost column in Figure 2.6. It is noteworthy that all selected angles produce similar values of  $|Sf|$  due to the lack of orientation-related features in the selected location.

The response of  $|Sf|$  at an edge is promising in Figure 4.4 for values of  $\alpha < 0$ . This event relates the detection of a vertical edge with  $|Sf|$  extrema at an angle of  $\frac{\pi}{2}$  corresponding to Pixel 3. For values of  $\alpha > 0$  we conjecture smaller scales are favored at almost all angles due to the proximity to another coarse region (white rectangle in Figure 4.1).

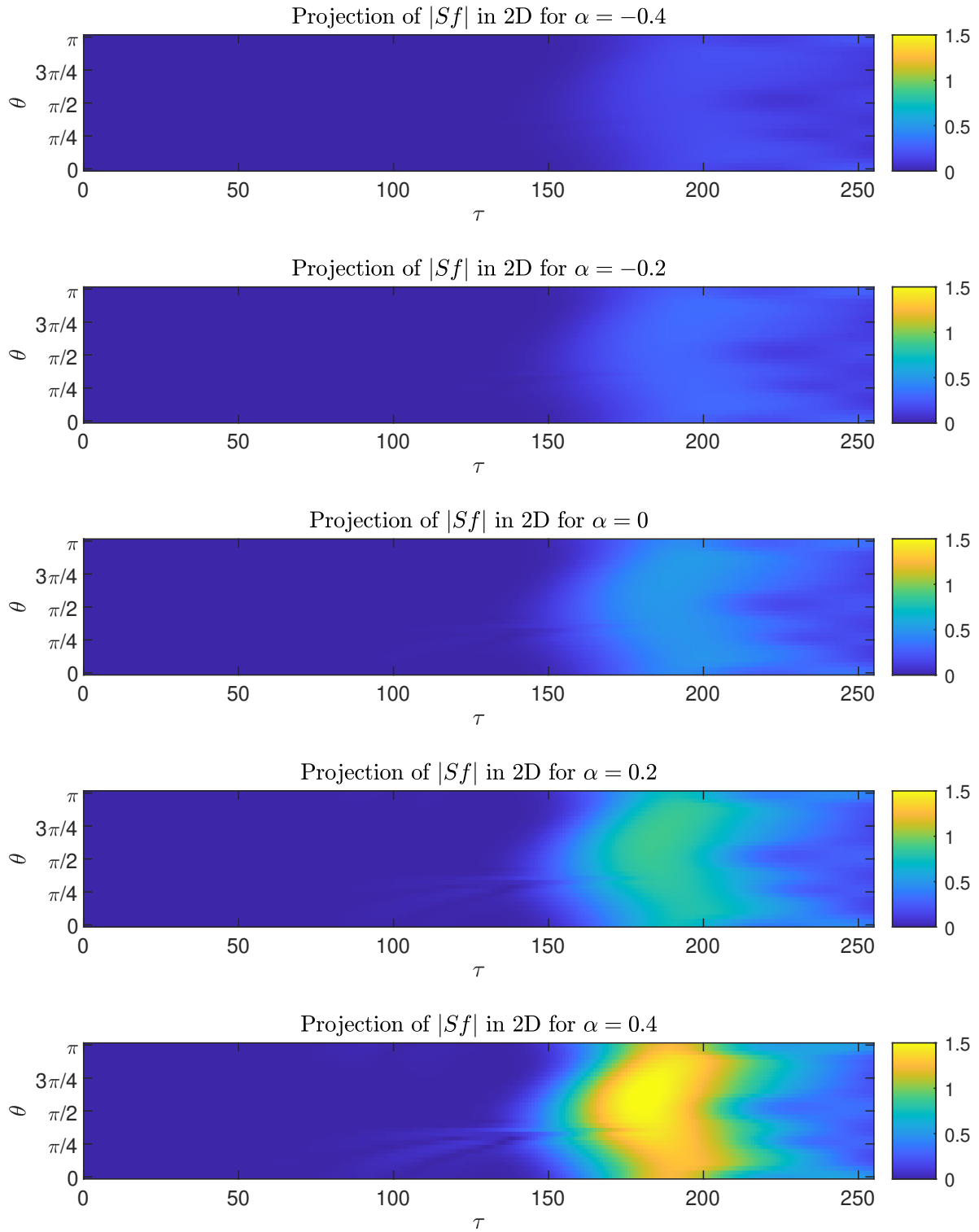


## Visualization of $|Sf|$ with Varying $\alpha$ Values



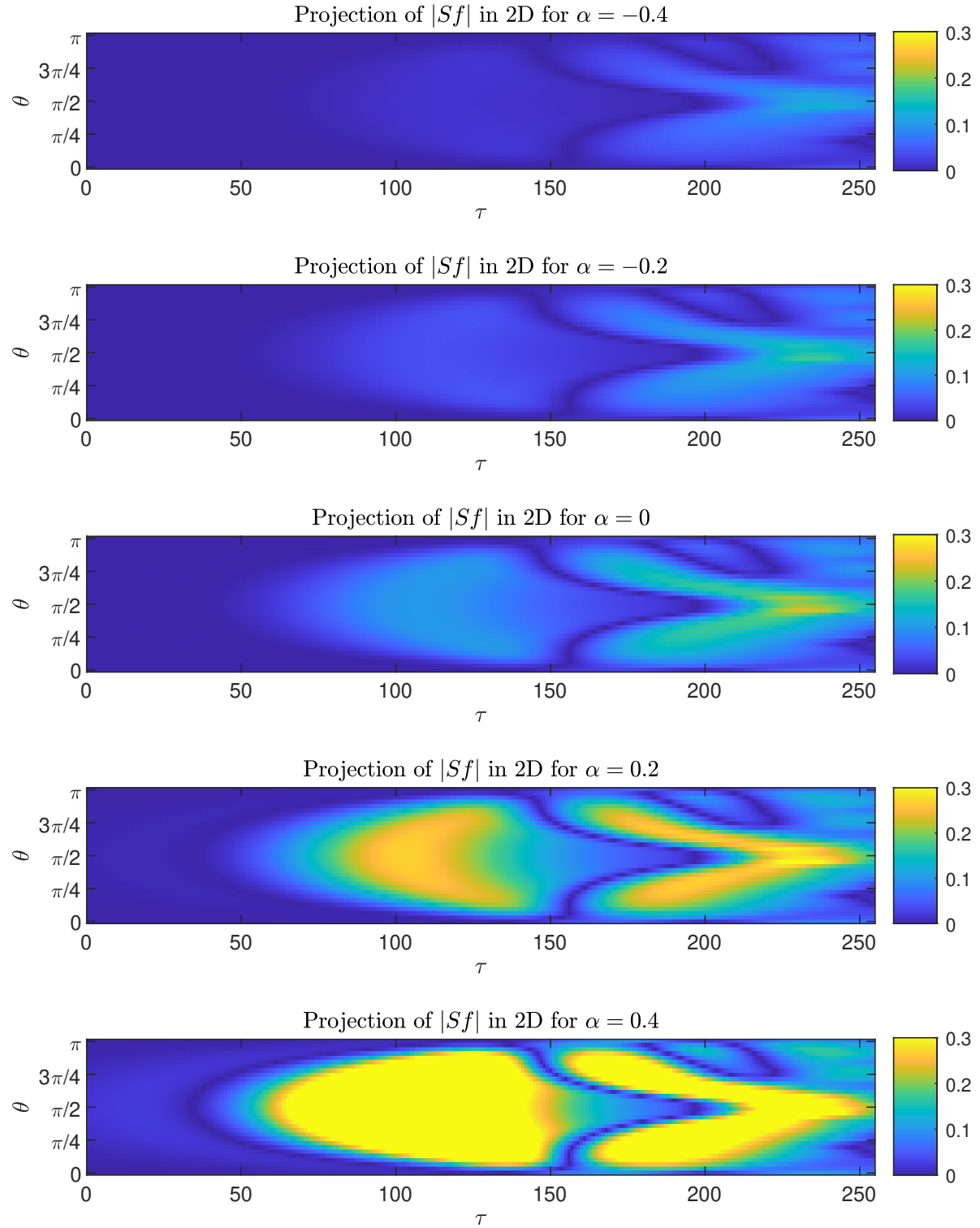
**Figure 4.2.** Computations of  $|Sf|$  at Pixel 1 with various values of  $\alpha$  showcasing extrema bias.

## Visualization of $|Sf|$ with Varying $\alpha$ Values



**Figure 4.3.** Computations of  $|Sf|$  at Pixel 2 with various values of  $\alpha$  showcasing extrema bias.

## Visualization of $|Sf|$ with Varying $\alpha$ Values



**Figure 4.4.** Computations of  $|Sf|$  at Pixel 3 with various values of  $\alpha$  showcasing extrema bias.

## CHAPTER 5

### FINAL REMARKS

Initially our work began with the analysis of local regularity in images. We have that Triebel-Lizorkin and Besov spaces are the most general in the sense that other well-known spaces are special cases, e.g., Sobolev, Hardy, and Hölder-Zygmund spaces [34], from which functional analysis tools may be implemented to better understand problems related to harmonic analysis and partial differential equations [2].

Throughout this work we have presented motivation from both a biology and computer vision point of view for studying localized behavior in images which hinges on a suitable selection of a kernel, in our case the Gaussian kernel  $K_t$ . We hope the link between seemingly unrelated topics is now evident and that the interdisciplinary study is meaningful if one wishes to tackle problems regarding the understanding of computer vision models, neural activity in the human visual cortex, as well as function spaces.

Our computations of both  $Sf(x, t)$  and  $Sf(x, t, \theta)$  presented in Figures 2.6, 4.2, 4.3, and 4.4 give numerical evidence for detection of notable low level descriptors where  $\alpha$  can be selected dependant on the application in mind. Although we discussed the influence of  $\alpha$  regarding extrema bias in  $Sf$  by providing several evaluations with varying  $\alpha$ , in practice this is not ideal. One can potentially explore an automated method for choosing  $\alpha$ . In the case of  $Sf(x, t)$ , a suggested approach is to calculate  $|Sf|_{\text{avg}}$  within a loop constantly updating  $\alpha$ . The purpose of the loop is to compute the total variation of  $|Sf|_{\text{avg}}$  and recover the value of  $\alpha$  which produces the maximum total variation. This method yields global information about prominent scales in an image but can also be localized within a sub-region of  $f$ .

To conclude we now describe potential future work. Regarding clustering algorithms, an exploration of distinct metrics might be fruitful. The distance metric utilized in Figure 3.3 is the squared Euclidian distance. Here one might also study scales by converting  $Sf$  into a vector of zeros and ones, where we place a one if a local maximum is evident, then the  $\ell^1$  or Hamming metrics may be of interest. Finally, due to the current popularity of machine learning localized scale selection may be useful to for data reduction regarding feature detection problems akin to [35].

## BIBLIOGRAPHY

- [1] R. ACHANTA, A. SHAJI, K. SMITH, A. LUCCHI, P. FUA, AND S. SÜSSTRUNK, *Slic superpixels compared to state-of-the-art superpixel methods*, IEEE transactions on pattern analysis and machine intelligence, 34 (2012), pp. 2274–2282.
- [2] D. R. ADAMS AND L. I. HEDBERG, *Function spaces and potential theory*, vol. 314, Springer Science & Business Media, 2012.
- [3] C. BEALL, B. J. LAWRENCE, V. ILA, AND F. DELLAERT, *3d reconstruction of underwater structures*, in 2010 IEEE/RSJ International Conference on Intelligent Robots and Systems, IEEE, 2010, pp. 4418–4423.
- [4] J. BIGÜN, G. H. GRANLUND, AND J. WIKLUND, *Multidimensional orientation estimation with applications to texture analysis and optical flow*, IEEE Transactions on Pattern Analysis & Machine Intelligence, 13 (1991), pp. 775–790.
- [5] T. BROX AND J. WEICKERT, *A tv flow based local scale measure for texture discrimination*, in European Conference on Computer Vision, Springer, 2004, pp. 578–590.
- [6] J. CANNY, *A computational approach to edge detection*, IEEE Transactions on pattern analysis and machine intelligence, (1986), pp. 679–698.
- [7] G. N. DESOUSA AND A. C. KAK, *Vision for mobile robot navigation: A survey*, IEEE transactions on pattern analysis and machine intelligence, 24 (2002), pp. 237–267.
- [8] L. A. ELREFAEI, A. ALHARTHI, H. ALAMOUDI, S. ALMUTAIRI, AND F. AL-RAMMAH, *Real-time face detection and tracking on mobile phones for criminal detection*, in 2017 2nd International Conference on Anti-Cyber Crimes (ICACC), IEEE, 2017, pp. 75–80.
- [9] J. GÅRDING AND T. LINDBERG, *Direct computation of shape cues using scale-adapted spatial derivative operators*, International Journal of Computer Vision, 17 (1996), pp. 163–191.
- [10] D. H. HUBEL AND T. N. WIESEL, *Receptive fields of single neurones in the cat's striate cortex*, The Journal of physiology, 148 (1959), p. 574.
- [11] P. W. JONES AND T. M. LE, *Local scales and multiscale image decompositions*, Applied and Computational Harmonic Analysis, 26 (2009), pp. 371–394.
- [12] S. N. KALITZIN, B. M. TER HAAR ROMENY, AND M. A. VIERGEVER, *Invertible orientation bundles on 2d scalar images*, in International Conference on Scale-Space Theories in Computer Vision, Springer, 1997, pp. 77–88.
- [13] T. KANUNGO, D. M. MOUNT, N. S. NETANYAHU, C. D. PIATKO, R. SILVERMAN, AND A. Y. WU, *An efficient k-means clustering algorithm:*

- Analysis and implementation*, IEEE transactions on pattern analysis and machine intelligence, 24 (2002), pp. 881–892.
- [14] T. LE AND F. MÉMOLI, *Local scales on curves and surfaces*, Applied and Computational Harmonic Analysis, 33 (2012), p. 3.
  - [15] A. LIKAS, N. VLASSIS, AND J. J. VERBEEK, *The global k-means clustering algorithm*, Pattern recognition, 36 (2003), pp. 451–461.
  - [16] T. LINDEBERG, *Detecting salient blob-like image structures and their scales with a scale-space primal sketch: A method for focus-of-attention*, International Journal of Computer Vision, 11 (1993), pp. 283–318.
  - [17] T. LINDEBERG, *Feature detection with automatic scale selection*, International journal of computer vision, 30 (1998), pp. 79–116.
  - [18] T. LINDEBERG, *Invariance of visual operations at the level of receptive fields*, PLoS One, 8 (2013), p. e66990.
  - [19] T. LINDEBERG, *Discrete approximations of affine gaussian receptive fields*, 2017.
  - [20] D. G. LOWE, *Distinctive image features from scale-invariant keypoints*, International journal of computer vision, 60 (2004), pp. 91–110.
  - [21] D. MARR, *Vision: A computational investigation into the human representation and processing of visual information*, MIT press, 2010.
  - [22] F. MERCIOL AND S. LEFÈVRE, *Fast image and video segmentation based on alpha-tree multiscale representation*, in 2012 Eighth International Conference on Signal Image Technology and Internet Based Systems, IEEE, 2012, pp. 336–342.
  - [23] B. G. OSGOOD, *Lectures on the Fourier transform and its applications*, vol. 33, American Mathematical Soc., 2019.
  - [24] S. A. PAPERT, *The summer vision project*, (1966).
  - [25] S. M. RIAD, *The deconvolution problem: An overview*, Proceedings of the IEEE, 74 (1986), pp. 82–85.
  - [26] L. G. ROBERTS, *Machine perception of three-dimensional solids*, PhD thesis, Massachusetts Institute of Technology, 1963.
  - [27] A. SERNA AND B. MARCOTEGUI, *Attribute controlled reconstruction and adaptive mathematical morphology*, in International Symposium on Mathematical Morphology and Its Applications to Signal and Image Processing, Springer, 2013, pp. 207–218.
  - [28] P. SOILLE, *Constrained connectivity for hierarchical image partitioning and simplification*, IEEE transactions on pattern analysis and machine intelligence, 30 (2008), pp. 1132–1145.
  - [29] C. STEGER, *An unbiased detector of curvilinear structures*, IEEE Transactions on pattern analysis and machine intelligence, 20 (1998), pp. 113–125.
  - [30] D. STRONG AND T. CHAN, *Edge-preserving and scale-dependent properties of*

*total variation regularization*, Inverse problems, 19 (2003), p. S165.

- [31] M. SYAKUR, B. KHOTIMAH, E. ROCHMAN, AND B. SATOTO, *Integration k-means clustering method and elbow method for identification of the best customer profile cluster*, in IOP Conference Series: Materials Science and Engineering, vol. 336, IOP Publishing, 2018, p. 012017.
- [32] H. TRIEBEL, *Theory of function spaces ii*, Bull. Amer. Math. Soc, 31 (1994), pp. 119–125.
- [33] M. VULOVIĆ, E. FRANKEN, R. B. RAVELLI, L. J. VAN VLIET, AND B. RIEGER, *Precise and unbiased estimation of astigmatism and defocus in transmission electron microscopy*, Ultramicroscopy, 116 (2012), pp. 115–134.
- [34] D. YANG AND W. YUAN, *A new class of function spaces connecting triebel–lizorkin spaces and  $q$  spaces*, Journal of Functional Analysis, 255 (2008), pp. 2760–2809.
- [35] M. ZHOU, J. SCOTT, B. CHAUDHURY, L. HALL, D. GOLDFOF, K. W. YEOM, M. IV, Y. OU, J. KALPATHY-CRAMER, S. NAPEL, ET AL., *Radiomics in brain tumor: image assessment, quantitative feature descriptors, and machine-learning approaches*, American Journal of Neuroradiology, 39 (2018), pp. 208–216.