Title

Physics-inspired Computational Imaging for Machine Vision, Drug Development and Cancer Immunotherapy

Permalink

https://escholarship.org/uc/item/3zq9f9mg

Author

Suthar, Madhuri

Publication Date

2021

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Physics-inspired Computational Imaging for Machine Vision, Drug Development and Cancer Immunotherapy

A dissertation submitted in partial satisfaction

of the requirements for the degree

Doctor of Philosophy in Electrical Engineering

by

Madhuri Suthar

2021

ABSTRACT OF THE DISSERTATION

Physics-inspired Computational Imaging for Machine Vision, Drug Development and Cancer Immunotherapy

by

Madhuri Suthar

Doctor of Philosophy in Electrical Engineering

University of California, Los Angeles, 2021

Professor Bahram Jalali, Chair

Traditional algorithms prevalent in computational imaging and signal processing are hand-crafted empirical rules synthesized to achieve a desired goal. In contrast, our approach is to craft qualitatively new algorithms by emulating laws of physics. Here, we show that Non-Linear Schrodinger Equation (NLSE), the master equation in optical physics can be exploited to invent a new class of computational imaging algorithms with best-in-class performance. We demonstrate a new contrast enhancement algorithm that is computationally efficient, achieves superior color gamut performance, and is able to support real-time video enhancement at 4K and 8K resolutions. We also show how the NLSE operator becomes an edge detection algorithm with exceptional performance in low light levels. In certain cases, these algorithms have the potential to be implemented in physical optics.

We demonstrate efficacy of these algorithms in solving a variety of problems for different real-world applications. Specifically, we have developed CytoLive, an award-winning real-time live cell tracking tool utilizing our NLSE-guided algorithms to analyze time-lapse microscopy videos acquired under low light conditions. This tool preserves inherent cell behavior by overcoming photo-toxicity and

photo-bleaching and has the potential for accelerating research in drug discovery.

Next, we discuss CytoEye, a cancer immunotherapy toolbox that mitigates the computational overload of analyzing giga-pixel sized pathology images of tumor micro-environment. Quantitative features extracted by this tool have the capability to predict whether or not patients respond to therapy – an important step toward personalized cancer immunotherapy.

The dissertation of Madhuri Suthar is approved.

Jonathan Kao

Fabien Scalzo

Vwani Roychowdhury

Ali Khademhosseini

Bahram Jalali, Committee Chair

University of California, Los Angeles

2021

To my family

for their unconditional love, support and encouragement

LIST OF FIGURES

xxiv

List of Tables

xxxii

xxxiii

from the lab members, especially Zhouya Bai, Shuqian Sun, Tingyi Zhou, Yiming Zhou and Callen MacPhee who have also supported me in their own ways, for which I am very grateful.

I also had many great mentors over the course of my undergraduate studies and project training who led me to this Ph.D. in Electrical Engineering. In particular, I am very grateful of Prof. Vishnu Priye, IIT-ISM Dhanbad and Prof. Shalabh Gupta, IIT Bombay.

Finally and most importantly, I am grateful to my parents, Dr. Indira Prabhakar, Dr. Om Prakash Suthar and my brother Dr. Chitranjan Suthar for being supportive of my scientific and academic endeavors since forever. I am also grateful to my husband Abhinav Agarwal for always having my back and lending a sympathetic ear. I would also like to express my deep gratitude and appreciation to my entire family, my extended family Prof. Ashok Kumar Agarwal, Uma Agarwal, Abhishek Agarwal, and my friends Ankita, Shreya and Akankshya for their unconditional love, support, and encouragement.

Vita

2010–2014	Bachelor of Technology (B.Tech), Electronics and Communication Engineering, IIT-ISM Dhanbad, India.
2014–2015	Research Assistant, Department of Electrical Engineering, IIT Bombay, India.
2015–2016	Master's of Science (M.S.), Electrical Engineering, UCLA, USA.
2016–present	Graduate Student Researcher, Department of Electrical and Computer Engineering, UCLA, USA.
Jun-Dec 2019	Image Processing Intern, Dolby Laboratories Inc., Sunnyvale, USA.

Publications

M. Suthar, B. Jalali, "Nonlinear Schrodinger Framework for Image and Video Enhancement," 2021 (Journal, in preparation)

M. Suthar, H. Asghari and B. Jalali, Feature enhancement in visually impaired images, IEEE Access 6, pp. 1407-1415, 2018 (Journal, Published)

M. Suthar and B. Jalali, Phase-Stretch Adaptive Gradient-Field Extractor (PAGE), Coding Theory. IntechOpen, 2020. (Book Chapter, Published).

B. Jalali, **M. Suthar**, H. Asghari and A. Mahjoubfar, Physics-Based Feature Engineering, SPRINGER, 2019 (Book Chapter, Published).

M. Suthar, B. Jalali, "Physics-inspired Image and Video Enhancement," ICCV 2021 (Conference, under review).

M. Suthar, B. Jalali, Natural algorithm for image and video enhancement, AI and Optical Data Science II, Paper 11703-38, SPIE Photonics West, 2021 (Conference, Published).

M. Suthar, B. Jalali, Optical Computing of Mathematical Derivatives using Dispersion and Coherent Detection. In IEEE Conference on Lasers and Electro-Optics Europe European Quantum Electronics Conference (CLEO/Europe-EQEC), 2019 (Conference, Published).

M. Suthar, Z. Bai, T. S. Nowicki, A. Ribas, C. K. Lonappan and B. Jalali, AI Curated Digital Pathology for Cancer Immunotherapy, SPIE Medical Imaging (2019) (Conference, Published).

M. Suthar, J. Mendelson, A. Mahjoubfar, C. K . Lonappan, B. Jalali, Natural computational imaging, Optical Data Science: Trends Shaping the Future of Photonics I, Paper 10551-16, SPIE Photonics West, 2018 (Conference, Published).

M. Suthar, J. Mendelson, A. Mahjoubfar, C. K . Lonappan, K. Roy, S. Mitchell, A. Hoffmann, B. Jalali, Automated hyper-dimensional live cell tracking, Imaging, Manipulation, and Analysis of Biomolecules, Cells, and Tissues XVI, Paper 10497-21, SPIE Photonics West (2018) (Conference, Published).

B. Jalali, **M. Suthar**, H. Asghari and A. Mahjoubfar, Optics-Inspired Computing, In Proceed- ings of the 5th International Conference on Photonics, Optics and Laser Technology - Volume 1: PHOTOPTICS, 2017 (Conference, Published).

M. Suthar, A. Mahjoubfar, K. Seals, E. W. Lee and B. Jalali, Diagnostic tool for pneumothorax, 2016 IEEE Photonics Society Summer Topical Meeting Series (SUM), Newport Beach, CA, pp. 218-219. doi: 10.1109/PHOSST.2016.7548806, 2016 (Conference, Published).

B. Jalali, **M. Suthar**, "Nonlinear Schrodinger Framework for Image and Video Enhancement," World and U.S. Patent (Applied), 2021.

CHAPTER 1

Introduction

1.1 Motivation

Every day, the world creates nearly 2.5 Exabyte (10^{18} bytes) of data. Surprisingly, 90% of the data present in the world today has been created in the last two years alone highlighting the exponential rise in the amount of digital data [31]. To process this enormous data, the field of electronics has experienced a "digital revolution" leading to development of high performance digital signal processors (DSPs). However, a large fraction of this digital data is generated by translating real world analog signals into their digital representations using electronic Analog-to-Digital Converters (ADCs). These electronic ADCs cannot capture fast waveforms due to limited temporal resolution as well as impose restrictions on real time operations owing to their slower speed. Additionally, processing massive amounts of data in data centers accounts for 50-60% of their electricity budget and a rapidly growing fraction of total electricity consumption.

This calls for development of new computing technologies that offer speed, energy efficiency and ease of implementation. Fortunately, nature and in particular, photonics can provide a solution to certain class of problems. We believe that inspirations from nature need not be limited to design of physical machines but should be extended to creation of new computational algorithms. We expect this new paradigm to lead to a new class of algorithms that are energy efficient while providing unprecedented functionality.

Traditional algorithms prevalent in computational imaging and signal processing are hand-crafted empirical rules synthesized to achieve a desired goal. In contrast, our approach is to design algorithms by emulating physical laws of nature. As such, the central theme of this research work has been to develop novel computational imaging algorithms that have roots in physical optics, i.e. these algorithms emulate physical phenomenon. This results in a new class of computational imaging algorithms that could never have been imagined using conventional techniques. Although the equations governing these physical phenomenon are complex representations but fortunately there are well established approximations, and as such the implemented algorithms are computationally fast. Additionally, in principle these physics inspired algorithms emulating physical phenomenon can be implemented on optical systems, thereby enabling computation at the speed of light. In section 1.2, we first present a brief background on the evolution of this research work and it connection to optics. And in section 1.3, we discuss an overview of the this dissertation.
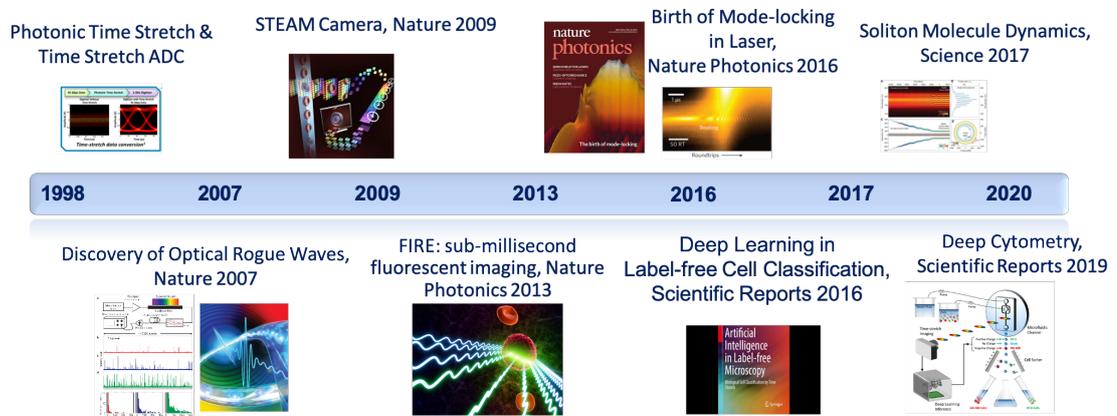
Figure 1.1: **Photonic Time Stretch**: A high throughput real time instrumentation technology that employs dispersion to slow down an analog signal in time and has been employed widely to slow down ultra-fast phenomena. Picture insets courtesy of [1–10]

1.2 Roots in Physics

Photonic time stretch, a temporal signal processing technology, employs group-velocity delay (GVD) dispersion to slow down an analog signal in time and thereby, compresses the bandwidth of an analog signal allowing digital processing of fast waveforms which are otherwise not supported by slower electronic ADCs [1, 32, 33]. This method also known as the time-stretch dispersive Fourier transform (TS-DFT) has been the most successful solution to solve the critical problems associated with electronic ADCs in terms of temporal resolution as well as dynamic range and has been instrumental in the development of highest performance ADCs [34]. Time stretch technology has been employed widely for real time acquisition of ultrafast signals as shown in Figure 1.1 but the most remarkable application of this technology has been to study the relativistic electron structure [4].

Time stretch spectrometer is an extension of time stretch technology for high throughput single-shot spectroscopy that has led to observation of non-repetitive ultrafast events in optical systems such as optical rogue waves [3], soliton molecules [5] and mode locking in lasers [7]. Similarly, time stretch camera [2], a MHz-frame-rate bright-field imager, uses amplified dispersive Fourier transform to analyze images and has enabled the detection of rare cancer cells in blood with false positive rate of one cell in a million [35]. By integrating artificial intelligence with time stretch imaging, label-free cancer cell detection with record accuracy has been achieved [8, 9]. These high throughput imaging systems generate a torrent of data which can be compressed by foveated sampling using warped stretch that exploits signal sparsity [36].

The master Non-Linear Schrodinger Equation (NLSE) equation that describes the physics behind the pulse propagation in an optical medium [37] is represented as:

$$\frac{\partial A}{\partial z} + \frac{\alpha}{2}A + i\frac{\beta_2}{2}\frac{\partial^2 A}{\partial T^2} - \frac{\beta_3}{6}\frac{\partial^3 A}{\partial T^3} = i\gamma|A|^2 A \qquad (1.1)$$

where A is the input signal that is function of time T and coefficients alpha α, beta β_2 and gamma γ are real numerical parameters standing for loss, group velocity dispersion, self-phase modulation, respectively. As we will show later in the text that it is also this master equation that is used to craft qualitatively new computational imaging algorithms by exploring the properties of physical laws of nature. Particularly, we will demonstrate a new computationally efficient contrast enhancement algorithm and an edge detection algorithm with exceptional performance in low light levels.

1.3 Dissertation Overview

In the first part of this thesis, we describe how Non-Linear Schrodinger Equation (NLSE) can be exploited to invent a new class of computational imaging algorithms with best-in-class performance and in the second part, we discuss application of these algorithms to solve real world imaging problems. In Chapter 2, we show that the Non-Linear Schrodinger Equation, the master equation used in optics, plasma physics and a number of other disciplines, has useful properties for contrast enhancement of digital images, demonstrate its efficacy and benchmark it against popular state-of-the-art algorithms. We also show how the NLSE operation can be extended for edge detection in digital images and suggest other interesting algorithms.

One of the major open problems in computational imaging is feature detection in visually impaired images. In Chapter 3, we describe a potential solution using Phase Stretch Transform (PST), a new edge detection method that is inspired by the physics of the photonic time stretch technique, a real time data acquisition technology. PST emulates 2D propagation through a medium with group velocity dispersion, followed by coherent (phase) detection. The algorithm performs exceptionally well as an edge and texture extractor. We mathematically derive the

intrinsic nonlinear transfer function and demonstrate how PST has an inherent equalization ability resulting in high dynamic range of operation for feature detection in visually impaired images. We also propose a hybrid method by combining our method with the conventional techniques which further improves the dynamic range. Additionally, our results also propose a new optical computing paradigm for the computation of mathematical derivatives via dispersion.

Emulated by an algorithm, certain physical phenomena have useful properties that can be exploited for image transformation. For example, image denoising can be achieved by emulating the propagation of an image through the heat diffusion equation. Similarly, by emulating propagation of an image through a birefringent diffractive propagation, the original image can be embedded into a set of feature maps that select semantic information at different scale, orientation, and spatial frequency. This decomposition method, which we term as Phase-stretch Adaptive Gradient-field Extractor (PAGE) is introduced in Chapter 4. We also demonstrate applications of this algorithm in edge detection and extraction of semantic information from medical images as well as electron microscopy images.

In Chapter 5, we demonstrate efficacy of the above mentioned physics-inspired computational imaging algorithms in solving a variety of problems for different real-world applications. Specifically, we demonstrate two software tools, 1) Cyto-Live, an award-winning real-time live cell tracking tool used to accelerate Drug Discovery and 2) CytoEye, a high throughput pathology image analysis tool for use cancer immunotherapy.

Finally, in Chapter 6 we conclude with our findings on developing novel computational imaging algorithms that have roots in physics- a fundamentally different approach to craft algorithms. We also present ideas for potential physical implementations of these algorithms as part of future works. The latest version of the codes relevant to this research work can be found at the official Jalali Lab Github page [38].

CHAPTER 2

Nonlinear Schrodinger Framework for Image and Video Enhancement

Non-Linear Schrodinger Equation is the master equation used in optics, plasma physics and a number of other disciplines. In this Chapter, we show that the equation has useful properties for contrast enhancement of digital images, demonstrate its efficacy and benchmark it against popular state-of-the-art algorithms. Two implementations are reported, one with fixed parameters and the other with parameters learnt by a neural network in an unsupervised fashion. We demonstrate that the algorithm achieves excellent color gamut performance, high computational efficiency and is able to support 8K video at 100 FPS frame rate on a NVIDIA Titan X GPU. Finally, we also show how the NLSE operation can be extended for edge detection in digital images.

2.1 Motivation

The smartphone photography is on a meteoric rise as evident from the fact that people all around the globe took more than 1.4 trillion digital photos and videos in the year 2020 alone mostly using their smart phones [39]. Typically smart phone captured photos suffer from impairments such as non-uniform illumination, low contrast and noise due to the limitation of the camera's optics or due to sub-optimal lighting conditions. Poor ambient lighting conditions pose a major challenge for smart phone camera photography due to the small camera aperture

6
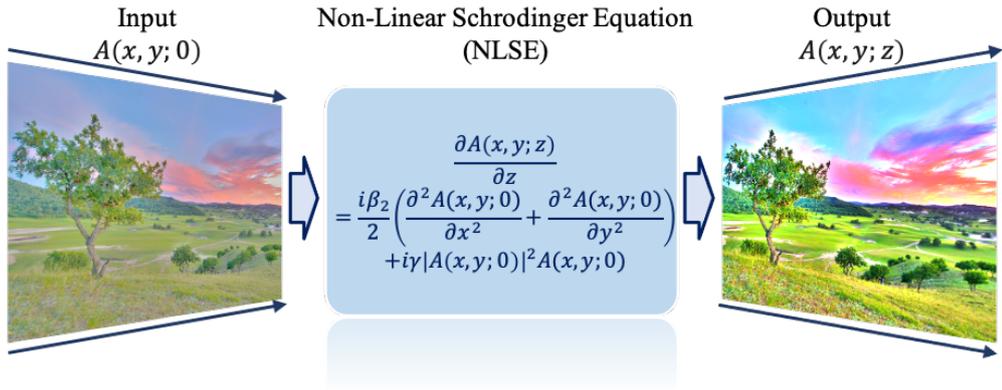
Figure 2.1: **Conceptual description of the approach behind Nonlinear Schrodinger Framework for <u>I</u>mage and <u>V</u>ideo <u>E</u>nhancement (SIVE)**. The input is transformed by propagation through the Nonlinear Schrodinger Equation (NLSE), the master equation used in optics, plasma physics and a number of other disciplines. Under certain conditions, described mathematically in this Chapter, the contrast of the image is enhanced. We show that this transformation is numerically efficient.

which limits the number of photons reaching the image sensor resulting in a low signal-to-noise ratio (SNR).

In extreme low-light conditions, SNR decreases to a point that the sharpness function extracted from the high frequency content of the captured image gets flattened out [40]. This makes it difficult for the camera's passive auto-focus (AF) to correctly focus thus degrading the visual aesthetics of the captured image as well as the performance of any subsequent machine vision tasks [41].

The low SNR problem further escalates in case of video acquisition where the frame rate can be high and therefore, the exposure time per frame will be low. Existing approaches to increasing the SNR for enabling AF in sub-optimal lighting conditions is to increase the camera exposure time resulting in longer sensor integration time. Clearly, this approach cannot work for video acquisition where

a high frame rate (short exposure time), is required to avoid motion blur [42]. Moreover, the smart phones being hand-held devices are not stable enough for long exposure times. This calls for development of an advanced image and signal processing (ISP) pipeline for image/video enhancement that is computationally efficient and thus requires low power for implementation on a resource constrained battery powered smart phone processor.

Traditional algorithms prevalent in computational imaging and signal processing are hand-crafted empirical rules synthesized to achieve a desired goal. In contrast, our approach is to craft qualitatively and fundamentally new algorithms that have superior and unprecedented performance compared to conventionally designed algorithms by emulating laws of physics. For example, Phase Stretch Transform emulates propagation of image through a dispersive optical medium and has inbuilt intrinsic equalization ability for feature detection in low contrast visually impaired images [25, 29].

Contrast enhancement improves the visual quality by transforming the input image/video data to make optimal use of the dynamic range available on a given display or an output device. In this Chapter, a novel computational imaging paradigm is proposed in which the image is transformed by subjecting it to propagation through a 2D nonlinear medium with the propagation modeled by the Non-Linear Schrodinger Equation (NLSE). The overall framework is illustrated in Figure 2.1.

This is an entirely new approach to image and video contrast enhancement and the resulting algorithm is extremely computationally efficient to the extent that can process 8K video at 100 frames per second. We call this algorithm Nonlinear S̲chrodinger Framework for I̲mage and V̲ideo E̲nhancement (SIVE). We also show how the NLSE operator becomes an edge detection algorithm with exceptional performance in low light levels. In certain cases, these algorithms have the potential to be implemented in physical optics.

8

Overall, the contributions of this Chapter are fourfold. 1) Non-Linear Schrodinger Equation (NLSE), the master equation in optical physics is exploited to invent a new class of computational imaging algorithms with best-in-class performance. Particularly, we propose a novel physics inspired method for image enhancement that offers significantly faster computational speed over state-of-the-art algorithms. It also improves both the objective and subjective image quality and in principle can be physically implemented as it emulates a physical phenomenon. We also show extension of the NLSE operation to an edge detection method.

2) Because of its extremely low latency and high computational efficiency, we are able to demonstrate image enhancement of real-time video at 4K and 8K resolutions. 3) To improve its performance over diverse lighting conditions, we show a neural network guided implementation where the network learns the parameters of the algorithm in an unsupervised fashion. 4) We demonstrate superiority of our proposed method over the state-of-the-art algorithms for color gamut expansion. All of these attributes make our algorithm superior to existing methods, and make it suitable for integration in standard ISP pipeline of a smart phone camera. For example, as our proposed method is computationally efficient it can help in maintaining longer battery life as well as improving the accuracy of many real-time machine vision tasks on a smart phone.

The rest of the Chapter is organized as follows. In section 2.2 we discuss current state of the art in the field of contrast enhancement. The physics foundations of our proposed Nonlinear S̲chrodinger Framework for I̲mage and V̲ideo E̲nhancement (SIVE) algorithm is discussed in detail in section 2.3 and its adaptive version utilizing a neural network is detailed in section 2.4. Section 2.5 compares the performance with state-of-the-art in contrast enhancement. We present the effect of dispersion that leads to an edge detection method in section 2.6. Finally we summarize our findings in section 2.7.

2.2 Prior Works

Since the primary focus of the present Chapter is on contrast enhancement, here we will review the prior art in this field. To address the issue of contrast enhancement, many algorithms have been proposed for increasing the contrast of images with low dynamic range to obtain enhanced images with higher perceptual quality. In general, these methods can be categorized into three main types: (i) traditional histogram-based, (ii) retinex theory-based and (iii) data-driven (deep learning-based) methods. Due to the simplicity, ease-of-implementation and effectiveness, histogram-based methods have been widely adopted. The histogram equalization (HE) [11] which flattens the histogram and stretches the dynamic range of the intensity levels, is one of the most popular techniques used for contrast enhancement. However, as we will illustrate in next section, HE usually results in over-enhancement that distorts the intrinsic colour space of the original image. To overcome this limitation, there are other sophisticated histogram-based enhancement methods such as contrast limited adaptive histogram equalization (CLAHE) [12] which constructs the enhanced image by mapping the distribution of pixel intensity with some constraints and an adaptive control. However, when the histogram has spikes, CLAHE often results in losing intrinsic image details and produces unnatural artifacts [43]. Other histogram based methods such as Fuzzy-Contextual Contrast Enhancement (FCCE) [13] use computationally intensive approach to obtain a transfer function for the enhanced image using fuzzy contextual information from the low contrast image.

Retinex theory [44] assumes that the scene in a human eye is a product of illumination and reflectance layer such that illumination is a measure of the intensity of incident light on the objects in the scene and the reflectance represents the physical characteristics of objects. These methods first decompose the low contrast image as a product of illumination and reflectance, either using a simplified physical

model of light reflection [45] or treat it as an optimization problem [14, 15, 46, 47]. For example, structure-revealing low-light image enhancement (SRIE) [14] estimates reflectance and illumination simultaneously using a weighted variational model. Then, the estimated illumination is adjusted for uniformity and projected back to the recovered reflectance layer to obtain high contrast images. Methods such as low-light image enhancement via illumination map estimation (LIME) [15], only estimates illumination by imposing a structure prior and then, use reflection to produce the enhanced results. These methods have been extensively studied and developed in the past few decades and have demonstrated promising results for a narrow range of test scenarios. However, there are several concerns that restrict their widespread usage. First and foremost, since these methods involve the reconstruction of illumination and reflectance layer by an image decomposition model or an optimization model based on retinex theory, these are extremely computationally intensive their use in a camera ISP pipeline is restricted. Second, these methods heavily rely on knowing constraints specified by the scene. Last, the decomposed reflectance layer is often contaminated with residual illumination information, and the estimated illumination contains halo artifacts which results in poor image quality [46].

Recently, data driven methods have shown leading performance on various computational imaging tasks, ranging from low-level demosaicing, denoising and image sharpening to high-level image adjustment and color enhancement, some of which can solve multiple tasks and can be easily integrated in an end-to-end fashion in a camera ISP pipeline [13, 48–52]. For the task of contrast enhancement, GLobal illumination Aware and Detail-preserving Network (GLAD-Net) [17], Multi-Branch Low-Light Enhancement Network (MBLLEN) [16] and Zero-Reference Deep Curve Estimation (Zero-DCE) [18] have shown promising performance. GLADNet consists of two steps, global illumination estimation step in which a encoder-decoder network produces an illumination estimation and de-

tail reconstruction step in which a convolutional network utilizes the input image and the estimated illumination from the previous step to enhance the output. On the other hand, MBLLEN method produces an enhanced image via fusion of features learned from low contrast image. Zero-DCE method, a highly computationally efficient contrast enhancement method uses a set of carefully formulated non-reference loss functions for training. Since these methods implicitly learn the statistics of images, they often fail with high error when operating outside the regime for which the model is trained.

2.3   The Physics Foundation

We propose a novel physics-inspired method which formulates image and video enhancement as a process of propagation of image through a 2D nonlinear optical medium with engineered linear and nonlinear optical properties (Figure 2.1). First, we describe the physics involved in our Nonlinear Schrodinger Framework for Image and Video Enhancement (SIVE) method. We then describe the conversion of the continuous-variable equations to discrete domain and re-parameterization of the physics equations leading to the formulation suitable for image processing.

We start with the two-dimensional Non-Linear Schrodinger Equation (NLSE) of pulse propagation through an optical fiber [37], derived from Gross-Pitaevskii equation:

$$\frac{\partial A(x,y;z)}{\partial z} = \frac{i\beta_2}{2}\left(\frac{\partial^2 A(x,y;0)}{\partial x^2} + \frac{\partial^2 A(x,y;0)}{\partial y^2}\right) \qquad (2.1)$$
$$+ i\gamma|A(x,y;0)|^2 A(x,y;0)$$

where β_2 is the dispersion coefficient and γ is the nonlinear coefficient. x, y, z are the spatial dimension variables such that the input signal is function of x, y and propagating along the z direction. Thereby, $A(x,y;0)$ is the input signal at location $z = 0$ and the solution to the above differential equation obtained after propagation through the optical medium is given by $A(x,y;z)$. We can rewrite

above NLSE in the form of dispersive operator $\mathbb{D}$ and nonlinear operator $\mathbb{N}$ as follows:

$$\frac{\partial A(x, y; z)}{\partial z} = (\mathbb{N} + \mathbb{D}) \, A(x, y; 0) \tag{2.2}$$

where the nonlinear operator is defined as follows:

$$\mathbb{N} = i\gamma |A(x, y; 0)|^2 \tag{2.3}$$

and similarly, the dispersive operator is defined as follows:

$$\mathbb{D} = \frac{i\beta_2}{2} \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right) \tag{2.4}$$

2.3.1 Dispersive Regime

In order to understand the individual impact of each of these operators (or the corresponding physical regime) on image enhancement, we independently solve the NLSE for each of these operators. We start with the dispersive operator and solve it as follows:

$$\begin{aligned} \frac{\partial A_D(x, y; z)}{\partial z} &= \mathbb{D} A(x, y; 0) \\ &= \frac{i\beta_2}{2} \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right) A(x, y; 0) \end{aligned} \tag{2.5}$$

where A_D is the solution of the reduced NLSE in the dispersive regime. We implement the dispersion operation in frequency domain by taking the Fourier transform of the above equation, as it leads to the solution easily:

$$\frac{\partial \tilde{A}_D(k_x, k_y; z)}{\partial z} = \frac{i}{2} \left(k_x^2 + k_y^2 \right) \tilde{A}(k_x, k_y; 0) \tag{2.6}$$

Here, $\tilde{A}_D$ and $\tilde{A}$ are output and input spectrum respectively obtained after the Fourier transform such that k_x and k_y are two dimensional spatial frequency variables. The above differential equation can be solved as follows:

$$\tilde{A}_D(k_x, k_y; z) = \exp\left(\frac{i * z}{2} \left(k_x^2 + k_y^2 \right) \right) \tilde{A}(k_x, k_y; 0) \tag{2.7}$$

As shown in the above equation, we now have a dispersive phase $\phi(k_x, k_y)$ that is function of spatial frequency k_x, k_y, given by:

$$\phi(k_x, k_y) = \frac{z}{2} \left(k_x^2 + k_y^2\right) \tag{2.8}$$

The solution of the reduced NLSE in the dispersive regime, A_D can be obtained by taking Inverse Fourier transform as follows:

$$A_D(x, y; z) = \int_{-\infty}^{\infty} \tilde{A}(k_x, k_y; 0) \exp\left(i\phi(k_x, k_y)\right)$$
$$\exp\left(ik_x x\right) \exp\left(ik_y y\right) dk_x dk_y \tag{2.9}$$

We rewrite the above equation by expanding the phase kernel using Euler expansion as follows:

$$A_D(x, y; z) = \int_{-\infty}^{\infty} \tilde{A}(k_x, k_y; 0) \Big(cos\left(\phi(k_x, k_y)\right)$$
$$+isin\left(\phi(k_x, k_y)\right)\Big) \exp\left(ik_x x\right) \exp\left(ik_y y\right) dk_x dk_y \tag{2.10}$$

In this regime, the dispersive phase is very small, and as such we use the following approximation, $cos\theta + isin\theta = 1 + i\theta$. Thereby, the equation can be written as:

$$A_D(x, y; z) = \int_{-\infty}^{\infty} \tilde{A}(k_x, k_y; 0)$$
$$\Big(1 + i\,\phi\left(k_x, k_y\right)\Big) \exp\left(ik_x x\right) \exp\left(ik_y y\right) dk_x dk_y \tag{2.11}$$

or,

$$A_D(x, y; z) = \int_{-\infty}^{\infty} \tilde{A}(k_x, k_y; 0)$$
$$\Big(1 + i\frac{z}{2}\left(k_x^2 + k_y^2\right)\Big) \exp\left(ik_x x\right) \exp\left(ik_y y\right) dk_x dk_y \tag{2.12}$$

By implementing the Fourier transform property of differentiation, we can compute the output as follows:

$$A_D(x, y; z) = A(x, y; 0) + +i\frac{z}{2}\left(\frac{\partial^2 A(x, y; 0)}{\partial x^2} + \frac{\partial^2 A(x, y; 0)}{\partial y^2}\right) \tag{2.13}$$

The interesting aspect of the above transformation comes from the complex behavior of the output. The real part of dispersive output is equal to the input

14

signal while the imaginary component of the dispersive output gives us second order gradients of the input along x and y direction (Fourier transform property of differentiation). We will show later how this dispersive operation leads to Phase Stretch Transform [22–27], an edge detection method with a natural equalization property making it ideal for feature extraction in visually impaired images.

2.3.2   Nonlinear Regime

We now compute the output of the nonlinear operator $\mathbb{N}$, from the following differential equation:

$$\frac{\partial A_N(x, y; z)}{\partial z} = \mathbb{N}\left(A(x, y; 0)\right) = i\gamma|A(x, y; 0)|^2 A \tag{2.14}$$

The solution to the above differential equation gives us the output $A_N(x, y; z)$ at location z described by the following equation:

$$A_N(x, y; z) = A(x, y; 0) * \exp\left(i\phi_{NL}\right) \tag{2.15}$$

The nonlinear operator $\mathbb{N}$ imparts a nonlinear phase ϕ_{NL} on the input image (a real quantity) such that this self-phase term is a function of the input signal $|A(x, y; z)|^2$, nonlinear coefficient γ and physical propagation distance z given by the following expression:

$$\phi_{NL} = \gamma * |A(x, y; 0)|^2 * z \tag{2.16}$$

which can be written in form of input signal power $P_0 = |A(x, y; z)|^2$, nonlinear coefficient γ and physical propagation distance L_{eff} as follows:

$$\phi_{NL} = \gamma * |A(x, y; 0)|^2 * dz = \gamma * P_0 * L_{eff} \tag{2.17}$$

We revise the above equation in terms of the nonlinear refractive index (physical optical property) of the medium as:

$$\phi_{NL} = \gamma * P_0 * L_{eff} = k_0 * n * L_{eff} \tag{2.18}$$

Where $k_0 = \frac{2\pi}{\lambda}$ is wave number, λ is the wavelength and the refractive index n is a nonlinear function of input signal A. We generalize this dependency to various orders of input but, for simplicity, here we show up to second order which gives the expression of refractive index n as follows:

$$n = n_0 + n_1|A| + n_2|A|^2 + O(|A|^n) \tag{2.19}$$

The second term can be identified with Pockel's effect in nonlinear optics [53]. Similarly, the third term is known as Kerr effect [54]. The nonlinear phase can now be written as,

$$\phi_{NL} = k_0 * \left(n_0 + n_1|A| + n_2|A|^2\right) * L_{eff}$$
$$= J_0 + J_1|A| + J_2|A|^2 \tag{2.20}$$

where we have parametrized the nonlinear phase with $J_i = k_0 L_{eff} n_i$ for $i = \{0, 1, 2, ..\}$. And therefore, the operation of nonlinear operator $\mathbb{N}$ can be computed as:

$$A_N(x, y; z) = \mathbb{N}A(x, y; 0)$$
$$= A(x, y; 0)\left(\cos\left(\phi_{NL}\right) + i * \sin\left(\phi_{NL}\right)\right) \tag{2.21}$$

2.4 Image Enhancement via 2D Nonlinear Propagation

We first consider a non-dispersive and lossless medium and therefore, $\alpha = \beta_2 = 0$. In this scenario the optical properties (refractive index) of the medium is dependent only on the input signal and this input dependent refractive index imposes a nonlinear transformation adding a phase which is proportional to the input intensity as shown previously.

We would like to point out here that our input image is a real-valued vector and the optical propagation induces an imaginary component that makes the output complex-valued vector. We now define Nonlinear Schrodinger Framework for Image and Video Enhancement (SIVE) algorithm as the imaginary component

16

of the output computed by NLSE operation on the input image:

$$\hat{A} = \text{SIVE}\{A(x,y;0)\} = \text{imag}\{A(x,y;z)\} \tag{2.22}$$

and

$$A(x,y;z) = \mathbb{N}A(x,y;0) \tag{2.23}$$

Here $\hat{A}$ is the contrast enhanced output after propagation of input A through the nonlinear optical medium such that the phase parameters J_0, J_1 and J_2 control the amount of enhancement. The induced pixel-dependent nonlinear phase converts the real-valued input image into a complex-valued array, the complex component of which is the original image raised to 2nd and 3rd (and higher) powers of the input.

2.4.1 Proof-of-concept example for contrast enhancement

To give the reader a glimpse of the results before we engage in the detailed discussions, here we show in Figure 2.2, a test image from the SICE dataset [55]. For computing the SIVE enhanced output, we use fixed value of parameters $J_0 = 10^{-4}$, $J_1 = 10^{-3}$ and $J_2 = 10^{-2}$ that we found to work for most images. When computing the nonlinear phase using these parameter values, the image brightness is first normalized for each image to the range [0,1]. Unless otherwise stated, for all our SIVE simulations we use these values for the preliminary results. In the next section we describe a neural network optimization that learns the optimum parameters in a non-supervised fashion.

Our method is orders of magnitude faster than other computationally expensive state-of-the-art methods as evident from the runtime performance curve shown in Figure 2.3(A) for this test image. For LIME, SRIE and FCCE methods, the open sourced codes are available in Matlab, and for HE and CLAHE, Matlab has highly optimized built-in functions. Therefore we used the Matlab environment for these methods and computed the runtime using the Matlab timeit

Figure 2.2: **Comparison of proposed SIVE algorithm with the state-of-the-art methods**– Our proposed method expands the dynamic range of input image (A) producing an enhanced image (B) with a wide color space yet preserving the naturalness. As seen HE (C) distorts the intrinsic color space of the image while other methods (D-J) produce an enhanced image with comparatively low contrast. SIVE = nonlinear S̲chrodinger framework for I̲mage and V̲ideo E̲nhancement, HE = Histogram Equalization [11], CLAHE = Contrast Limited Adaptive Histogram Equalization [12], FCCE= Fuzzy-Contextual Contrast Enhancement [13], SRIE = Structure-Revealing low-light Image Enhancement [14], LIME = Low-light image enhancement via Illumination Map Estimation [15], MBLLEN = Multi-Branch Low-Light Enhancement Network [16], GLADNet = GLobal illumination Aware and Detail-preserving Network [17], Zero-DCE = Zero-Reference Deep Curve Estimation [18].
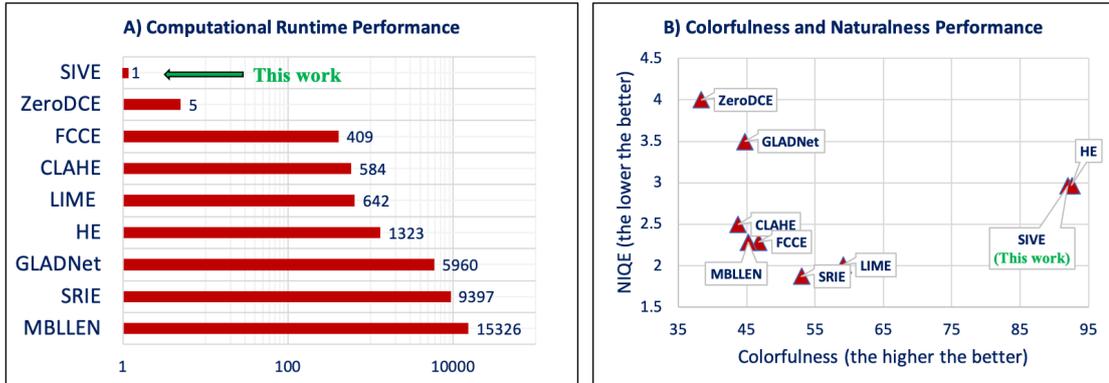
Figure 2.3: **Performance comparison of proposed SIVE method with the state-of-the-art methods** – As shown above in (A), our proposed method is orders of magnitude faster than the state-of-the-art methods. The runtime is computed for the test image shown in Figure 2.2 of dimension 800×1200×3 pixels and is normalized w.r.t. to SIVE (being the smallest) and rounded off to nearest integer. Also, as can be seen from the performance evaluation metric plot (B), our proposed method SIVE achieves almost highest colorfulness [19] but yet preserves naturalness of the image with a much lower Naturalness Image Quality Evaluator (NIQE) [20] than methods such as Zero-DCE and GLADNet. SIVE = nonlinear Schrodinger framework for Image and Video Enhancement, HE = Histogram Equalization [11], CLAHE = Contrast Limited Adaptive Histogram Equalization [12], FCCE= Fuzzy-Contextual Contrast Enhancement [13], SRIE = Structure-Revealing low-light Image Enhancement [14], LIME = Low-light image enhancement via Illumination Map Estimation [15], MBLLEN = Multi-Branch Low-Light Enhancement Network [16], GLADNet = GLobal illumination Aware and Detail-preserving Network [17], Zero-DCE = Zero-Reference Deep Curve Estimation [18].

function on an Intel Core i7 CPU (3.6 GHz). For our SIVE, as well as MBLLEN, GLADNet and Zero-DCE, computation time is evaluated in a Python environment on a NVIDIA GeForce GTX Titan X GPU.

We use a no-reference image quality score knows as the measure of Colorfulness as described in [19] to evaluate the perceptual impact of enhancement on image. A higher colorfulness value represents a higher image quality. The value of colorfulness of the original image increases from C=45.23 to a much higher value of C=91.97 (for SIVE) during the enhancement process as shown in Figure 2.3(B). We also consider naturalness image quality evaluator (NIQE) [20] that measures the distance between the natural scene statistics and the enhanced image such that a lower NIQE value represents a higher image quality. As shown, our proposed method has a low NIQE value =2.96, demonstrating the high visual quality of enhanced result while also preserving the intrinsic image details.

2.4.2 SIVE-Net for Adaptive Contrast Enhancement

While SIVE works well as an open-loop algorithm with fixed preset parameters, it would perform better if its parameters are optimized for different image conditions. In this section, we show a deep learning technique where the solution of the nonlinear Schrodinger equation is guided to the optimum output defined by the choice of a loss function.

To enable adaptivity, we incorporate an unsupervised learning methodology for generalizing SIVE parameters J_0, J_1 and J_2 as function of spatial coordinate x, y. If these physical parameters have the same values for all pixels it can over-/under enhance local regions. To address this problem, we formulate a Physics-AI symbiotic model that learns a pixel-wise mapping for physical parameters J_0, J_1 and J_2 to produce an enhanced image with much wider dynamic range, as shown

in Figure 2.4. Hence, Eq. (8) can be reformulated as:

$$\phi_{NL}(\text{x},\text{y}) = J_0(\text{x},\text{y}) + J_1(\text{x},\text{y})|A| + J_2(\text{x},\text{y})|A|^2 \qquad (2.24)$$

where J_0, J_1 and J_2 are physical parameters with the same size as the input image. To learn these physical parameters, we propose and demonstrate an adaptive version of the SIVE algorithm called SIVE-Net. The input to the SIVE-Net is the image while the outputs are a set of pixel-wise physical parameter array J_0, J_1 and J_2 of SIVE.
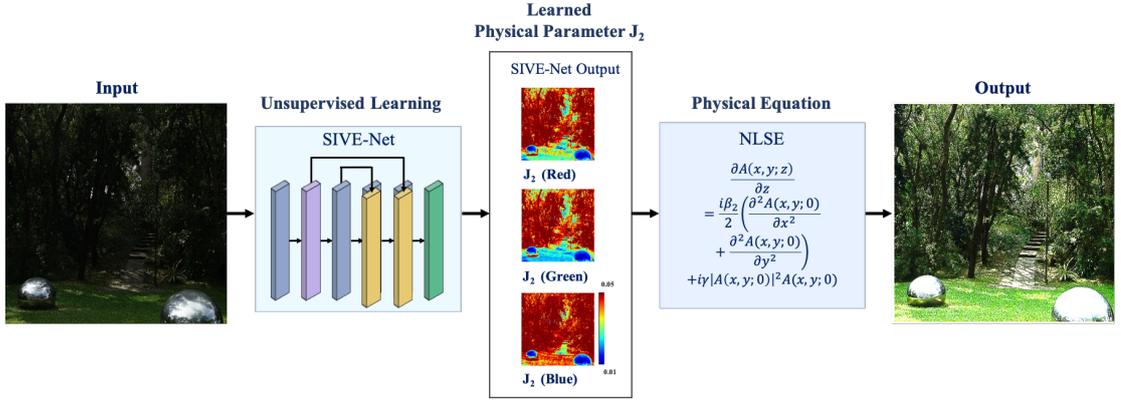
Figure 2.4: **Unsupervised learning framework for parameter optimization for our proposed method SIVE**. The NLSE parameter is learned through a deep neural network that outputs a mapping for physical parameter J_2 corresponding to each pixel of the image. The learned J_2 parameters are a type of features extracted from the image. Next this learnt physical parameter is used in the physics equations to produce the enhanced output. The learnt features guide the NLSE into the desired output. Batch normalization is applied to the input images. And the output is re-scaled to the original dynamic range, for eg. the output is re-scaled to a range of 0-255 for a 8-bit input.

For our network architecture, we employ a convolutional neural network of two convolutional layers with symmetrical concatenation. Each layer consists of 64 convolutional kernels of size 3×3 and stride 1 followed by the ReLU activation

function. To keep our model end-to-end differentiable, we remove down sampling layers as these operations break the relations of neighboring pixels. The last convolutional layer is followed by the *tanh* activation function and normalization in the range [0.01,0.05], which produces a feature map for the SIVE parameter J_2. We then use this learned physical parameters to empirically compute $J_0 = 0.01 * J_2$ and $J_1 = 0.1 * J_2$. We note that the SIVE-Net has only 152,899 parameters for an input image of size 400×600×3 which makes this network lightweight and an ideal choice for image and video enhancement on computational resource-limited devices, such as smartphones.

In learning the optimal SIVE parameters using common error metrics such as Peak Signal-to-Noise Ratio (PSNR), Structural Similarity (SSIM), and Mean Absolute Error (MAE) is prone to the availability of ground truth. Therefore, we propose a set of non-reference loss functions that measure the enhancement performance at local and global levels.

Inverse Loss L_I: The transformation of the image should be such that it enhances the under exposure regions and avoids local over exposure. As such the SIVE parameters should be proportional to the inverse of the local brightness. In this case, we can define the inverse loss as:

$$L_I = \frac{1}{N} \sum_{n=1}^{N} \left(\left| J_{2n} - \bar{A}_n \right| \right) \tag{2.25}$$

Where $\bar{A}_n$ is average intensity of the 8x8 local region of the inverted image, and J_{2n} is the average intensity of the learned SIVE parameter. Here, N is the total number of non-overlapping local regions of size 8x8.

Gradient Loss L_G: Natural scenes consists of large number of edges and corners. [56]. Preserving the edge information is vital during the enhancement. To ensure that learned physical parameters are consistent with the original image,

we propose a gradient loss as:

$$L_G = \frac{1}{N} \sum_{n=1}^{N} \left| \nabla_x J_{2n} - \nabla_x \bar{A}_n \right| + \left| \nabla_y J_{2n} - \nabla_y \bar{A}_n \right| \qquad (2.26)$$

Where $\nabla_x J_{2n}$ and $\nabla_y J_{2n}$ is the spatial derivative of the learned SIVE parameters along the x and y direction respectively of non-overlapping regions of size 8x8. Similarly, $\nabla_x \bar{A}_n$ and $\nabla_y \bar{A}_n$ are the spatial derivatives of the inverted input image along the x and y directions, respectively of the n-th non-overlapping region of size 8x8. We use the L1-norm to compute the distance between gradients of predicted SIVE parameters with those of the inverse image.

Exposure Control Loss L_E: To minimizes the distance between the average intensity of a local region to the well-exposedness level L , we use the exposure control loss as previously defined in [18].Therefore, the exposure control loss L_E can be defined as:

$$L_E = \frac{1}{N} \sum_{n=1}^{N} \left(\left| \hat{A}_n - L \right| \right) \qquad (2.27)$$

Here also we keep the non-overlapping regions of size 8×8 and calculate the average intensity of the enhanced image $\hat{A}_n$. We set L to 0.6 in our experiments.

Color Constancy Loss L_C: According to the Gray-World color constancy hypothesis [18, 57] color in each sensor channel averages to gray over the entire image. We design here a color constancy loss to ensure that color channels of the enhanced image do not deviate from the natural statistics. The color constancy loss L_C can be expressed as:

$$L_C = \sum_{\forall(i,j)\epsilon\sigma} \left(\left| \frac{\hat{A}_i}{\hat{A}_j} - \frac{A_i}{A_j} \right| \right) \qquad (2.28)$$

Where $\hat{A}_i$ denotes the average intensity over all pixels of i-th color channel in the enhanced image, A_i denotes the same for the input image, and (i,j) represents a pair of channels from color space $\sigma = \{(R,G),(R,B),(G,B)\}$, which comprises of permuted pairs of channels from the RGB color space. This loss maintains the intrinsic color distribution of the original image during the enhancement process.

Therefore, the total loss L_{Total} can be expressed as:

$$L_{\text{Total}} = W_I L_I + W_G L_G + W_E L_E + W_C L_C \qquad (2.29)$$

where W_I, W_G, W_E and W_C are the weights of the individual losses.
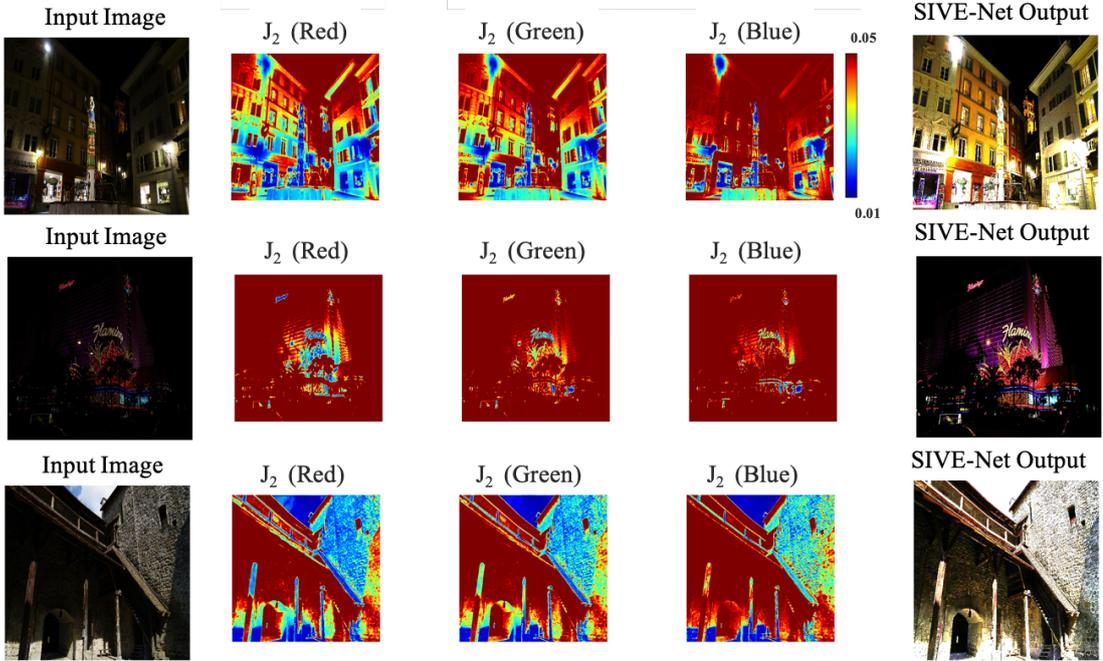
Figure 2.5: **Adaptive enhancement achieved by SIVE-Net**. The input images are test images from the SICE dataset which pass through the trained SIVE-Net to produce NLSE parameters (abstractions from the input image). The low light regions are enhanced more and the already enhanced regions in the input image are preserved.

Implementation Details: Our framework is implemented with PyTorch on an NVIDIA GeForce GTX Titan X GPU. For improving the fidelity of enhancement using SIVE, we incorporate both under and over-exposed images into our training set. To this end, we use 3022 images of multi-exposure sequences from the SICE dataset Part1 [55] that covers diverse image qualities and content. We randomly split these images of different exposure levels into two parts, 2,400 images

for training and the rest for validation. All these images are resized to 400x600x3 pixels and also scaled in the range [0,1]. In the experiment, training is done using the ADAM optimizer [58] with a learning rate of $a = 0.002$, $b1 = 0.9$, $b2 = 0.99$ and $e = 10^{-8}$ for network optimization. We also use the learning rate decay strategy, which reduces the learning rate to 95% before the next epoch. A batch size of 8 is applied. The weights W_I, W_G, W_E and W_C are set as 20, 4 3.5 and 2 respectively. As shown in Figure 2.5, the network learns an optimal mapping for the physical parameter J_2 to produce a contrast enhanced image governed by the NLSE.

2.5   Performance Evaluation

We evaluate the performance of our proposed method SIVE with existing state-of-the-art image and video enhancement methods through extensive experiments. Overall, we have done three major sets of experiments as follows:

1) Histogram based methods (HE, CLAHE, FCCE), retinex theory based methods (SRIE, LIME) and data driven methods (GLADNet, MBBLLEN, Zero-DCE), are used for comparison of computational performance during runtime. The results of these existing methods are reproduced by using published source codes with recommended parameters.

2) We next conduct experiment on 4K and 8K videos and demonstrate that our method can support real-time video processing.

3) We show another set of comparison on the task of image enhancement for color gamut expansion, for resolution enhancement and for noise suppression during the contrast enhancement process.

Computational performance: To evaluate the computational performance of our algorithm, we perform quantitative experiments on standard image sets from previous works including Kodak [59] (24 images), LIME [15] (10 images),

Method	Kodak	LIME	MEF	DICM
SIVE	0.0005	0.0006	0.0004	0.0005
SIVE-Net	0.002	0.003	0.002	0.003
ZeroDCE	0.003	0.004	0.003	0.005
HE	0.3	0.4	0.13	0.26
CLAHE	0.13	0.17	0.06	0.11
SRIE	3.3	14.03	4.27	5.45
LIME	0.14	0.2	0.09	0.14
FCCE	0.08	0.12	0.03	0.06
GLADNet	1.3	1.79	0.6	1.1
MBBLLEN	6.3	17.68	3.9	6.6

Table 2.1: **Quantitative comparisons in terms of average runtime (in seconds) performance**. Best value is highlighted in blue, second best value in magenta and worst value in red. As seen, SIVE is 6x times and SIVE-Net is 1.5x faster than state-of-the-art method Zero-DCE. Both SIVE and SIVE-Net are orders of magnitude faster than other techniques.

MEF [60] (17 images) and DICM [21] (69 images). For HE, CLAHE, FCCE, LIME and SRIE, we used the MATLAB environment to compute the runtime where as for SIVE, SIVE-Net, GLADNet, MBBLLEN, Zero-DCE we evaluated the runtime in a python environment. For Matlab simulations, the runtime is computed using Matlab timeit function on Intel Core i7 CPU at 3.6 GHz. The computational time for our method is by far the lowest when compared to the state-of-the-art algorithms.

As seen in Table 2.1, SIVE is 6x times and SIVE-Net is 1.5x faster than state-of-the-art method Zero-DCE. Both SIVE and SIVE-Net are orders of magnitude faster than other techniques.

Method	Kodak	LIME	MEF	DICM
SIVE	59.92	48.5	31.94	51.10
SIVE-Net	60.63	48.98	32.36	50.8
ZeroDCE	32.54	58.03	34.87	41.76
HE	48.25	44.15	28.91	32.45
CLAHE	36.74	38.23	22.65	33.14
SRIE	44.48	51.03	32.36	42.07
LIME	54.26	68.5	49.80	55.08
FCCE	37.98	45.61	25.32	34.95
GLADNet	38.43	38.82	28.79	35.18
MBBLLEN	51.25	54.99	31.64	43.98

Table 2.2: **Quantitative comparisons in terms of measure of Colorfulness [19] (the higher the better) averaged on all sample images from a given dataset**. Best value is highlighted in blue, second best value in magenta and worst value in red. As seen, both SIVE and SIVE-Net have better performance than most of the other methods as indicated by a relatively high value of the average measure of Colorfulness.

We also compare the enhancement performance using the measure of Colorfulness [19], such that the higher the colorfulness, the better. We compute the average measure of Colorfulness by considering all sample images from a given dataset. As seen in Table 2.2, SIVE-Net gives the best performance on this metric for Kodak dataset and second best for DICM dataset. SIVE also gives second best performance for Kodak dataset. In general, both SIVE and SIVE-Net have better performance than most of the other methods for LIME and MEF datasets. Even though ZeroDCE has best performance for LIME and MEF datasets, it does not perform well on the Kodak dataset.
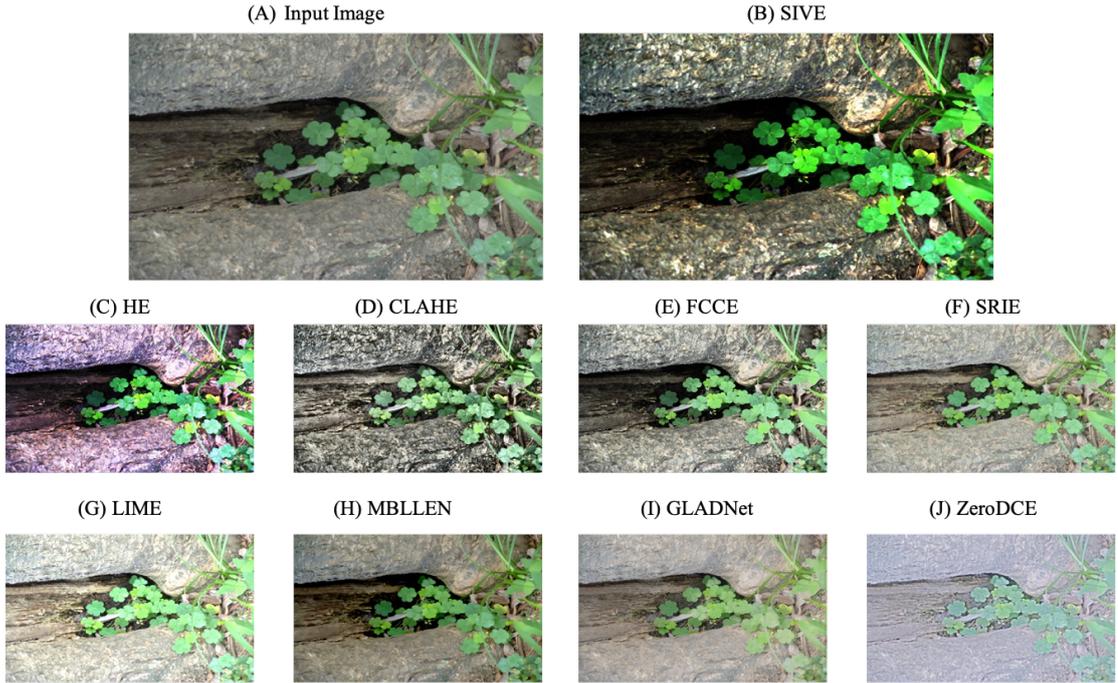
Figure 2.6: **Visual comparisons on an example image from DICM** [21] **dataset**. Nonlinear Schrodinger Framework for Image and Video Enhancement (SIVE) enhances the input image, widens the color gamut producing an output image with better perceptual quality but also preserves the inherent structure details. SIVE = nonlinear Schrodinger framework for Image and Video Enhancement, HE = Histogram Equalization [11], CLAHE = Contrast Limited Adaptive Histogram Equalization [12], FCCE= Fuzzy-Contextual Contrast Enhancement [13], SRIE = Structure-Revealing low-light Image Enhancement [14], LIME = Low-light image enhancement via Illumination Map Estimation [15], MBLLEN = Multi-Branch Low-Light Enhancement Network [16], GLADNet = GLobal illumination Aware and Detail-preserving Network [17], Zero-DCE = Zero-Reference Deep Curve Estimation [18].

In Figure 2.6, we show visual image enhancement comparison of our method SIVE with the state-of-the-art methods for one example image from DICM [21]

28

dataset. As shown, SIVE enhances the image, widens the color gamut but also preserves the inherent structure details.

Real-time 4K and 8K Video Processing: For demonstrating real-time video enhancement using SIVE, we use a publicly available video from SONY, with a playback frame rate of 30 FPS and 4K resolution, as shown in Figure 2.7. We show here that SIVE processes frames faster than the playback time (1/ frame rate) = 0.034 sec, and therefore, enable real-time video enhancement. The processing runtime is less than 0.001 second averaged over 100 frames on a NVIDIA GeForce GTX Titan X GPU.
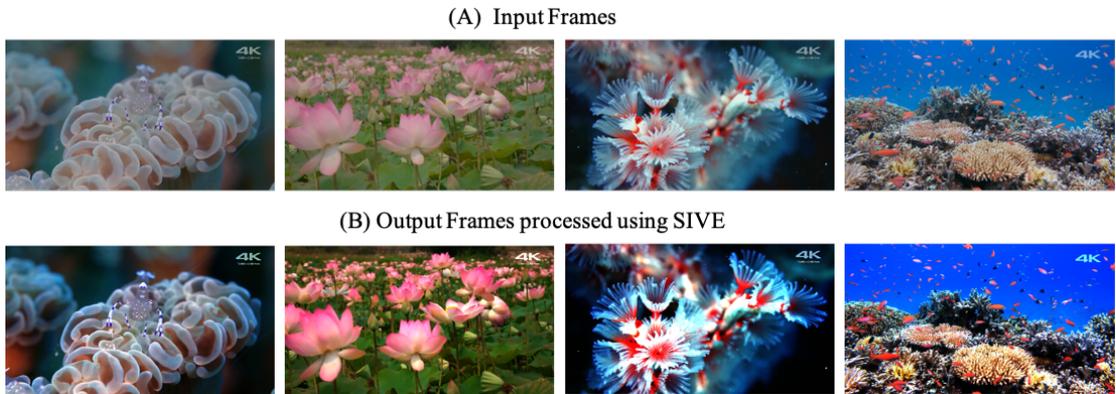
Figure 2.7: **Real-time video processing using Nonlinear Schrodinger Framework for Image and Video Enhancement (SIVE)**: We show here that SIVE has the potential for use in live video broadcasting as the processing overhead computed for an input video with a decent frame rate of 30 FPS is zero. The computation was carried out on a NVIDIA GeForce GTX Titan X GPU. We can process a 8K video on this GPU in less than 0.01 second, to achieve a frame rate greater than 100 FPS.

Method	Runtime (in seconds)	Achievable Frame Rate (in FPS)
SIVE	0.001	1000
ZeroDCE	0.03	30
CLAHE	4	0.25
FCCE	4	0.25
LIME	5	0.2
HE	12	0.08
GLADNet	30	0.03
SRIE	250	0.04
MBLLEN	300	0.003

Table 2.3: **Quantitative comparisons in terms of average runtime (in seconds) performance, and possible 4K video broadcasting frame rate**. Best value is highlighted in blue and worst value in red.

Specifically we show that SIVE is capable of handling 4K real-time video processing with a frame rate of 1000 FPS while the second most computationally efficient method ZeroDCE [18] can only support a 30 FPS frame rate as shown in Table 2.3. Moreover, we can process a 8K video in less than 0.01 second on this GPU and achieve a frame rate greater than 100 FPS, enabling live video streaming.

We also quantify the enhancement performance using non reference metrics like measure of Colorfulness [19] and naturalness image quality evaluator (NIQE) [20] for the two 4K video frames shown in Figure 2.8 and 2.9. The quantitative evaluation for these frames in Figure 2.8 and 2.9 is shown in Table 2.4 and 2.5, respectively. As seen in the tables, SIVE provides a very high measure of Colorfulness and a good (low) NIQE score.
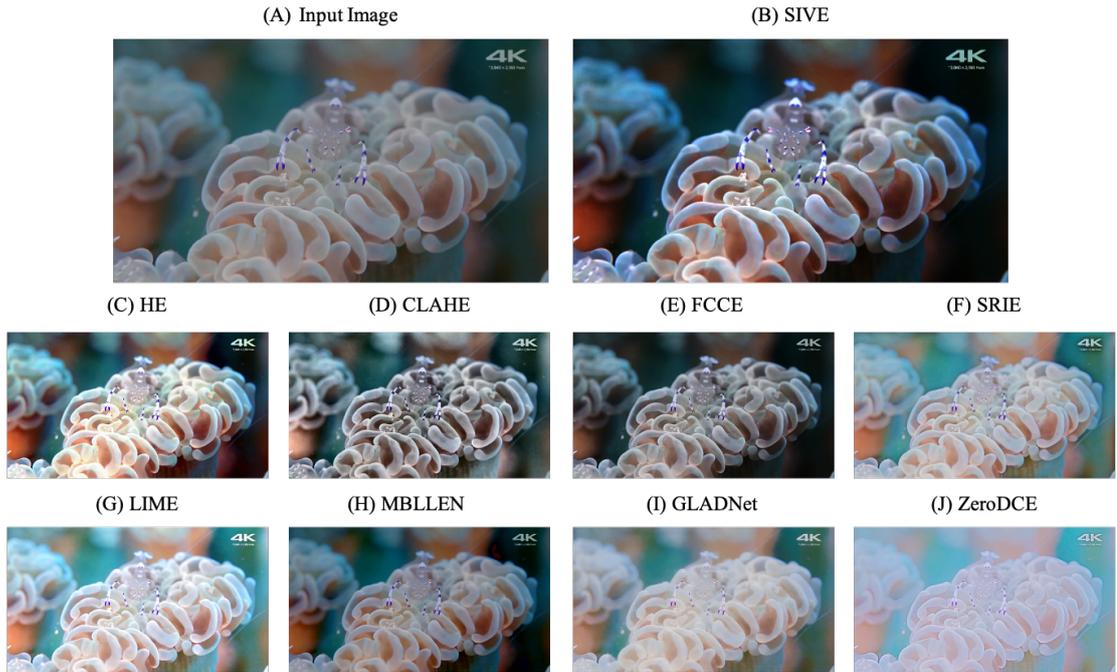
Figure 2.8: **Performance comparison of our proposed method SIVE with state-of-the-art methods for real-time video enhancement**. As seen, our proposed method SIVE enhances the input frame in real-time, bringing out the contrast and also, enabling possible live video broadcasting. SIVE = nonlinear Schrodinger framework for Image and Video Enhancement, HE = Histogram Equalization [11], CLAHE = Contrast Limited Adaptive Histogram Equalization [12], FCCE= Fuzzy-Contextual Contrast Enhancement [13], SRIE = Structure-Revealing low-light Image Enhancement [14], LIME = Low-light image enhancement via Illumination Map Estimation [15], MBLLEN = Multi-Branch Low-Light Enhancement Network [16], GLADNet = GLobal illumination Aware and Detail-preserving Network [17], Zero-DCE = Zero-Reference Deep Curve Estimation [18].

Method	Measure of Colorfulness (the higher, the better)	NIQE (the lower, the better)
SIVE	36.32	3.04
ZeroDCE	40.3	3.12
HE	41.77	2.69
CLAHE	24.95	2.67
SRIE	41.05	3.09
LIME	48.14	2.84
FCCE	20.12	2.69
GLADNet	30.29	3.13
MBLLEN	35.71	4.37

Table 2.4: **Quantitative comparisons in terms of measure of Colorfulness [19] and naturalness image quality evaluator (NIQE) [20] for image shown in Figure 2.8**. As the input image itself has a very narrow color space, SIVE has moderate measure of colorfulness. While a low NIQE value for SIVE indicates that the enhanced image has statistics similar to natural images. Best value is highlighted in blue and worst value in red. SIVE = nonlinear Schrodinger framework for Image and Video Enhancement, HE = Histogram Equalization [11], CLAHE = Contrast Limited Adaptive Histogram Equalization [12], FCCE= Fuzzy-Contextual Contrast Enhancement [13], SRIE = Structure-Revealing low-light Image Enhancement [14], LIME = Low-light image enhancement via Illumination Map Estimation [15], MBLLEN = Multi-Branch Low-Light Enhancement Network [16], GLADNet = GLobal illumination Aware and Detail-preserving Network [17], Zero-DCE = Zero-Reference Deep Curve Estimation [18]

Figure 2.9: **Performance comparison of our proposed method SIVE with state-of-the-art methods for real-time video enhancement**. Our proposed method SIVE enhances the input frame by expanding the input color gamut in real-time and providing an output image with better perceptual quality. SIVE = nonlinear $\underline{\text{S}}$chrodinger framework for $\underline{\text{I}}$mage and $\underline{\text{V}}$ideo $\underline{\text{E}}$nhancement, HE = Histogram Equalization [11], CLAHE = Contrast Limited Adaptive Histogram Equalization [12], FCCE= Fuzzy-Contextual Contrast Enhancement [13], SRIE = Structure-Revealing low-light Image Enhancement [14], LIME = Low-light image enhancement via Illumination Map Estimation [15], MBLLEN = Multi-Branch Low-Light Enhancement Network [16], GLADNet = GLobal illumination Aware and Detail-preserving Network [17], Zero-DCE = Zero-Reference Deep Curve Estimation [18].

Method	Measure of Colorfulness (the higher, the better)	NIQE (the lower, the better)
SIVE	68.79	3.22
ZeroDCE	38.35	3.8
HE	43.73	3.295
CLAHE	34.06	3.293
SRIE	49.12	3.8
LIME	56.47	3.53
FCCE	35.66	3.3
GLADNet	41.67	3.50
MBLLEN	49.45	4.07

Table 2.5: **Quantitative comparisons in terms of measure of Colorfulness [19] and naturalness image quality evaluator (NIQE) [20] for image shown in Figure 2.9.** As seen, SIVE has lowest NIQE value and highest measure of Colorfulness indicates that the enhanced image has statistics similar to natural images and best color gamut respectively. Best value is highlighted in blue and worst value in red.

Figure 2.10: **Color gamut expansion using our proposed SIVE method** - Our proposed method expands the input gray level in a uniform manner as shown in the color cloud preserving input image details. We also compare the Peak Signal-to-Noise Ratio (PSNR) between the enhanced outputs and reference DCI-P3 color space image. Both the PSNR (=17.84) and Colorfulness (=119.47) is highest for SIVE. Performance comparison of our proposed method SIVE with state-of-the-art methods for color gamut expansion. Our method expands the color gamut of input image and is also computationally efficient.

Color Gamut Expansion: In color theory, the gamut is the part of the color space that can be represented on the given display, or produced by an enhancement process. Display technologies are now able to produce images with a much wider color gamut than ever seen before [61]. We show here that using SIVE, we can exploit the full color potential of these new displays. As an example, we consider an input image having a narrow color space, sRGB [62], and a reference image with a wider color space, Apple's DCI-P3 [63]. We show that by application of SIVE on a sRGB image, we can extend the color gamut in an efficient manner and produce an enhanced image with a wide color gamut that is very similar to Apple's DCI-P3 as shown in Figure 2.10. This is also reflected from the measure of Colorfulness which is highest for SIVE (C=119.47) followed by HE (C=99.47). On the other hand, second most computationally efficient Zero-DCE produces the worst output image (C=56.10), shrinking the original color cloud of sRGB image (C=73.71). MBLLEN (C=97.3) also produces an output with much wider color space but is computationally the most expensive method. Therefore, SIVE produces the best output for color gamut enhancement as also reflected from the high value of measure of Colorfulness [19] shown in Table 2.6.

We use DCI-P3 as a reference image for all our quantitative evaluation of color gamut expansion and evaluate the enhancement performance using reference based metric Structural Similarity (SSIM) for evaluation and compare the enhanced image with respect to the DCI-P3 image as the reference image. As shown in Table 2.6, SSIM for SIVE is comparatively high indicating its closeness to DCI-P3 image. While SSIM is highest for MBLLEN, it is computationally the most expensive method (15000x times slower than SIVE). Therefore, our proposed method SIVE produces an enhanced output with better perceptual quality and is also computationally efficient.

36

Method	Computation Runtime (in seconds)	Measure of Colorfulness (the higher, the better)	NIQE (the lower, the better)	SSIM (DCI-P3 reference) (the higher, the better)
SIVE	0.0004	119.47	3.36	0.81
ZeroDCE	0.005	56.10	3.0	0.63
HE	0.4	99.47	3.5	0.8
CLAHE	0.13	69.49	3.46	0.84
SRIE	3.9	79.64	3.2	0.89
LIME	0.15	92.2	3.3	0.79
FCCE	0.09	71.96	3.02	0.87
GLADNet	1.04	64.43	3.4	0.8
MBLLEN	6.13	97.3	3.4	0.9

Table 2.6: **Quantitative comparisons in terms of runtime (in seconds) performance, measure of Colorfulness [19], naturalness image quality evaluator (NIQE) [20] and Structural Similarity (SSIM), for images shown in Figure 4.11**. SIVE is computationally efficient, produces a wide color gamut image, with moderate NIQE and SSIM score. MBLLEN produces a much wider color space image and has a high SSIM value but is the most computationally expensive (slowest) method. Best value is highlighted in blue and worst value in red. SIVE = nonlinear Schrodinger framework for Image and Video Enhancement, HE = Histogram Equalization [11], CLAHE = Contrast Limited Adaptive Histogram Equalization [12], FCCE= Fuzzy-Contextual Contrast Enhancement [13], SRIE = Structure-Revealing low-light Image Enhancement [14], LIME = Low-light image enhancement via Illumination Map Estimation [15], MBLLEN = Multi-Branch Low-Light Enhancement Network [16], GLADNet = GLobal illumination Aware and Detail-preserving Network [17], Zero-DCE = Zero-Reference Deep Curve Estimation [18].

We also do additional quantitative comparison for SIVE, HE and Zero-DCE for reference based metric Peak Signal-to-Noise Ratio (PSNR). Calculated PSNR for SIVE is highest equal to 17.84 followed by HE (PSNR =15.13). ZeroDCE (third most computationally efficient method after SIVE and SIVE-Net) has the lowest PSNR (=8.6). Clearly, SIVE has the best performance for color gamut expansion. Additionally, we also consider naturalness image quality evaluator (NIQE) [20] that measures the distance between the natural scene statistics and the enhanced image (lower NIQE value means a higher image quality). As shown, our proposed method SIVE has an excellent NIQE value =3.36, demonstrating that SIVE offers best in class performance for color gamut expansion as well as computational efficiency.

Resolution Enhancement: To demonstrate resolution enhancement using our proposed method SIVE, we conduct a experiment where we simulate the optical lens degradation by applying a Gaussian filter to the input image. resulting in a low contrast as well as a blurry image. As shown in Figure 2.11, our method improves the contrast of the input image as well as helps in improving the resolution. This feature can help the passive auto-focus (AF) system in the camera to correctly focus in images with low SNR (captured under low-light conditions) and improve image quality.

Noise Suppression: All contrast enhancement methods have the inherent drawback of amplifying the noise during the gray-level adjustment. We evaluate the performance of our image enhancement method SIVE on low-light input image with added noise as shown in Figure 2.12. As can be observed in this figure, the noise hidden in very low-light condition is really intense. While other methods sufficiently enhance the visibility of low-light regions in the image, they also intensively amplify the noise. By a visual comparison of result, it can be observed that the proposed method SIVE removes the noise in dark regions but also enhances the low-light regions.
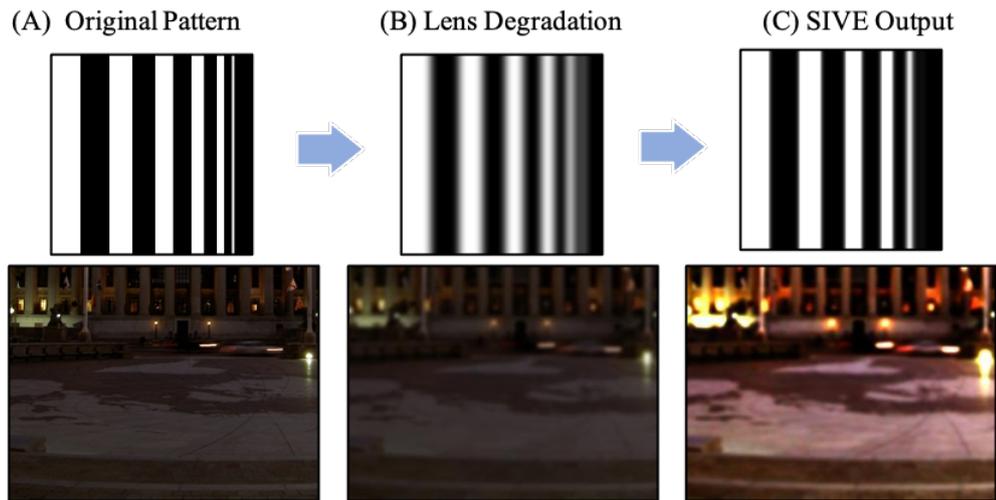
Figure 2.11: **Simultaneous resolution and contrast enhancement using our proposed Nonlinear Schrodinger Framework for Image and Video Enhancement (SIVE) method**- Our proposed method increases the resolution of features in the input image as well as expands the dynamic range of input gray level. For example, in the image shown above, the high frequency components in the scene such as the curves in the floor can be enhanced using our proposed method SIVE. SIVE = nonlinear Schrodinger framework for Image and Video Enhancement, HE = Histogram Equalization [11], CLAHE = Contrast Limited Adaptive Histogram Equalization [12], FCCE= Fuzzy-Contextual Contrast Enhancement [13], SRIE = Structure-Revealing low-light Image Enhancement [14], LIME = Low-light image enhancement via Illumination Map Estimation [15], MBLLEN = Multi-Branch Low-Light Enhancement Network [16], GLADNet = GLobal illumination Aware and Detail-preserving Network [17], Zero-DCE = Zero-Reference Deep Curve Estimation [18].
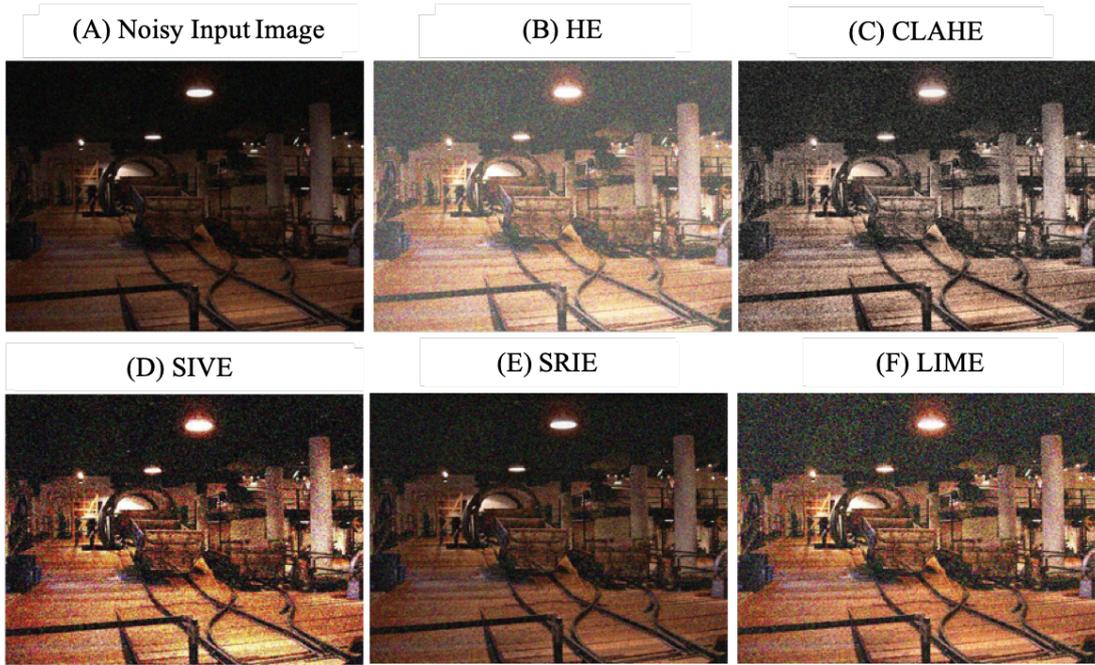
Figure 2.12: **Noise Suppression using our proposed Nonlinear Schrodinger Framework for Image and Video Enhancement (SIVE) method**: Our proposed method SIVE expands the input gray level but does not intensifies the noise in the image unlike other methods that enhance the noise in the image especially in the dark regions of the image. While SRIE does not amplify the noise however it fails to enhance the contrast of the image. Performance comparison of our proposed method SIVE with state-of-the-art methods for color gamut expansion. Our method expands the color gamut of input image and is also computationally efficient. SIVE = nonlinear Schrodinger framework for Image and Video Enhancement, HE = Histogram Equalization [11], CLAHE = Contrast Limited Adaptive Histogram Equalization [12], FCCE= Fuzzy-Contextual Contrast Enhancement [13], SRIE = Structure-Revealing low-light Image Enhancement [14], LIME = Low-light image enhancement via Illumination Map Estimation [15], MBLLEN = Multi-Branch Low-Light Enhancement Network [16], GLADNet = GLobal illumination Aware and Detail-preserving Network [17], Zero-DCE = Zero-Reference Deep Curve Estimation [18].

Input

Non-Linear Schrodinger Equation (NLSE)

$$\frac{\partial A(x,y;z)}{\partial z} + \frac{\alpha}{2} A(x,y;0)$$

$$= \frac{i\beta_2}{2}\left(\frac{\partial^2 A(x,y;0)}{\partial x^2} + \frac{\partial^2 A(x,y;0)}{\partial y^2}\right) + i\gamma |A(x,y;0)|^2 A(x,y;0)$$

PST
Dispersive
Regime
$(\alpha = \gamma = 0, \beta \neq 0)$

SIVE
Non-Linear Regime
$(\alpha = \beta = 0, \gamma \neq 0)$

Hybrid: SIVE+PST
Non-Linear and
Dispersive Regime
$(\alpha = 0, \beta \neq 0, \gamma \neq 0)$
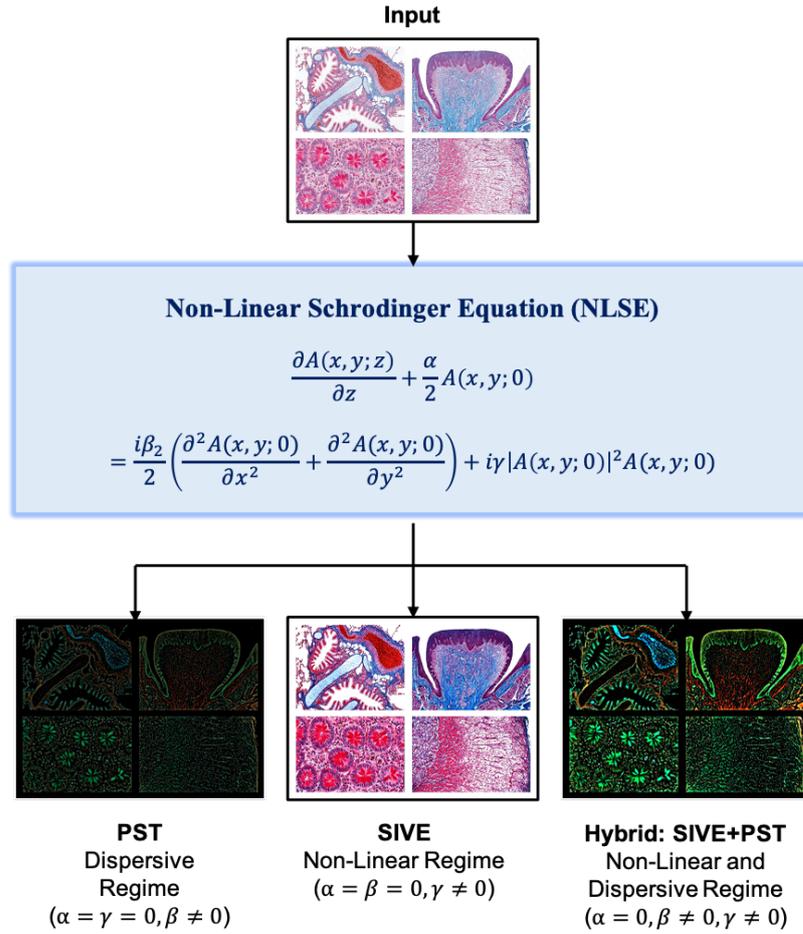
Figure 2.13: **Demonstration of NLSE operation on a digital image under various regimes**: In the dispersive regime, the NLSE operation results in an edge detection method called as PST, while for the nonlinear regime, the NLSE operator produces a contrast enhanced image. And in the event of a nonlinear and dispersive regime, NLSE operator produces a contrast enhanced edge output. SIVE = nonlinear Schrodinger framework for Image and Video Enhancement, PST = Phase Stretch Transform [22–27]

2.6   Edge Detection using Nonlinear Schrodinger Equation in Dispersive Regime

In this section, we consider the effect of dispersive propagation on digital images and show that the the solution of NLSE in the dispersive regime leads to the Phase Stretch Transform [22–27], a high performance edge and texture extraction algorithm. To proceed, we consider a dispersive medium that is linear and lossless and therefore, $\alpha = \gamma = 0$. Thereby, the NLSE can be written as:

$$\frac{\partial A(x,y;z)}{\partial z} = \frac{i\beta_2}{2} \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right) A(x,y;0) \qquad (2.30)$$

where $A(x,y;z)$ is the solution of the reduced NLSE in the dispersive regime. As shown previously, the solution is computed as follows:

$$A(x,y;z) = A(x,y;0) + i\frac{z}{2} \left(\frac{\partial^2 A(x,y;0)}{\partial x^2} + \frac{\partial^2 A(x,y;0)}{\partial y^2} \right) \qquad (2.31)$$

Next we convert the above continuous domain 2D equation to discrete domain 2D equation as follows:

$$A_0[m,n] = IFFT^2 \left(FFT^2\{A[m,n]\} \cdot \phi[u,v] \right) \qquad (2.32)$$

In the above equations, $A_0[m,n]$ is the output image, n and m are the spatial variables, FFT^2 is the two dimensional Fast Fourier Transform, and u and v are spatial frequency variables. The function $\phi[u,v]$ is the discrete domain dispersive phase propagator and FFT, $IFFT$ is Fast Fourier Transform and Inverse Fast Fourier Transform respectively. For simplicity, we first analyze 1D discrete dispersive operation defined as:

$$A_0(n) = IFFT \left(FFT\{A(n)\} \cdot \phi(u) \right) \qquad (2.33)$$

such that the discrete phase propagator is defined as quadratic function of frequency given by:

$$\phi(u) = exp(s \cdot u^2) \qquad (2.34)$$

such that s is a strength variable controlling the amount of phase applied. We substitute this phase as follows:

$$A_0(n) = IFFT\left(FFT\{A(n)\} \cdot exp(s \cdot u^2)\right)$$
$$= IFFT\left(FFT\{A(n)\} \cdot (cos(s \cdot u^2) + i \cdot sin(s \cdot u^2))\right) \quad (2.35)$$

As shown before, for the dispersive regime to produce edge output, the value of s is very small (i.e. phase is small). As such, we can rewrite the above equation as:

$$A_0(n) = IFFT\left(FFT\{A(n)\} \cdot (1 + i \cdot s \cdot u^2)\right) \quad (2.36)$$

And by implementing the Fourier transform property of differentiation, we can compute the output as follows:

$$A_0[n] = A[n] + i \cdot s \cdot \frac{d^2 A[n]}{dn^2} \quad (2.37)$$

As can be seen from the above results, the real part of the dispersive propagation output is same as the input where as the imaginary component is the second order derivative of the input. As discussed previously, Phase Stretch Transform [22–27] is a physics-inspired edge detection that emulates light propagation through a dispersive medium. PST operator is defined as the phase of the dispersive output as follows:

$$PST\{A[n]\} \triangleq \angle\{A_0[n]\}$$
$$= tan^{-1}\left(\frac{s \cdot \frac{d^2 A[n]}{dn^2}}{A[n]}\right) \quad (2.38)$$

where $\angle\{\}$ is the angle operator. The PST operator computes edge information from the input digital image as shown in Figure 2.13. And finally, for the case dispersive and nonlinear regime, the solution to the NLSE equation can be obtained using Split-Step Fourier Method (SSFM), where first the nonlinear part is solved independently and then the solution obtained is used as input to solve the dispersive part. This results in an enhanced edge map as shown in Figure 2.13.

43

2.7   Conclusion

In this chapter, we have outlined a computational image processing pipeline inspired by the Nonlinear Schrodinger Equation (NLSE). The technique is implemented in such a way that it enhances both the color as well as contrast of the incoming frame/image. The technique works by subjecting the image to virtual propagation through a two-dimensional NonLinear Schrodinger Equation (NLSE). The propagation imparts a brightness-dependent pixelwise phase onto the image transforming it into a complex-valued quantity containing both real and imaginary components. The effect is a type of Self Phase Modulation (SPM) and more generally an amplitude-modulation (AM) to phase-modulation (PM) conversion (AM to PM conversion). The Imaginary component of the output is the desired contrast-enhanced image. In one implementation, the parameters of the NLSE equation are manually set.

The second implementation utilizes artificial intelligence where the parameters of the Equation are learned by a neural network in an unsupervised fashion and are then used to compute the output image. The concept is validated with numerous images and videos illustrating benefits of the disclosed technology in different applications and under various conditions. Experimental results on standard datasets show superior computational efficiency compared to the state-of-the-art image enhancement methods. Finally, we also show that our method achieves best-in-class performance for color gamut expansion.

CHAPTER 3

Feature Enhancement in Visually Impaired Images

One of the major open problems in computer vision is feature detection in visually impaired images. In this Chapter, we describe a potential solution using Phase Stretch Transform, a new computational approach for image analysis, edge detection and resolution enhancement that is inspired by the physics of the photonic time stretch technique. We mathematically derive the intrinsic nonlinear transfer function and demonstrate how it leads to (1) superior performance at low contrast levels and (2) a re-configurable operator for hyper-dimensional classification. We prove that the Phase Stretch Transform equalizes the input image brightness across a range of intensities resulting in high dynamic range of operation for feature detection in visually impaired images. We also show further improvement in the dynamic range by combining our method with the conventional techniques. Finally, our results propose a new paradigm for the computation of mathematical derivatives via group delay dispersion operations.

3.1 Introduction

Feature detection in images plays a critical role in the field of computer vision for solving problems associated with object recognition, image registration, content-based image retrieval and deep learning [64–66]. Prior works for improving feature detection in images have focused on the use of grey level statistics of the image [64]

and on application of edge detection methods [67]. Color distinctiveness and color models [68, 69] and scale selections [70] in images have also been exploited for enhancing the feature detection. The main goal of feature detection is to classify objects more accurately and at the same time be robust to varying viewing conditions that include changes in illumination, environmental conditions, object orientation, and the zoom factor of the camera. Environmental conditions can severely impair detection and localization of objects in images. For instance, under foggy conditions, acquired images suffer from visual impairments such as reduced contrast, blur and noise which leads to lower resolution [71, 72]. This constitutes a major bottleneck for many computer vision applications including autonomous vehicles. The emerging imaging technologies such as High Dynamic Range (HDR) hold promise to solve feature detection problems in the field of computer vision. However, the slow frame rate of these technologies restricts their practice in self-driven cars, autonomous robotics and other real-time applications.

The Phase Stretch Transform (PST) was recently introduced as a computational approach for signal and image processing [27, 28]. PST is a physics-based algorithm that has its roots in photonic time stretch technique [1, 32–34, 73], a method for real-time measurements of ultra-fast events and one that has enabled the discovery of optical rogue waves [3], observation of relativistic electron structure [4], label-free cancer cell detection with record accuracy [8, 9] and optical data compression [36]. The algorithm mimics the propagation of electromagnetic waves through a diffractive medium with engineered 3D dispersive property (refractive index) [27, 28]. This optics-inspired algorithm has superior properties that can be exploited to develop advanced algorithms for feature extraction from digital images as shown previously in [22, 25, 29, 30]. Here, we discuss in detail the nonlinear behavior of PST and demonstrate how this behavior can be used to solve problems related to feature detection for computer vision applications. PST can be applied to both digital images as well as time series data [73] and has been used for

46

edge detection in biomedical images to develop diagnostic assistant tools [23, 24] and Synthetic Aperture Radar (SAR) images [74]. PST has also been applied for resolution enhancement in super-resolution localization microscopy for imaging of a single molecule [26]. The transform drastically improved the localization of point spread function, reduced the computational time by 400% and increased the emitter density by the same amount. The algorithm has been open sourced on GitHub and Matlab Central File Exchange [75] and has received extraordinary endorsements both by the software as well as image processing community. The transform exhibits superior performance over conventional derivative based edge operators in particular for visually impaired images. It is able to reveal features invisible to human eye and to conventional algorithms used today. Because of these unique intrinsic properties offered by PST, it has promising application for feature enhancement in visually impaired images which is discussed in detail in this Chapter.

We first show visually that PST has an inherent equalization ability that gives a response ideal for feature detection in low contrast regimes of visually impaired images. To do this, we apply our edge detection algorithm on two road traffic images taken under foggy conditions, as shown in Figure 3.1. The figure depicts how our edge detection algorithm could significantly improve the feature detection in case of visually impaired images by outperforming the conventional edge detection methods based on derivative of the image. The conventional derivative based method is unable to capture details with small contrast in the bright but low resolution areas of the image whereas our technique successfully detects features in these low contrast visually impaired regions of the image. The warp and strength parameters of the PST kernel as described in [27, 28] for feature detection in these images are 22 and 500, respectively. As we will show in our mathematical formulations, this property emerges because PST's transfer function has an inherent equalization ability, derived analytically in the next section. Finally, we

(A) Original image   **(B) Feature detection using smooth derivative**   **(C) Feature detection using PST**
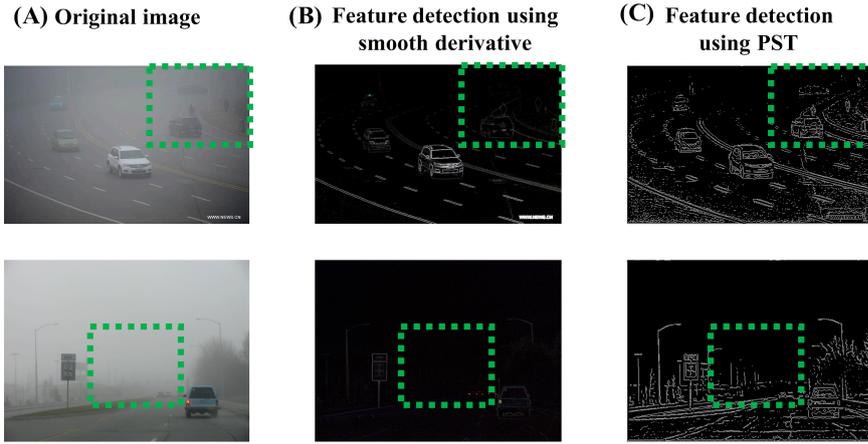
Figure 3.1: **Comparison of feature detection using conventional derivative based edge operator to the case of feature detection using Phase Stretch Transform (PST) in case of visually impaired images.** Original traffic images taken in a foggy weather are shown in (A). Detected features using conventional derivative based edge operator and PST operator are shown in (B) and (C), respectively. It can be seen that the conventional derivative based edge detection operator fails to visualize the low contrast details in the visually impaired regions of the images (as shown in green dashed boxes). However, PST captures these low contrast details in the low resolution regions (as shown in green dashed boxes) due to its unique re-configurable mechanism that detects features over a wide dynamic range. The strength of features detected using PST over both low and high resolution regions of the images is consistent unlike derivative operator as also shown in [25].

demonstrate the superior performance of PST at low light levels and its application to HDR images. Towards the end, we also propose a new paradigm for the computation of mathematical derivatives via group delay dispersion operations that has potential applications in optical computing.

3.2 From Optical Physics to Digital Algorithms

Photonic time stretch technique can be understood by considering the propagation of an optical pulse through a dispersive optical fiber. The optical pulse propagation in an optical fiber is governed by non-linear Schrodinger equation as discussed in [76]. By disregarding the loss and non-linearity in an optical fiber and considering only the group velocity dispersion, this equation upon integration reduces to:

$$E_o\left(z,t\right) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \tilde{E}_i\left(0,\omega\right) \cdot \left[e^{\frac{-j\beta_2 z \omega^2}{2}} \right] \cdot e^{j\omega t} d\omega \qquad (3.1)$$

where $\tilde{E}_i(0,\omega)$ is the input pulse spectrum, β_2=GVD parameter, z is propagation distance, $E_o\left(z,t\right)$ is the reshaped output pulse at distance z and time t. The response of a dispersive element in a time-stretch system can be approximated as a phase propagator as presented in [73, 77],

$$H\left(\omega\right) = e^{i\phi(\omega)} = e^{i\sum_{m=0}^{\infty} \phi_m(\omega)} = \prod_{m=0}^{\infty} H_m\left(\omega\right) \qquad (3.2)$$

Therefore, Eq. 3.1 for a pulse that propagates through the time-stretch system and is reshaped into a temporal signal with a complex envelope can be written as follows:

$$E_o\left(t\right) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \tilde{E}_i\left(\omega\right) \cdot H\left(\omega\right) \cdot e^{j\omega t} d\omega \qquad (3.3)$$

By considering sufficient linear dispersion, the stationary phase approximation can be satisfied resulting in a mapping of spectrum to time. Using this spectro-temporal mapping, we can evaluate the sparsity in the spectrum of a signal [73]. In particular, one can create an information gearbox for matching the time-bandwidth of fast real-time optical data to that of the much slower electronics. These photonic hardware accelerators proposed in [73] act as a means to boost the speed and reduce the power consumption of electronics.

The time stretch operation $\mathbb{S}\left\{E_i\left(t\right)\right\}$ on an input pulse $E_i\left(t\right)$ defined above in terms of phase can be extended to operation on amplitude of a signal in optical

domain as following:

$$\mathbb{S}\left\{E_i\left(t\right)\right\} = \int_{-\infty}^{+\infty} F\left\{E_i\left(t\right)\right\} \cdot \left[e^{j\phi(\omega)} \cdot \tilde{L}\left(\omega\right)\right] \cdot e^{j\omega t} d\omega \tag{3.4}$$

where $e^{j\phi(\omega)}$ is the phase filter and $\tilde{L}\left(\omega\right)$ is the amplitude filter. And for a discrete signal, the stretch operation can be defined as:

$$\mathbb{S}\left\{E_i\left[t\right]\right\} = IFFT\left\{FFT\left\{E_i\left[t\right]\right\} \cdot \tilde{K}\left[\omega\right] \cdot \tilde{L}\left[\omega\right]\right\} \tag{3.5}$$

where $\tilde{K}\left[\omega\right]$ is the phase filter or propagator, $\tilde{L}\left[\omega\right]$ is the amplitude filter, IFFT is Inverse Fast Fourier Transform and FFT is Fast Fourier Transform.

The dispersion operation on a 1D temporal signal is equivalent to parallax diffraction on a 2D input space. This led us to analyze the application of above mentioned dispersion-based stretch operations to digital images which resulted in an optics-inspired edge detection algorithm called Phase Stretch Transform (PST) [27, 28]. PST is a qualitatively new method for feature engineering and is discussed at length in the next section.
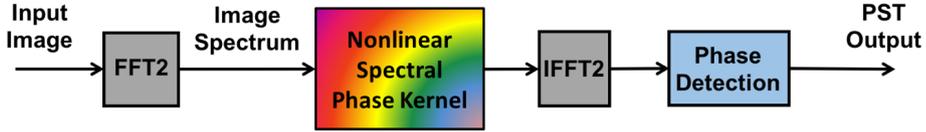
Figure 3.2: **Operation Principle of Phase Stretch Transform (PST)**. PST is a physics-inspired feature detection method that transforms the image by emulating propagation through a dispersive medium [27, 28]. The input image is processed in frequency domain. The nonlinear spectral phase kernel encodes frequency components into the spatial phase of the output image such that high frequency components have higher phase. Upon phase detection of the output followed by thresholding operation, high phase values corresponding to high frequency components survive, producing a edge map.

3.3 Phase Stretch Transform

Phase Stretch Transform (PST) is a recently introduced computational approach for signal and image processing that emerged out of the research on Photonic Time Stretch. This algorithm transforms an image by emulating propagation of electromagnetic waves through a diffractive medium with an engineered 3D dispersive property (refractive index profile) [27, 28].

As shown in Figure 3.2, by applying a nonlinear spectral phase kernel to the image spectrum (by operating 2D Fast Fourier Transform (FFT2) on the image), the frequency distribution is mapped to the spatial phase of the output image. The nonlinear phase kernel encodes a high phase value to high frequency component. The spatial output after the 2D Inverse Fast Fourier Transform (IFFT2) is no longer a real quantity but instead has a complex phase associated with it. Upon phase detection of the output followed by thresholding operation, high phase values corresponding to high frequency components survive. Hence, edges i.e. high frequency components in an image are detected. The time stretch stretch operation $\mathbb{S}\{\}$ for an image can be represented as follows:

$$\mathbb{S}\{E_i[x,y]\} = IFFT^2\left\{FFT^2\{E_i[x,y]\} \cdot \tilde{K}[u,v] \cdot \tilde{L}[u,v]\right\} \qquad (3.6)$$

In the above equations, $E_i[x,y]$ is the input image, x and y are the spatial variables, FFT^2 is the two-dimensional Fast Fourier Transform, $IFFT^2$ is the two-dimensional Inverse Fast Fourier Transform and, u and v are spatial frequency variables. The function $\tilde{K}[u,v]$ is called the warped phase kernel and the function $\tilde{L}[u,v]$ is a localization kernel implemented in frequency domain for image processing. PST operator is defined as the phase of this Warped Stretch Transform output as follows,

$$PST\{E_i[x,y]\} \triangleq \measuredangle\langle \mathbb{S}\{E_i[x,y]\}\rangle \qquad (3.7)$$

where $\measuredangle\langle\cdot\rangle$ is the angle operator.

(A) Original Image   **(B) Feature detection using smooth derivative**   **(C) Feature detection using PST**
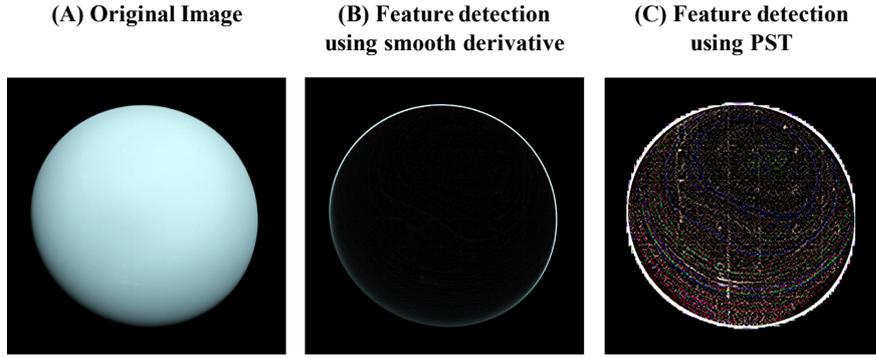
Figure 3.3: **Comparison of feature detection using conventional derivative based edge operator to the case of feature detection using Phase Stretch Transform (PST).** The derivative is the fundamental operation used in the popular Canny, Sobel and Prewitt edge detection methods. These derivative based methods are unable to capture the low contrast on the surface of the Uranus planet. On the other hand, PST extracts these surface variations efficiently as shown in [22, 29, 30]

PST has unique intrinsic properties which are not offered by the state-of-the-art algorithms. To validate this claim, we refer to Figure 3.3 that shows an image of the planet Uranus processed by the conventional derivative based edge operator and by the PST. The derivative method is the underlying function utilized by the popular Canny, Sobel and Prewitt algorithms. The result clearly shows the dramatic advantage offered by the optics-inspired PST. The inherent equalization ability of PST gives a response ideal for feature detection in low contrast visually impaired images. Figure 3.4 compares the effect of feature detection using conventional derivative based edge operator with feature detection using PST on another image of planet Uranus captured from a different viewing angle. Conventional derivative based operator fails to visualize the low contrast in the bright areas of the image over the surface of the planet Uranus. However, PST can clearly show these small intensity changes even in the intensity-saturated areas due to its natural equalization mechanism. These surface variations over the planet are consistent with the edges detected in the Fig 3 highlighting the efficiency of PST.

| (A) Original Image | (B) Feature detection using smooth derivative | (C) Feature detection using PST |
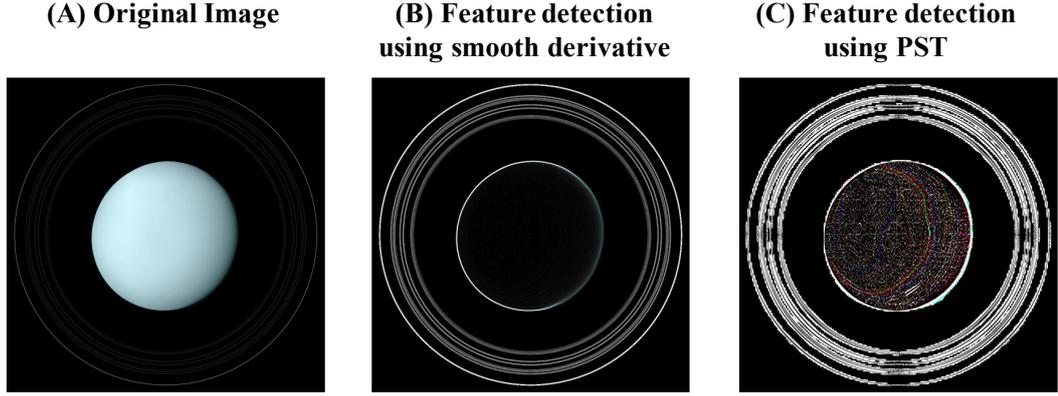
Figure 3.4: **Comparison of feature detection using conventional derivative based edge operator to the case of feature detection using Phase Stretch Operator (PST) on an image of the planet Uranus captured from a different angle of view as compared to the one shown in Figure 3.3.** Original image is shown in (A). Results of feature detection using conventional derivative based edge operator and PST operator are shown in (B) and (C), respectively. PST is able to locate the low contrast on the surface of the planet which are consistent with the edges located in Figure 3.3 as shown in [22, 29, 30]

3.4 Mathematical foundations of Phase Stretch Transform

The superior performance of Phase Stretch Transform (PST) in the low contrast regime is proved here mathematically by deriving closed-form analytical expressions for its transfer function. Mathematical results reveal that the transform has an inherent intensity equalization property leading to high dynamic range performance. Analytical results are also supported by numerical simulations confirming the dynamic range enhancement. As define previously, the stretch operator $\mathbb{S}\{\}$ on an input image act as follows:

$$E_o[x, y] = S\{E_i[x, y]\} \triangleq IFFT2\left\{\tilde{K}[u, \ v] \cdot \tilde{L}[u, v] \cdot FFT2\{E_i[x, y]\}\right\} \quad (3.8)$$

where $E_o[x, y]$ is a complex quantity defined as,

$$E_o[x, y] = |E_o[x, y]| \; e^{j\theta[x,y]} \qquad (3.9)$$

Here for simplicity, we assume the localization kernel $\tilde{L}[u, v] = 1$. Now, without the loss of generality and in order to keep the notations manageable in what follows, we consider operation of PST on 1D data, i.e.,

$$PST\{E_i[x]\} \triangleq \measuredangle\{S\{E_i[x, y]\}\} = \measuredangle\{E_o[x]\} = \measuredangle\left\langle IFFT\left\{\tilde{K}[u] \cdot FFT\{E_i[x]\}\right\}\right\rangle \qquad (3.10)$$

The warped phase kernel $\tilde{K}[u]$ is defined by a phase function that has a nonlinear dependence on frequency, u,

$$\tilde{K}[u] = e^{j \cdot \varphi[u]} \qquad (3.11)$$

By expanding the phase term in the warped phase kernel $\tilde{K}[u]$ using Taylor series we have,

$$\tilde{K}[u] = e^{\left(j \sum_{m=2}^{M} \frac{\varphi^{(m)}}{m!} u^m\right)} \qquad (3.12)$$

where $\varphi^{(m)}$ is the mth- order discrete derivative of the phase $\varphi[u]$ evaluated for $u = 0$ and values of m are even numbers. PST phase term $\varphi[u]$ only contains even-order terms in its Taylor expansion due to the even symmetry of the phase term $\varphi[u]$ as first considered in [27, 28]. By using the expansion of warped phase kernel as described in Eq. 3.12, output complex-field, $E_o[x]$, can be calculated as follows,

$$E_o[x] = IFFT\left\{\tilde{E}_i[u] \quad \times \quad \tilde{K}[u]\right\}$$

$$= IFFT\left\{\tilde{E}_i[u] \quad \times \quad e^{\left(j \sum_{m=2}^{M} \frac{\phi^{(m)}}{m!} u^m\right)}\right\} \qquad (3.13)$$

where $\tilde{E}_i[u]$ is the spectrum of the input computed using Fast Fourier transform (FFT). Simulation show that when the applied phase is small, PST works best. Under these conditions, we can use small value approximation for the applied phase kernel. Therefore, the phase term in Eq. 3.13 can be simplified to,

$$E_o[x] = IFFT\left\{ \tilde{E}_i[u] \times \left[1 + j\left(\sum_{m=2}^{M} \frac{\phi^{(m)}}{m!} u^m \right) \right] \right\} \qquad (3.14)$$

$$\rightarrow \qquad E_o[x] \approx \left[1 \times E_i[x] + j \sum_{m=2}^{M} \frac{(-1)^{m/2}\ \phi^{(m)}}{m!\ (2\pi)^m} E_i[x]^{(m)} \right] \qquad (3.15)$$

where $E_i[x]^{(m)}$ is the mth-order discrete mathematical derivative of the input $E_i[x]$. As the input is a real quantity, we can calculate the output phase as,

$$PST\{E_i[x]\} = \measuredangle \{E_o[x]\}$$

$$\approx \quad \tan^{-1}\left\{ \frac{\sum_{m=2}^{M} \frac{(-1)^{m/2}\ \phi^{(m)}}{m!\ (2\pi)^m} E_i[x]^{(m)}}{E_i[x]} \right\} \qquad (3.16)$$

Finally, since the phase is restricted to small values ($\tan^{-1}\theta \approx \theta$), above equation can be simplified to,

$$PST\{E_i[x]\} \approx \left\{ \frac{\sum_{m=2}^{M} \frac{(-1)^{m/2}\ \phi^{(m)}}{m!\ (2\pi)^m} E_i[x]^{(m)}}{E_i[x]} \right\} \qquad (3.17)$$

We see that the transfer function of PST is consisting of a summation of even order mathematical derivatives of the input in the numerator divided by the input amplitude (brightness) in the denominator. A hyper dimensional feature set corresponding to different measures of the curvature of the edge is computed in the numerator while the denominator renders the response nonlinear in such a way that low-light-levels in the input are enhanced. Above analytical results are

55

derived by considering a general expression for the phase kernel and only when the applied phase of the PST kernel is small. We now expand our analytical findings by considering different scenarios that reveal further insights into the unique properties of PST.

Case 1: Let's consider the Phase Kernel $\tilde{K}[u]$ as a quadratic function of frequency variable u. Under this condition $\tilde{K}[u] = u^2$ and by using small phase approximation, the phase term in Eq. 3.14 can be simplified to,

$$E_o[x] = IFFT\left\{ \ \tilde{E}_i[u] \ \times \ [1 + j \ (u^2)] \ \right\} \tag{3.18}$$

$$\rightarrow E_o[x] \approx \left[1 \ \times \ E_i[x] - \ j\frac{1}{(2\pi)^2} * \frac{d^2 E_i[x]}{dx^2}\right] \tag{3.19}$$

We also assume that the phase of the complex output is restricted to small values. Therefore, phase of the output can be simplified to,

$$PST\{E_i[x]\} = \angle E_o[x] \approx \frac{\frac{-1}{(2\pi)^2} * \frac{d^2 E_i[x]}{dx^2}}{E_i[x]} \tag{3.20}$$

Case 2: We consider here the same Phase Kernel (as a quadratic function of frequency variable u), $\tilde{K}[u] = u^2$ as discussed in Case 1. However, we do not restrict to small phase approximation. The exponential term in Eq. 3.13 can now be represented as,

$$E_o[x] = IFFT\left\{ \ \tilde{E}_i[u] \ \times \ [\cos(u^2) + j \ sin(u^2)] \ \right\} \tag{3.21}$$

We expand the sine and cosine terms using Euler expansion up to third order and then by applying small value approximation to the complex output of Eq. 3.16, the PST output can be computed as shown below

56

$$PST\left\{E_i\left[x\right]\right\} = \angle E_o\left[x\right] \approx \frac{\frac{-1}{(2\pi)^2} * \frac{d^2 E_i[x]}{dx^2} \quad + \quad \frac{1}{3!(2\pi)^6} * \frac{d^6 E_i[x]}{dx^6} \quad - \quad \frac{1}{5!(2\pi)^{10}} * \frac{d^{10} E_i[x]}{dx^{10}}}{E_i\left[x\right] \quad - \quad \frac{1}{2!(2\pi)^4} * \frac{d^4 E_i[x]}{dx^4} \quad + \quad \frac{1}{4!(2\pi)^8} * \frac{d^8 E_i[x]}{dx^8}}$$

$$(3.22)$$

The closed-form expression of the transfer function of PST shown in Eq. 3.16 relates the output to the input in the case of an arbitrary phase kernel valid under small phase approximation. For certain scenarios, the core functionality of PST as a feature detector can be established by the closed-form expression presented in Eq. 3.16. As we can see, the output of the PST operator is directly proportional to the even-order derivatives of the input with weighting factors of $\frac{(-1)^{m/2}}{m!} \frac{\phi^{(m)}}{(2\pi)^m}$. Each computed mathematical derivative highlights a different feature of the input. The weighting factors can be modified to selectively enhance features of interest. Hence, our transform acts as a re-configurable operator that can be tuned to emphasize different features in an input image.

The intrinsic equalization ability of PST can be seen in Eq. 3.16. The PST output is inversely related to the input brightness level valid under small phase approximation. Therefore, for a same contrast level, the PST output is higher in dark low-light-level areas of an image. This important observation from Eq. 3.16 confirms the fact that PST has an inherent property to equalize the detected output with the input brightness level and therefore, allows for a more sensitive feature detection as also reported in [25, 30].

Brightness level equalization is a well-studied technique to improve feature detection algorithms in High Dynamic Range (HDR) images (see [78] for example). One approach to achieve brightness level equalization in images is by applying a logarithmic (log) function to the input before application of feature detection algorithms. By applying a log function, we can achieve high gain for low brightness input. This brightness equalization results in a more sensitive feature detection. Fortunately, the PST operator has a built-in logarithmic behavior which gives it

excellent dynamic range. However, this property does not completely describe the functionality of our transform. As observed in Eq. 3.16, our transform computes a hyper-dimensional feature set for signal classification (shown in Figure 3.5). These results also demonstrate a method for computation of mathematical derivatives via group delay dispersion operations.
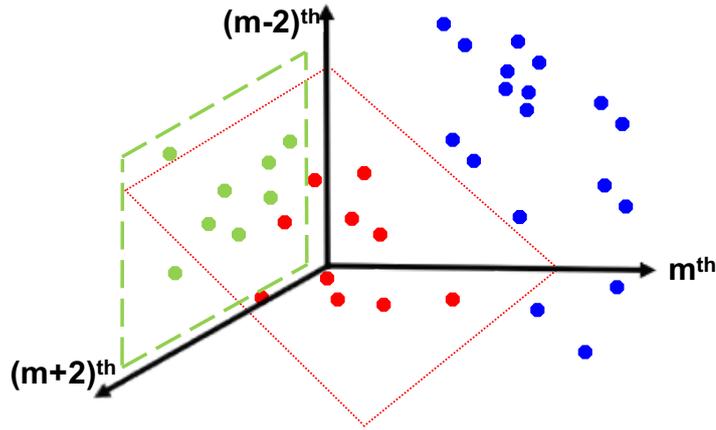
Figure 3.5: **Phase Stretch Transform (PST) as a hyper dimensional classifier.** PST operator can act as a reconfigurable operator to compute mth-order derivative. Here the dimensions are the order (even) of derivatives and hyper planes are shown as green and red boxes.

3.5 Clinical Decision Support Systems using PST

Medical images act as a very important source of information in order to understand the anatomy and organ function as well as aid in diagnosis of diseases. Lately, machine learning tools are unified with image processing techniques for application to medical imaging leading to production of computer-aided diagnostics (CAD) and decision making tools. As discussed previously, PST has superior performance over conventional edge detectors, therefore, its application to medical images for feature detection is promising for accurate segmentation and development of CAD tools.

One such application of PST is to develop a diagnostic assistant tool for pneumothorax [23,24]. Pneumothorax is a medical situation in which air leaks into the space between the lungs and the chest wall due to a chest injury or a lung disease or even sometimes due to certain medical procedures. The major risk factor that increases the mortality caused by pneumothorax is the failure to identify it at any early examination by a radiologist. By the application of PST to the chest x-rays' of patient suffering from pneumothorax, the boundary of collapsed lung is easily located otherwise difficult to be identified during an initial visual inspection by a radiologist as shown in Figure 3.6. This tool, which is first of its kind, traces the collapsed lung and aids the radiologist to take correct decision in this life-critical examination. With pneumothorax being very common in ventilated critically ill patients, it becomes important to develop tools for accurate diagnosis as failures in diagnosis can cause life threatening complications [79].
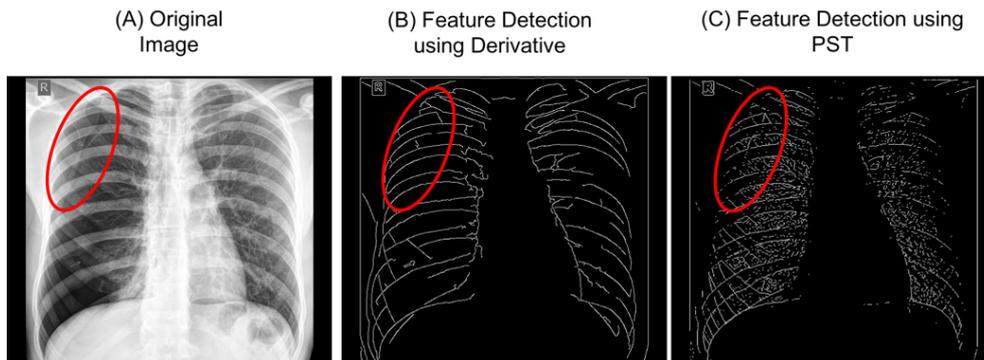
Figure 3.6: **Comparison of feature detection using Phase Stretch Transform (PST) algorithm with the features detected using a conventional derivative based edge operator for a lung x-ray image.** In the figure, (A) X-ray of lung of a patient suffering from pneumothorax (B) Edge detection using conventional derivative based edge operator (C) Edge detection using PST. The red oval indicates the region of collapsed lung. As shown, PST traces the low contrast lung edge with an equalized response due to intrinsic nonlinear behavior. Conventional derivative based edge operator work well only for high contrast regions and therefore, fails to locate the collapsed lung boundary as described in [23,24].
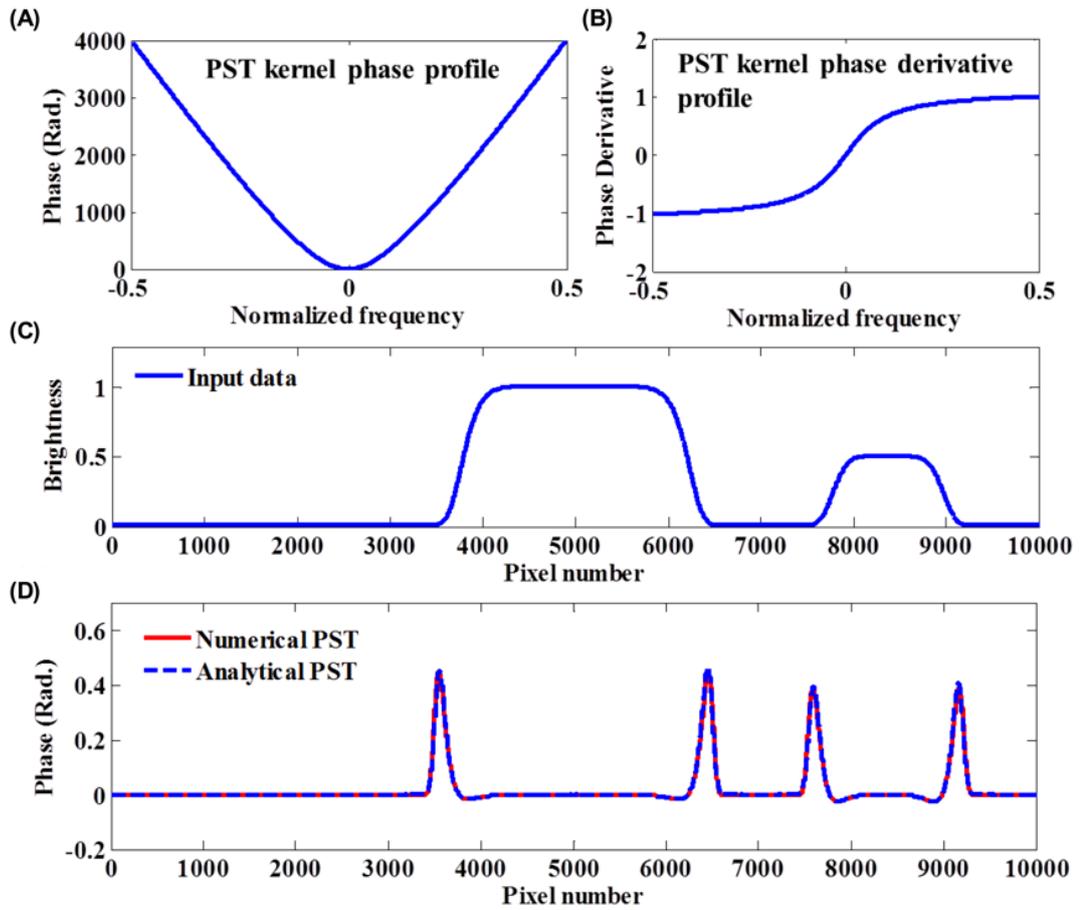
Figure 3.7: **Comparison of numerically simulated output of Phase Stretch Transform (PST) algorithm with the output given by the closed-form analytical expression derived in Eq.** **3.16.** The phase kernel of the PST operator and the corresponding derivative profile of the phase kernel are shown in (a) and (b), respectively. The simulated input 1D brightness data is shown in (c). Numerically calculated PST output data and the output data estimated by the closed-form expression of the PST transfer function, as derived in Eq. 3.16, is shown in (d) using red-solid and blue-dotted lines, respectively. The above simulation result validates the accuracy of the closed-form analytical expression of PST transfer function derived in Eq. 3.16

3.6   Simulation Results

In this section, we present simulation results that validate the closed-form analytical expression of the transfer function of PST derived in the previous section. In order to reinforce the new theory developed above, we also demonstrate several visual examples of operation of PST on HDR images. In our first simulation result, we show the PST output for a given 1D data and compare it to the output computed by the analytical expression derived in Eq. 3.16. Figure 3.7A shows the phase kernel $\varphi[u]$ of the PST operator designed for this simulation. The warp W, and strength S, parameters of the phase kernel are 12.5 and 4000, respectively. The derivative of the PST phase kernel is shown in Figure 3.7B. The simulated 1D input data is shown in Figure 3.7C. In Figure 3.7D, numerically simulated PST output is compared to the output estimated by the analytical expression of Eq. 3.16 using red-solid and blue-dotted lines, respectively. It is evident that the simulations match analytically derived results validating the accuracy of the closed-form analytical model of our algorithm as derived in Eq. 3.16.

In the next simulation example, we examine the effect of PST operation for feature detection on a signal with varying contrast levels at a constant brightness level and compare it to the case of using the conventional technique of differentiation to detect features in the same input. The warp W, and strength S, parameters of the PST operator are 12.15 and 0.48, respectively. The input was designed to have varying contrast levels at a constant brightness level, as shown in Figure 3.8A. We compare numerically simulated PST output to the output using differentiation in Figure 3.8B. As expected, the output of the differentiator is linearly proportional to the contrast level and is insensitive to the input brightness level. On the other hand, PST output is directly dependent on the input contrast level but inversely dependent on the constant brightness level. This nonlinear behavior is due to the inherent equalization mechanism of PST as described in Eq. 3.16.

61
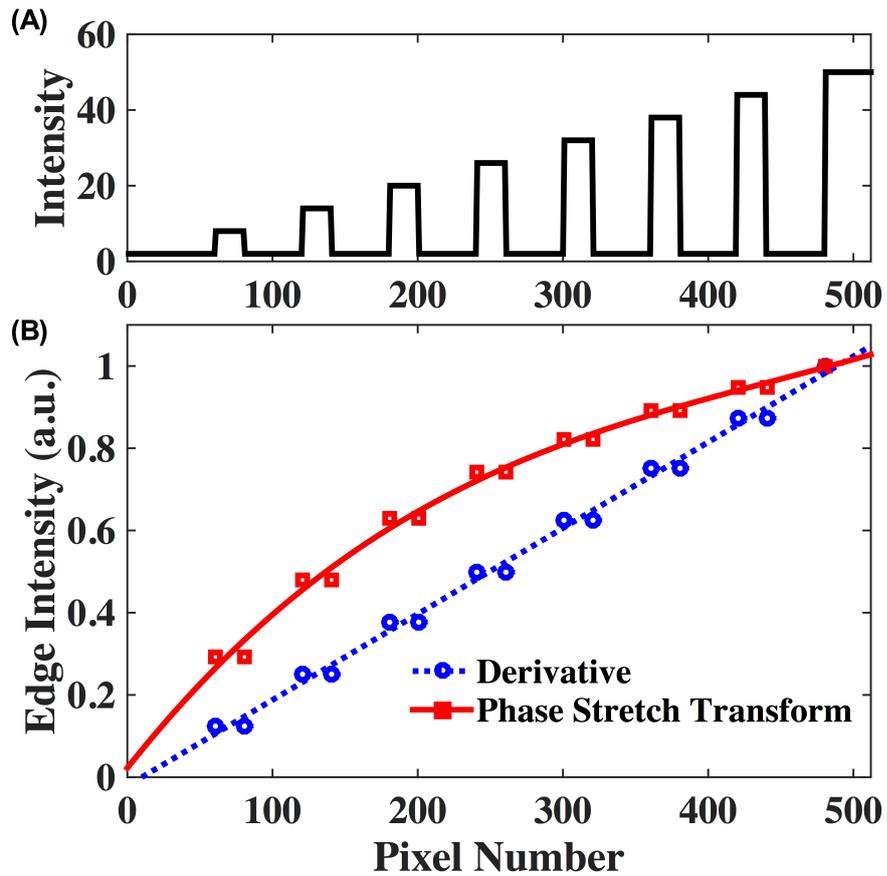
Figure 3.8: **Effect of Phase Stretch Transform (PST) on an input signal with various contrast levels at a constant brightness level.** The input 1D data which is designed to have various contrast levels at a constant brightness level is shown in (a). Numerically calculated PST output is compared to the output using differentiation in (b) using red-solid and blue-dotted lines, respectively. As can be seen, the output of the differentiator has a linear response to contrast level in the input and is completely insensitive to the input brightness level. On the other hand, PST output is non-linearly related to the contrast level in the input at fixed brightness

The behavior of PST operation for feature detection on a signal with a constant contrast level and various brightness levels is evaluated in Figure 3.9. The input data, designed to have a constant contrast level and various brightness levels, is shown in Figure 3.9B. The warp, W, and strength, S, factors used for the PST
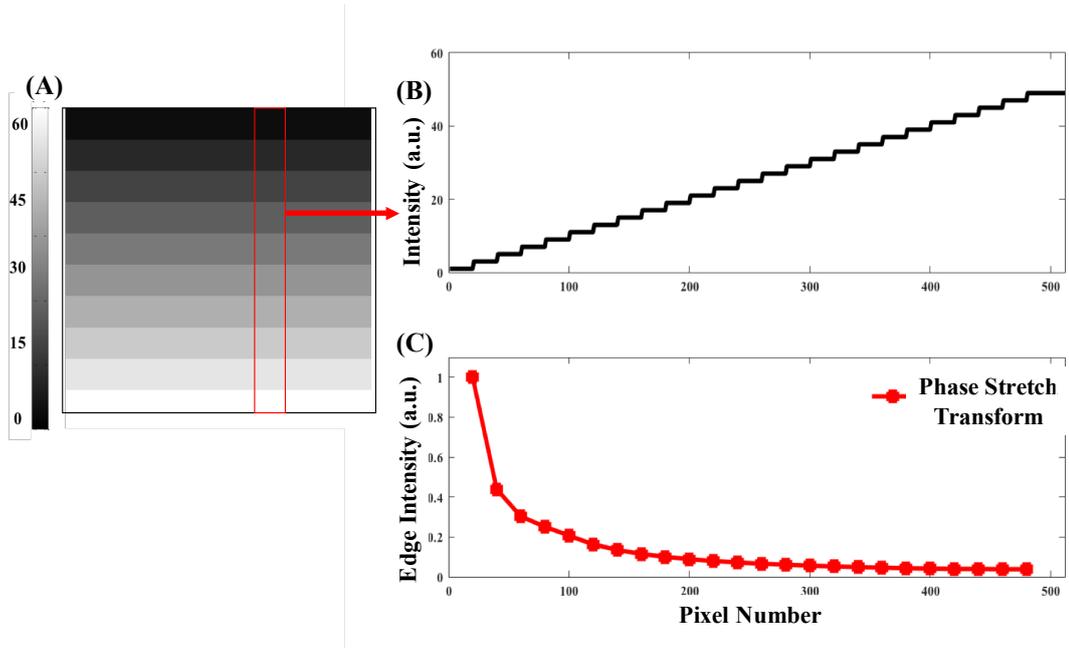
Figure 3.9: **Effect of Phase Stretch Transform (PST) on an input signal with a constant contrast level and various brightness levels.** The input data which is designed to have a constant contrast level and various brightness levels is shown in (B). Numerically calculated PST output data for feature detection in a signal with a constant contrast level varies with the input brightness level is shown in (C). This shows that PST output has an inverse dependence on the input brightness level.

operator are 12.15 and 0.48, respectively. The red solid line representing the output data confirms that the relation of PST to same contrast level at various brightness levels, is logarithmic as also estimated in Eq. 3.16. Therefore, the simulation result presented in Figure 3.9 further reinforce the accuracy of the closed-form equation to estimate the output of the PST algorithm [22, 29, 30], validating that PST has an inverse dependence on the input brightness level.

63

(A) Original image

(B) Feature detection using smooth derivative

(C) Feature detection using PST

Figure 3.10: **Comparison of feature detection using smooth derivative operator to the case of feature detection using Phase Stretch Transform (PST).** Original image is shown in (A). Smooth derivative operator is unable to efficiently visualize the low contrast details in the dark areas of the image. However, PST captures these contrast changes in low-light-level areas due to its intrinsic equalization property.

Figure 3.10 shows a visual example of using PST for feature enhancement on a 14 bit HDR image. The image has features of interest in extremely low-light-level regions, as seen in the red boxes. We now compare the performance of the derivative operator with PST for feature detection. The derivative operator was implemented from native smooth derivative function. For a fair comparison, both methods use the same localization kernel (a Gaussian function) with sigma factor of 2. The warp W, and strength S, parameters used for the PST operator are 12.15 and 0.48, respectively. Results of feature detection using smooth derivative operator and PST operator are shown in Figure 3.10B and Figure 3.10C, respectively.

The derivative operator is unable to unveil the small contrast details in the dark areas of the image, as can be seen in Figure 3.10B . However, PST extracts these low contrast details in dark areas due to its natural equalization mechanism, see dashed box in Figure 3.10C. It also can be observed that the intensity of detected edges in the case of smooth derivative is related linearly to the brightness level of the original image, compare solid box areas in Figure 3.10A and Figure 3.10B.

In contrast, PST has automatically equalized the brightness level in the solid box region in the image and outputs relatively constant feature intensity for that region, see Figure 3.10 C. We also note that PST has failed to visualize features in high contrast areas in the image. This is because of the inverse dependence on brightness level as derived in Eq. 3.16. This issue can be mitigated by setting a higher maximum threshold for detected features or by equalizing the image brightness before passing through PST operator.

We further examine the role of PST for feature detection in low-light-level and high-light-level regions by considering a line scan of a HDR image shown previously in Figure 3.10. The blue box in the Figure 3.11 demonstrates the response of PST to low-light-level regions where it outperforms conventional derivative based edge operator. Similarly, for high-light-level regions of the image (shown in green and
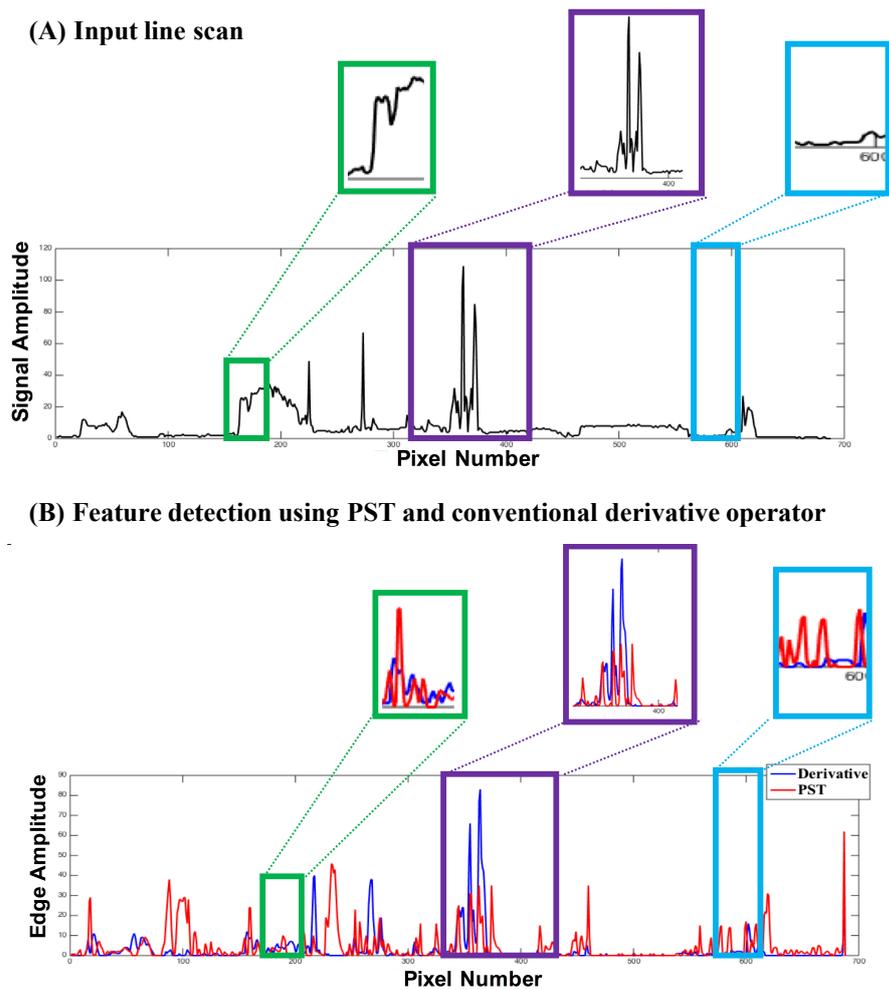
Figure 3.11: **Comparing feature detection using conventional derivative based edge operator and Phase Stretch Transform (PST) operator under low-light-level and high-light-level conditions.** (A) Original input line scan corresponding to Row 524 from the image shown in Figure 3.10. Feature detection of this input line scan using the derivative and the PST operator is shown in (B). The blue box demonstrates that the response of PST is higher than the derivative operator under low-light-level conditions. The green and purple box shows the response of PST and derivative operator for feature detection under high-light-level conditions. While PST enhances low contrast features under low-light-level as wells as under high-light-levels (see green box) unlike derivative operator which identifies high contrast features (see blue box).

66

purple box in the Figure 3.11), PST outperforms when the contrast is low (shown in green box in the Figure 3.11). On the contrary, the conventional derivative based edge operator response is dominating only in high contrast regions (shown in purple box in the Figure 3.11).

(A) Original image   **(B) Feature detection using
conventional edge detectors**   **(C) Feature detection using PST**
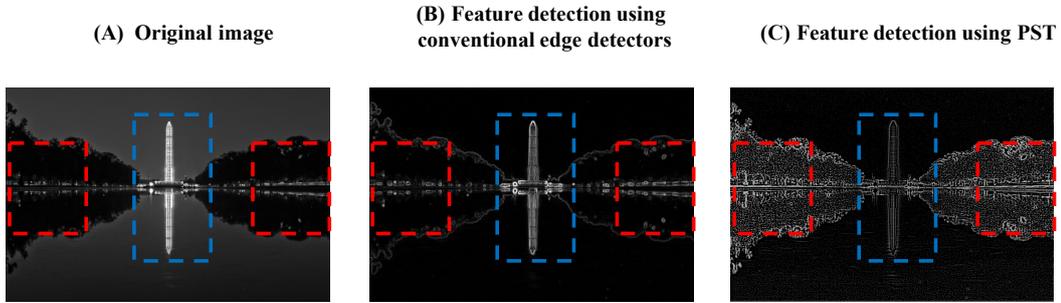
Figure 3.12: **Comparison of feature detection using the smooth derivative operator to the case of feature detection using the PST**. The original image is shown in (A). The smooth derivative operator is unable to efficiently visualize the features in the low contrast areas of the image (see red boxes in (B)) whereas PST detects these low contrast features (see red boxes in (C)).

We consider another 14-bit HDR image to show feature enhancement in low light level regions using PST. The image has sharp features in the form of edges of leaves and branches of trees in the low light level regions, see red dashed box in Figure 3.12(A). Results of feature detection using the smooth derivative operator and the PST operator are shown in Figure 3.12(B) and Figure 3.12(C), respectively. The warp and strength parameters used for the PST operator are 22.4 and 10.5, respectively. The conventional smooth derivative operator detects features only in high contrast regions, as shown in blue dashed box. The derivative operator fails to identify edges corresponding to low contrast areas of the image, see red dashed box in Figure 3.12(B). PST due to its inherent equalization mechanism detects these features as shown in red dashed box in Figure 3.12(C). It can be observed that PST does not have a strong edge response corresponding to high

contrast regimes such as the edges of Washington Monument in the image. We will show later that this issue can be resolved by combining the edges from both the methods.
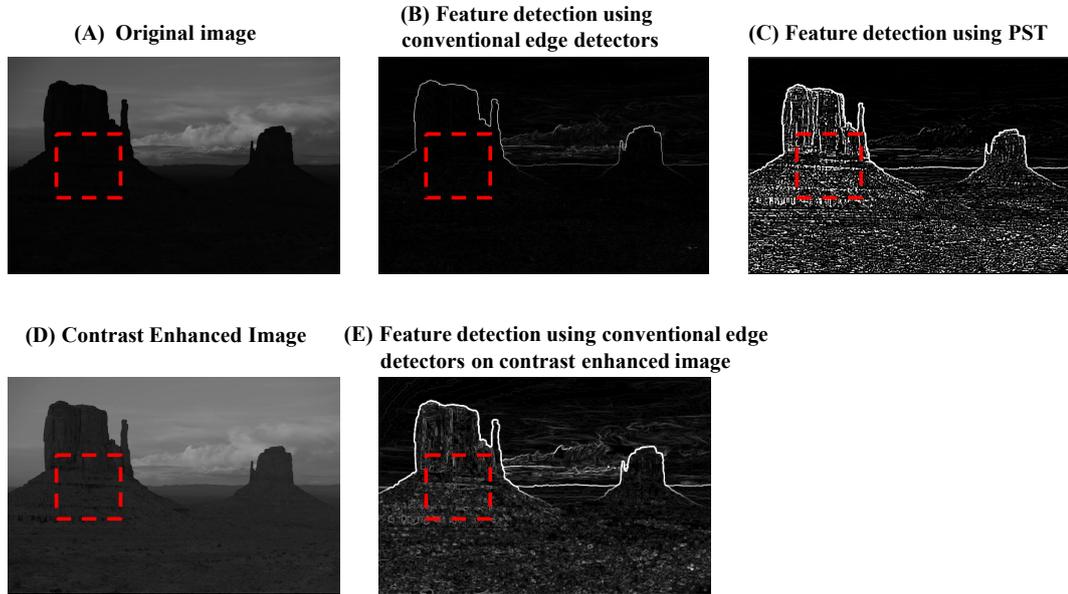
Figure 3.13: **Comparison of feature detection using the smooth derivative operator on a contrast enhanced image to the case of feature detection using the PST**. The original image is shown in (A). The smooth derivative operator detects features in the under-exposed regions of the image only after enhancing the contrast of the image. The edges detected by the conventional derivative operator in the under-exposed regions of the contrast enhanced image are consistent with the ones detected by the PST (see red boxes in (C) and (E)).

In order to evaluate the performance of PST compared to the previous methods, we show an example of using PST for feature enhancement in a 14-bit HDR image and then compare the output with the conventional techniques for feature enhancement in low contrast regions. The image of the rock has sharp surface variations in the low light level regions, as shown in the red dashed box. The smooth derivative operator fails to detect these surface features in low contrast regions, see Figure 3.13(B). We apply the standard intensity histogram equaliza-

tion technique [11] to enhance the contrast of the under-exposed regions of this HDR image, shown in Figure 3.13(D). The derivative operator now detects features in these low contrast regions, see Figure 3.13(E) which are visually consistent with the features detected by PST, see Figure 3.13(C). The warp and strength parameters used for the PST operator are 12.4 and 0.48, respectively.
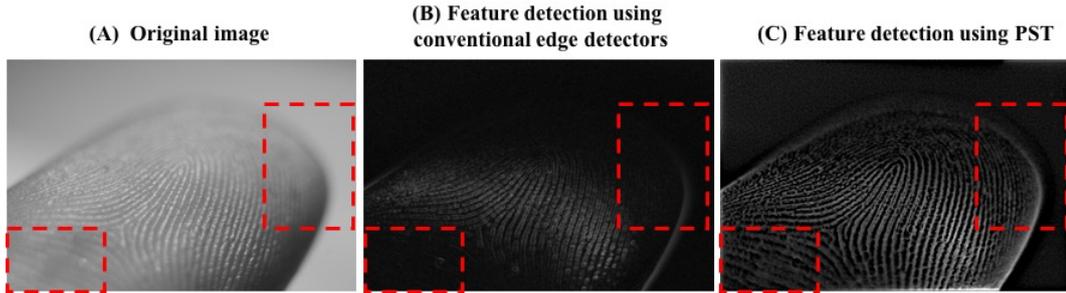
Figure 3.14: **Comparison of feature detection using the smooth derivative operator to the case of feature detection using the PST operator**. The original image is shown in (A). The smooth derivative operator is unable to detect features in the defocused areas of the image (as shown in red boxes in (B)). However, PST captures the fingerprint details in these blurred areas due to its unique re-configurable property (see (C)).

We consider another example of a defocused image, shown in Figure 3.14(A), to show that PST outperforms conventional edge detectors for feature enhancement in visually impaired images. By comparing the edge map from smooth derivative function to that from the PST operator, shown in Figure 3.14 (B) and Figure 3.14 (C), respectively, it is evident that the PST operator detects edges of finger prints in the blurred (low contrast) regions marked by red boxes. This opens up the possibility of using PST for fingerprint analysis in forensic science. The warp and strength parameters used here for the PST operator are 1000.2 and 50.4, respectively.
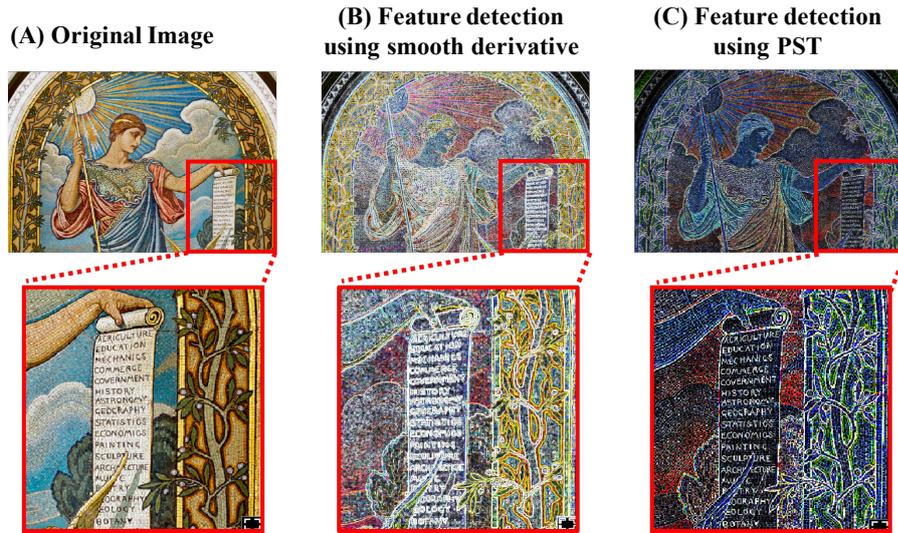
Figure 3.15: **Comparison of feature detection using conventional derivative based edge operator to the case of feature detection using Phase Stretch Transform (PST).** Original image is shown in (A). Results of feature detection using conventional detectors and PST operator are shown in (B) and (C), respectively. Enlarged view of the scroll in the painting, shown in the red boxes, establishes the superiority of PST to trace the edges of alphabets in the scroll.

Finally, Figure 3.15 shows another example application of PST for Optical Character Recognition (OCR). As shown in Figure 3.15, in this painting of "Minerva of Peace" there are optical characters in the scroll that need to be recognized (see red solid box in Figure 3.15 A). Results of feature detection using conventional derivative based edge operator and PST operator are shown in Figure 3.15B and 15C respectively. Clearly, conventional operator fails to efficiently visualize the sharp features of the alphabets in the scroll. However, PST traces the edges of alphabets efficiently and thus, provide more information on the contrast changes in dark areas due to its natural equalization mechanism, see Figure 3.15C. Conventional derivative based edge operator was implemented from find edge function in ImageJ software. The warp, W, and strength, S, parameters used for the PST operator are 13 and 0.4, respectively.

70

To further enhance the dynamic range of operation for feature detection, we introduced a hybrid system that combines the edge responses from the PST and the conventional derivative operator, enabling feature detection in low as well as high contrast regions [22,25]. As shown in Figure 3.16, the edge map of the hybrid system has edges in the high contrast regions such as the hand in the image and also in the low contrast regions such as the pattern on the curtain. It can be clearly seen that the feature detection of the hybrid system surpasses the capabilities of feature detection from the derivative and the PST operator.
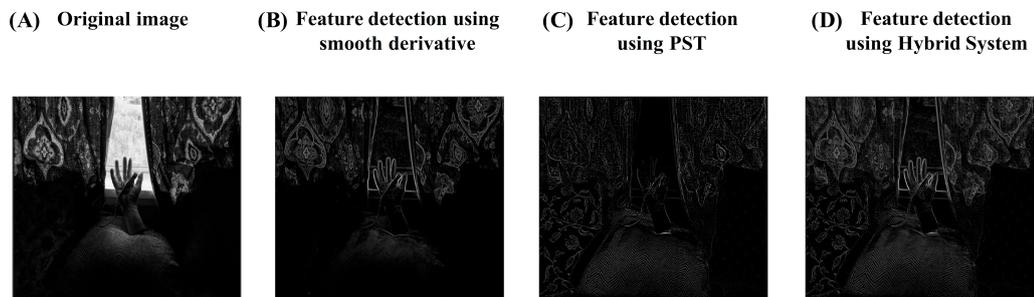
Figure 3.16: **Hybrid system that combines the edge maps from the smooth derivative operator and the PST**. The original image is shown in (A). Results of feature detection using the smooth derivative operator and the PST are shown in (B) and (C), respectively. The output of the hybrid system is shown in (D). Note that in (D), the strength of the detected features in both the high light level and low light level regions is same. The hybrid system selects the detected features in the darker regions using the PST and in the brighter regions using the smooth derivative operator and thereby, provides a wide dynamic range of operation.

3.7 Optical Computing of Mathematical Derivatives using Dispersion and Coherent Detection

Early attempts in the field of optical computing were futile because they aimed to construct an all-optical equivalent of the digital computer [80]. However, photonics does have potential for performing certain computing functions in the analog domain [80]. These analog photonic hardware accelerators precede an optical-to-electrical conversion and alleviate bandwidth, dynamic range and power consumption bottlenecks on the subsequent electronics [73]. Far from a general purpose computer, photonic hardware accelerators are custom designed to perform specific signal transformations in real time and in analog domain. The computational operations that have already been realized by these analog optical computing primitives include integration [81], data compression [36], time–bandwidth engineering [82], logarithm [83] and optical dynamic range compression [[84]. Optical implementation of temporal differentiation has been demonstrated in two ways: field differentiators realized using a micro-ring resonator [85] and intensity differentiators implemented using phase modulation [86].
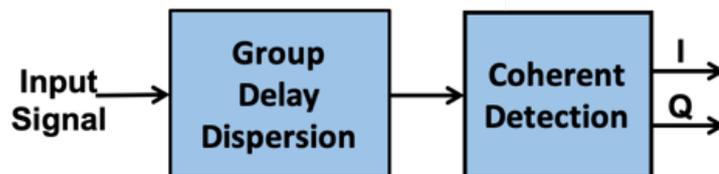
Figure 3.17: **Implementation of Proposed Differentiator**. The input signal passes through a dispersive medium that is designed to have specific dispersive property (refractive index profile). The complex output is detected using coherent detection method. The real part of the output is same as the input signal while the small imaginary component is the derivative of the input signal.

Inspired by the physics of Photonic Time Stretch technology, and its spin-off in the form of the Phase Stretch Transform, we proposed a novel computational approach for calculating mathematical derivatives as described in [87]. The proposed all-optical differentiator consists of group delay dispersion primitives and coherent detection as shown in Figure 3.17 .

The reconfigurability to compute different orders of differentiation is realized simply by controlling the applied group delay dispersion spectrum. The simulation results presented in Figure 3.18 demonstrate that the computation of any-order of differentiation of the input signal is possible. As shown in Figure 3.18, the imaginary component of the output gives the m-th order of differentiation while the real component of the output recovers the input signal. The proposed approach has promising applications in signal classification and is also easy to implement enabling realization of a real-time, low power and broad band optical differentiation functions.
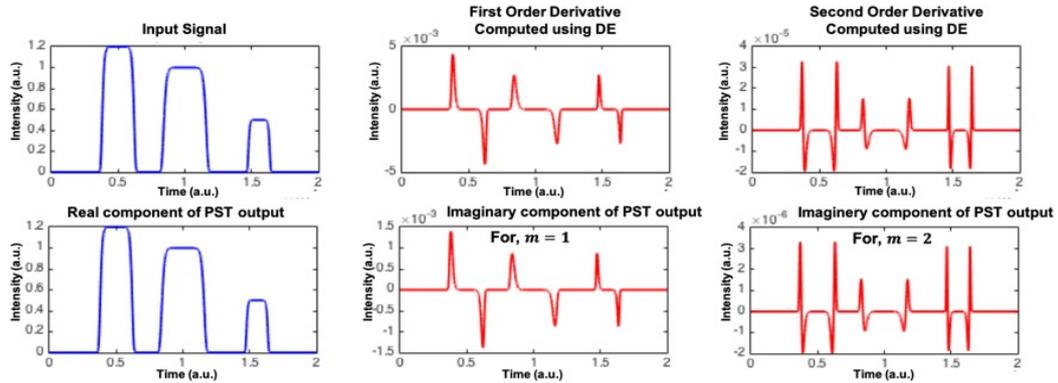
Figure 3.18: **Simulation results for computing mathematical derivatives using dispersion**. Different orders of differentiation on a calibrated signal using group delay dispersion. The input signal is recovered using the real component of the output while the imaginary component of the output shows the computed derivative corresponding to different order of differentiation based on the applied group delay dispersion. DE = Differential Equation, PST=Phase Stretch Transform.

3.8   Conclusions

In this chapter, we presented that the physics of light propagation in a dispersive or a diffractive media has natural properties that can be exploited for various applications. Photonic time stretch technology utilizes dispersion to slow down an analog signal in time. Phase Stretch Transform employs dispersion to extract features from the data. We also presented how this method is useful for edge detection in visually impaired images and showed the mathematical transform of this method inspired by the physics of photonic time stretch. We showed via analytical derivations as well as numerical simulations that this physics-inspired transform has an intrinsic equalization property. This inherent ability of PST significantly improves feature detection in visually impaired images and thereby, results in a high dynamic range of operation for feature extraction.

The phase kernel of PST can be tuned to compute different orders of mathematical derivative via group delay dispersion operations. This inbuilt reconfigurability of our transform can be used to generate a hyper-dimensional feature set consisting of different orders of derivative of input for signal classification. Furthermore, our results show a method for the computation of mathematical derivatives via group delay dispersion operations This novel approach for computing mathematical derivatives is based on group delay dispersion plus/and coherent detection. The reconfigurability to compute different orders of differentiation is realized simply by controlling the applied group delay dispersion. The simulation results presented demonstrate that the computation of any-order of differentiation of the input signal is possible. The proposed computational approach is also easy to implement enabling realization of a real-time, low power and broad band optical differentiator.

CHAPTER 4

Phase-stretch Adaptive Gradient-field Extractor

Emulated by an algorithm, certain physical phenomena have useful properties for image transformation. For example, image denoising can be achieved by propagating the image through the heat diffusion equation. Different stages of the temporal evolution represent a multi-scale embedding of the image. Stimulated by the photonic time stretch, a real-time data acquisition technology, the Phase Stretch Transform (PST) emulates 2D propagation through a medium with group velocity dispersion, followed by coherent (phase) detection. The algorithm performs exceptionally well as an edge and texture extractor, particularly in visually impaired images. Here, we introduce a decomposition method that draws inspiration from the birefringent diffractive propagation. This decomposition method, which we term as **P**hase-stretch **A**daptive **G**radient-field **E**xtractor (**PAGE**) embeds the original image into a set of feature maps that selects semantic information at different scale, orientation, and spatial frequency. We demonstrate applications of this algorithm in edge detection and extraction of semantic information from medical images, electron microscopy images of semiconductor circuits, optical characters and finger print images.

Towards the end, we also present a novel generalized approach for computing gradient fields using the physical phenomenon of diffraction followed by coherent (phase and amplitude) detection. The proposed diffraction differentiator can be reconfigured to compute different orders of differentiation by controlling the value of diffractive phase. The analysis of this physical operation is extended to digi-

tal domain which leads to definition of a new generalized feature decomposition method using diffraction known as **D**iffractive **G**radient **T**ransform ($\mathbb{DGT}$). This transform embeds an input signal into a set of **D**iffractive **G**radient **F**ields (g) which can act as feature maps for further analysis.

4.1 Introduction

Physical phenomena described by partial differential equations (PDE) have inspired a new field in computational imaging and computer vision [88]. Such physics-inspired algorithms based on PDEs have been successful for image smoothening and restoration. Image restoration can be viewed as obtaining the solution to evolution equations by minimizing an energy function. The most popular PDE technique for image smoothening treats the original image as the initial state of a diffusion process and extracts filtered versions from its evolution at different times. This embeds the original image into a family of simpler images at a hierarchical scale. Such a scale-space representation is useful for extracting semantically important information [89].

Physics based algorithms not only outperform their conventional counterparts, but also have enabled new applications. Usage of these algorithms range from feature detection in digital images [90–92], to 3D modelling of objects from 2D images [93, 94], to optical character recognition [95] as well as for restoring audio quality [96] .

Phase Stretch Transform (PST) is a physics inspired algorithm that emulates 2D propagation through a medium with group velocity dispersion, followed by coherent (phase) detection [27, 75]. The algorithm performs exceptionally well as edge and texture extractor, in particular in visually impaired images [25]. This transform has an inherent equalization ability that supports wide dynamic range of operation for feature detection [22, 25, 29]. It also exhibits superior properties over

conventional derivative operators, particularly in terms of feature enhancement in noisy low contrast images. These properties have been exploited to develop image processing tools for clinical needs such as a decision support system for radiologists to diagnose pneumothorax [23,24], for resolution enhancement in brain MRI images [97], single molecule imaging [26], and image segmentation [98].

PST emulates the physics of photonic time stretch [77], a real time measurement technology that has enabled observation as well as detection of ultrafast, non-repetitive events like optical rogue waves [3], optical fiber soliton explosions [5] and birth of mode locking in laser [7]. Further, by combining photonic time stretch technology with machine learning algorithms, a world record accuracy has been achieved for classification of cancer cells in blood stream [8,9].

The photonic time stretch employs group-velocity dispersion (GVD) in an optical fiber to slow down an analog signal in time by propagating a modulated optical pulse through the time stretch system which is governed by the following equation:

$$E_o\left(z,t\right) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \tilde{E}_i\left(0,\omega\right) \cdot \left[e^{\frac{-j\beta_2 z\omega^2}{2}}\right] \cdot e^{j\omega t} d\omega \qquad (4.1)$$

where, β_2=GVD parameter, z is propagation distance, $E_o\left(z,t\right)$ is the reshaped output pulse at distance z and time t. The response of dispersive element in the time-stretch system can be approximated a phase propagator $\tilde{K}\left[\omega\right] = e^{\frac{-j\beta_2 z\omega^2}{2}}$ which leads to the definition of PST for a discrete 2D signal as following:

$$\mathbb{PST}\left\{E_i\left[x,y\right]\right\} \triangleq \measuredangle\left\{IFFT^2\left\{FFT^2\left\{E_i\left[x,y\right]\right\} \cdot \tilde{K}\left[u,v\right]\right\}\right\} \qquad (4.2)$$

In the above equations, $E_i\left[x,y\right]$ is the input image, FFT^2 is 2D Fast Fourier Transform, $IFFT^2$ is 2D Inverse Fast Fourier Transform, x and y are the spatial variables and, u and v are spatial frequency variables. The function $\tilde{K}\left[u,v\right]$ is called the warped phase kernel implemented in frequency domain for image processing. PST utilizes the GVD dispersion to convert a real image to a complex quantity such that the spatial phase after the $IFFT^2$ operation is a function of

frequency. Upon thresholding, the high frequency edges survive. The phase kernel for the PST is designed by converting the 2D cartesian frequencies u and v to polar coordinates which results in a symmetric cartesian phase kernel. However, as digital images are fundamentally 2 dimensional, there is an inherent loss of information in the features detected by PST. This motivates us to develop a more comprehensive approach that captures angular as well as spatial frequency information in a semantic fashion.

In this chapter, we discuss Phase-stretch Adaptive Gradient-field Extractor (PAGE), a new physics inspired feature engineering algorithm that computes a feature set comprising of edges at different spatial frequencies, at different orientations, and at different scales and was originally introduced in [99]. These filters emulate the physics of birefringent (orientation-dependent) diffractive propagation through a physical medium with a specific diffractive property. In such a medium, the dielectric constant of the medium and hence, it's refractive index is a function of spatial frequency and the polarization in the transverse plane. To understand this analogy, we consider an optical pulse with two linearly orthogonal polarization's, $\tilde{E}_x$ and $\tilde{E}_y$, propagating through a dispersive diffractive medium such that:

$$\tilde{E}_i\left(z,t\right) = \tilde{E}_x + \tilde{E}_y \tag{4.3}$$

As the propagation constant $\beta = \frac{n.2\pi}{\lambda}$ is a function of refractive index (spatially varying), the two orthogonal polarizations $\tilde{E}_x$ and $\tilde{E}_y$ will have different propagation constants and hence, a phase difference at the output given by the following equation:

$$\triangle\phi = \phi_x - \phi_y = \triangle\beta \cdot l = \frac{\omega_m}{c}\mid n_x - n_y\mid \cdot L \tag{4.4}$$

By controlling the value of n_x and n_y, as well the dependence of refractive index on frequency $n_x(\omega)$ and $n_y(\omega)$, we are able to detect a semantic hyper-dimensional feature set from a 2D image. We demonstrate with several visual examples in the later part of this chapter that the above filter banks can be applied for image

78

processing and computer vision applications such as for detection of fabrication artifacts in semiconductor chips, development of clinical decision support systems, recognition of optical characters or finger prints. In particular, we show that PAGE features outperform the conventional derivative operators as well as directional Gabor filter banks.

Further, we address the dual problem of spatial resolution and dynamic range limitations in an imaging system. In an ideal imaging system, the numerical aperture and the wavelength of an optical set up are the only factors that determine the spatial resolution offered by the modality. But under non-ideal conditions, the number of photons collected from a specimen control its dynamic range (the ratio between the largest and the smallest value of a variable quantity) which in turn also limits the spatial resolution. This leads to the fundamental dual-problem of spatial resolution and dynamic range limitations in an imaging modality [100].

Certain approaches to improve the resolution of the imaging system include use of wide-field fluorescence microscopy [101, 102] which offers better resolution than con-focal fluorescence microscopy [103], multiple fluorophores [104, 105]. Also, various image processing techniques such as multi-scale analysis using wavelets [106, 107] have been proposed for improving the resolution while retaining important visual information post the image acquisition. We show later in the Chapter that we are able to alleviate this dual-problem by incorporating, in our algorithm, a local adaptive contrast enhancement operator, also known as Tone Mapping Operator (TMO) which leads to excellent dynamic range. Other steps of the proposed decomposition method are discussed at length in the next section. The organization of the Chapter is as follows. In Section 4.2, we describe the details of the proposed decomposition method. Experimental results, a generalized diffractive gradient framework and conclusions are presented in Sections 4.3, 4.4, and 4.5, respectively.
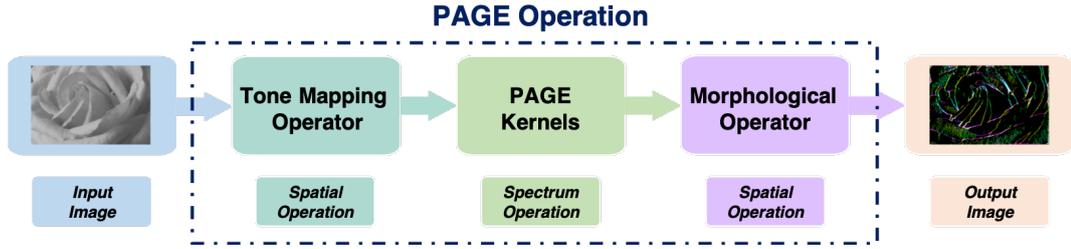
Figure 4.1: **Different steps of the Phase-stretch Gradient-field Extractor (PAGE) algorithm.** The pipeline starts with application of tone mapping in the spatial domain. This is followed by a smoothening and a spectral phase operation in the frequency domain. The spectral phase operation is the main component of the PAGE algorithm. The generated hyper-dimensional feature vector is thresholded and post-processed by morphological operations. PAGE embeds the original image into a set of feature maps that select semantic information at different scale, orientation, and spatial frequency.

4.2 Mathematical Framework

Different steps of our proposed decomposition method Phase-stretch Gradient-field Extractor (PAGE) for feature engineering are shown in Figure 4.1. The first step is to apply an adaptive tone mapping operator (TMO) to enhance the local contrast. Next, we reduce the noise by applying a smoothening kernel in frequency domain (this operation can also be done in spatial domain). We then apply a spectral phase kernel that emulates the birefringence and frequency channelized diffractive propagation. The final step of PAGE is to apply thresholding and morphological operations on the generated feature vectors in spatial domain to produce the final output. The PAGE output embeds the original image into a set of feature maps that select semantic information at different scale, orientation, and spatial frequency. We show in Figure 4.2 how PAGE embeds semantic information at different orientations for an X-ray image of a flower.

80
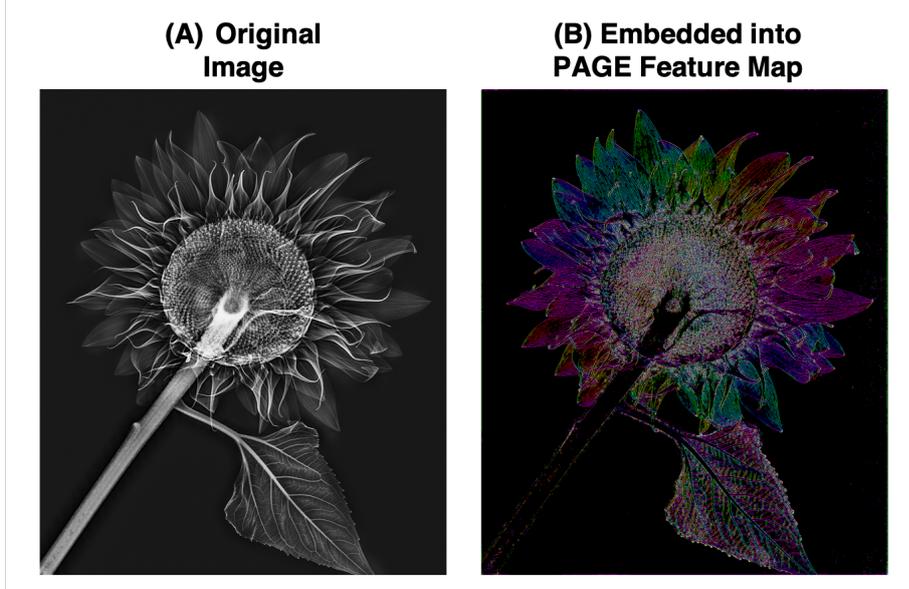
Figure 4.2: **The Phase-stretch Gradient-field Extractor (PAGE) feature map of an X-ray image.** The original image is shown on the left (A). PAGE embeds the original image into a feature map that selects semantic information at different orientations as shown in (B). The orientation of the edges is encoded into various color values here.

The sequence of steps of our physics-inspired feature extraction method, PAGE, can be represented by the following equations. We first define the birefringent stretch operator $\mathbb{S}\{\}$ as follows:

$$E_o\left[x,y\right] = \mathbb{S}\left\{E_i\left[x,y\right]\right\} = IFFT^2\left\{\widetilde{K}\left[u,v,\theta\right]\cdot\widetilde{L}\left[u,v\right]\cdot FFT^2\left\{TMO\left\{E_i\left[x,y\right]\right\}\right\}\right\}$$

$$(4.5)$$

where $E_o\left[x,y\right]$ is a complex quantity defined as,

$$E_o\left[x,y\right] = \left|E_o\left[x,y\right]\right|\ e^{j\theta\left[x,y\right]} \tag{4.6}$$

In the above equations, $E_i\left[x,y\right]$ is the input image, x and y are the spatial variables, FFT^2 is the two-dimensional Fast Fourier Transform, $IFFT^2$ is the two-dimensional Inverse Fast Fourier Transform, TMO is a spatially adaptive Tone

Mapping Operator and u and v are frequency variables. The function $\widetilde{K}[u, v, \theta]$ is called the PAGE kernel and the function $\widetilde{L}[u, v]$ is a smoothening kernel, both implemented in frequency domain. For all our simulations here, we consider $\widetilde{L}[u, v]$ to be low pass gaussian filter whose cut off frequency is determined by the sigma of the gaussian filter (σ_{LPF}).

The PAGE operator $\mathbb{P}\{\}$ can then be defined as the phase of the output of the stretch operation $\mathbb{S}\{\}$ applied on the input image $E_i[x, y]$:

$$\mathbb{P}\{E_i[x, y]\} = \measuredangle\Big\{\mathbb{S}\big\{E_i[x, y]\big\}\Big\} \tag{4.7}$$

where $\measuredangle\langle\cdot\rangle$ is the angle operator. In the next subsections, we discuss each of the above mentioned kernels in detail and demonstrate the operation of each step using simulation results.

4.2.1   Tone Mapping Operator (TMO)

A tone mapping operator (TMO) is applied to enhance the local contrast in the input image $E_i[x, y]$. This technique is a standard method in the field of image processing to solve the problem of limited contrast in an imaging system while still preserving important details and thereby, helps in improving the dynamic range of an imaging system via post processing. Figure 4.3 shows that by applying a tone mapping operator to the input image, an enhanced contrast can be achieved (see red boxes). While various TMO operators have been developed for adaptive contrast enhancement, here, we implement the TMO step by applying a Contrast Limited Adaptive Histogram Equalization (CLAHE) operator to the input image.

We operate on the input image using a TMO first, followed by smoothening operator (low pass filter) and not vice versa. The reason to follow this sequence of operation is as follows. Noise present in an image is mostly represented by the high frequency components in the spectrum. These high frequency components can be present at both low-light-level or at high-light-level in the spatial domain.

Because of the use of a tone mapping operator, the low-light-level features get over emphasized [108, 109]. This also leads to amplification of the image noise particularly in low-light scenarios. By applying a smoothening filter after the TMO operation, we aim to remove these noise artifacts from the contrast enhancement step. Alternatively, where any noise is left after the application of a smoothening kernel on the input image, it could be amplified by the TMO operation in the next step. Therefore, one may need to alternate between the smoothening step and TMO before obtaining a final satisfactory result [110].

(A) Original Image   **(B) Tone Mapping Operator**
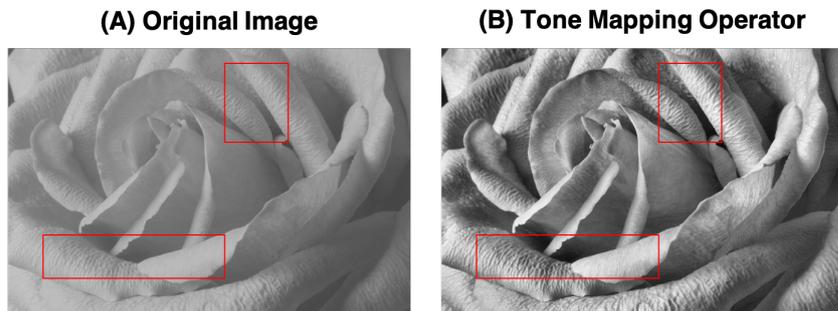
Figure 4.3: **Effect of Tone Mapping Operation (TMO)**: Tone Mapping Operation is implemented using Contrast Limited Adaptive Histogram Equalization (CLAHE) operator on the input image, that produces output image with better perceptual quality.

4.2.2 Phase-stretch Adaptive Gradient-field Extractor (PAGE) Kernel

Phase-stretch Adaptive Gradient-field Extractor (PAGE) filter banks are defined by the PAGE kernel $\widetilde{K}[u, v, \theta]$ and are designed to compute semantic information from an image at different orientations and frequencies. The PAGE kernel $\widetilde{K}[u, v, \theta]$, consists of a phase filter which is a function of frequency variable u and v, and a steerable angle variable θ which controls the directionality of the

response. We first define the translated frequency variable u' and v'

$$u' = u \cdot cos(\theta) + v \cdot sin(\theta) \tag{4.8}$$

$$v' = u \cdot sin\theta) + v \cdot cos(\theta) \tag{4.9}$$

such that the frequency vector rotates along the origin with θ

$$u' + jv' \Longleftarrow u + jv \tag{4.10}$$

We then define the PAGE kernel $\widetilde{K}[u, v, \theta]$ as a function of frequency variable u and v and steerable angle θ as follows:

$$\widetilde{K}[u, v, \theta] = \widetilde{K}[u', v'] = \exp\left\{j \cdot \phi_1(u') \cdot \phi_2(v')\right\} \tag{4.11}$$

where

$$\phi_1(u') = S_{u'} \cdot \frac{1}{\sigma_{u'}\sqrt{2\pi}} \cdot exp^{-(|u'|-\mu_{u'})^2/2\sigma_{u'}^2} \tag{4.12}$$

$$\phi_2(v') = S_{v'} \cdot \frac{1}{|v'|\sigma_{v'}\sqrt{2\pi}} \cdot exp^{-(ln(|v'|)-\mu_{v'})^2/2\sigma_{v'}^2} \tag{4.13}$$

There are two important things that should be noted here. First, we consider the modulus of our translated frequency variable u' and v' so that our kernel is symmetric for proper phase operation as discussed in [25]. Second, for all our simulation examples here, when we consider a bank of PAGE filters, we first normalize $\phi_1(u')$ and $\phi_2(v')$ in the range (0,1) for all values of θ and then, multiply the filter banks with $S_{u'}$ and $S_{v'}$ respectively, in order to make sure that the amplitude of each filter in the bank is same.

These filter banks can detect features at a particular frequency and/or in a particular direction. Therefore, by selecting a desired direction and/or frequency, a hyper-dimensional feature map can be constructed. We list all parameters in Table 4.1 that control different functionalities of our proposed decomposition method PAGE.

Notation	Variable
u and v	Spatial Frequency
θ	Steerable Angle
u' and v'	Translated Spatial Frequency
$\phi_1(\cdot)$	Normal Filter
$\phi_2(\cdot)$	Log Normal Filter
$S_{u'}$	Strength of ϕ_1 Filter
$S_{v'}$	Strength of ϕ_2 Filter
$\mu_{u'}$	Mean of normal distribution for ϕ_1 Filter
$\mu_{v'}$	Mean of log-normal distribution for ϕ_2 Filter
$\sigma_{u'}$	Sigma of normal distribution for ϕ_1 Filter
$\sigma_{v'}$	Sigma of log-normal distribution for ϕ_2 Filter
σ_{LPF}	Sigma of gaussian distribution for $\widetilde{L}[u,v]$ smoothening kernel
$Threshold(Min, Max)$	Bi-level feature thresholding for morphological operations

Table 4.1: **Different parameters of our physics-inspired feature decomposition method PAGE** The values of these parameters for Figure 2 simulation result are: $S_{u'} = 3.4$, , $S_{v'} = 1.2$, $\mu_{u'} = 0$, $\mu_{v'} = 0.4$, $\sigma_{u'} = 0.05$, $\sigma_{v'} = 0.7$, $\sigma_{LPF} = 0.1$ and $Threshold(Min, Max) = (-1, 0.0019)$. The number of filters considered for a 1 degree resolution is equals to 180.
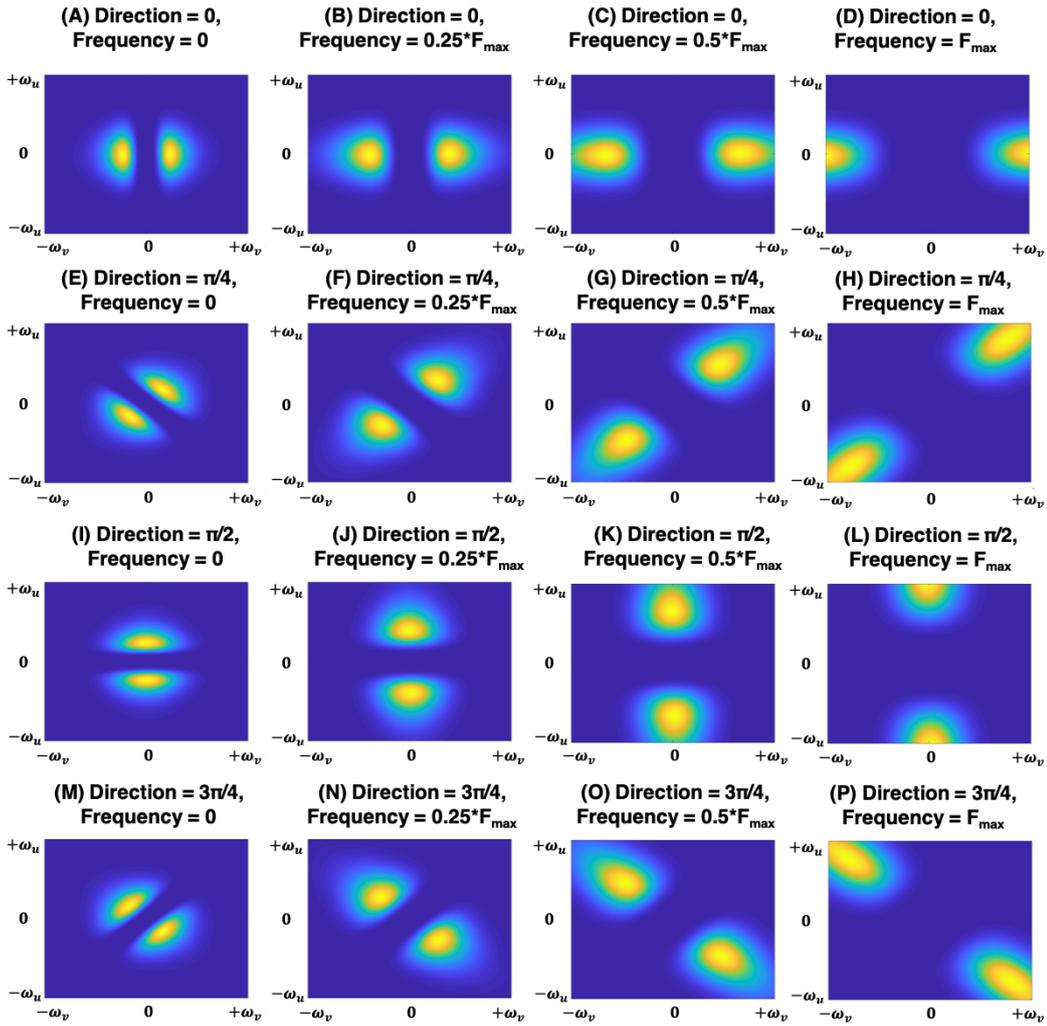
Figure 4.4: **Phase-stretch Gradient-field Extractor (PAGE) Filter Banks** (A)-(P) Phase filter banks as defined in Equations (8)-(13) for various frequencies and directions. The frequency variables u and v are normalized from $-\omega_u$ to $+\omega_u$ and $-\omega_v$ to $+\omega_v$, respectively. The center $\mu_{v'}$ of the phase kernel $S_{v'}$ is gradually increased for control over the frequency distribution. The values for steerable angle θ considered here are 0, $\pi/4$, $\pi/2$, $3\pi/4$.

Figures 4.4(A) to 4.4(P) show the generated phase profiles for $\phi_1(u') \cdot \phi_2(v')$ that select semantic information at different orientation and frequency as described in equations (10)-(13) using PAGE kernels. These phase kernels are

applied to the input image spectrum. Using the steerable angle, the directionality of edge response can be controlled in the output phase of the transformed image. The detected output response for each directional filter is thresholded using a bi-level method. This is done to preserve negative high amplitude values as well as positive high amplitude values.

4.2.2.1 Directionality

In order to detect features in a particular direction spread over the all the frequency components in the spectrum, we construct the PAGE filter banks by using equations (9)-(13) for $\widetilde{K}[u, v, \theta]$, $\phi_1(u')$ and $\phi_1(v')$ respectively. By controlling the value of sigma $\sigma_{u'}$ of normal distribution for $\phi_1(u')$ filter, we avoid any overlapping of directional filters as seen in Figure 4.5.
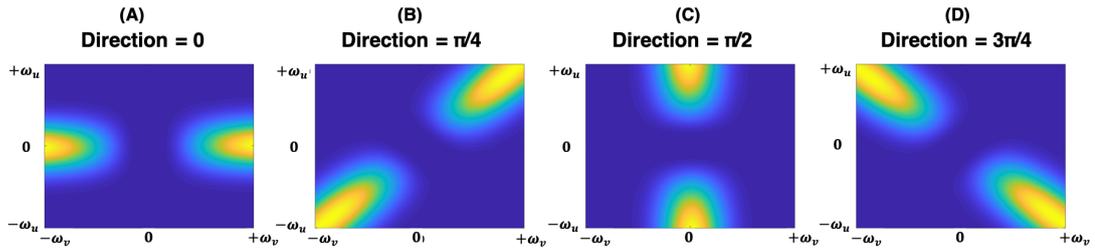
Figure 4.5: **Phase-stretch Gradient Field Extractor (PAGE) Directional Filter Banks** (A)-(D) The directional filter banks of PAGE computed using the definition in equations (9)-(13) for steerable angle $\theta = 0$, $\pi/4$, $\pi/2$ and $3\pi/4$ respectively. By monitoring the value of sigma $\sigma_{u'}$ of the normal filter $\phi_1(u')$, the angular spread of kernel $\widetilde{K}[u, v, \theta]$ can be controlled to avoid any overlapping of directional filters.
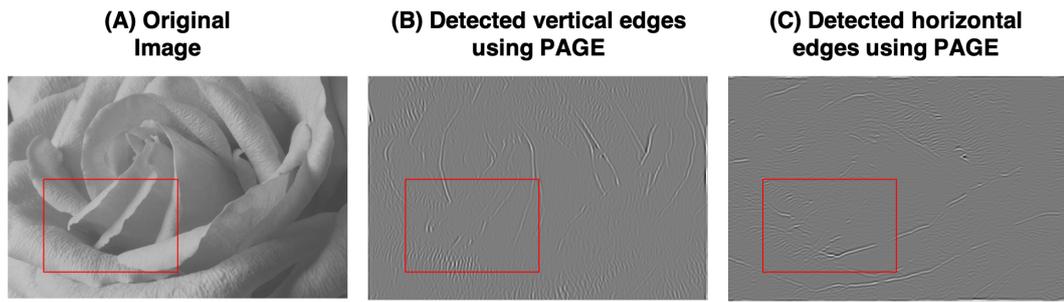
Figure 4.6: **Phase-stretch Gradient-field Extractor (PAGE) Directional Filter Banks Response** The original image is shown in (A). We design two directional PAGE filters here to detect vertical ($\theta = \pi/2$) and horizontal ($\theta = 0$) edges as shown in (B) and (C) respectively.
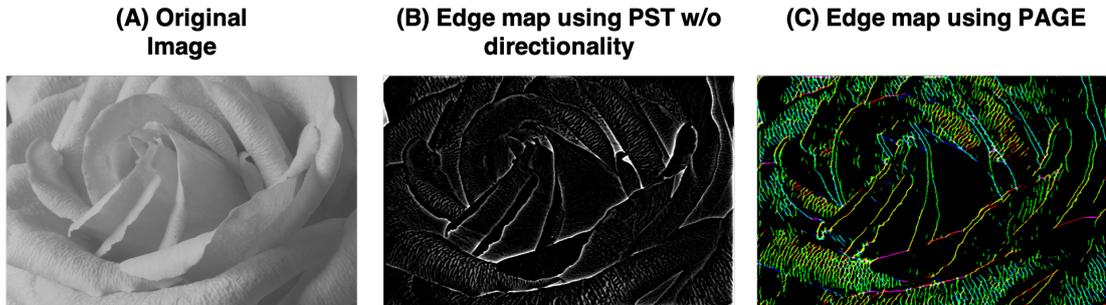
Figure 4.7: **Comparison of feature detection using Phase Stretch Transform (PST) and Phase-stretch Gradient-field Extractor (PAGE)** The original image is shown in (A). The output edge image obtained using PST without the support of directional response is shown in (B). The edge map obtained using PAGE filter banks that support edge detection at all frequencies is shown in (C). Different color values are used to show the orientation of the edges.

We first evaluate the performance of these kernel by qualitatively comparing the feature detection of PAGE with PST. The image under analysis is a gray-scale image of a rose. For a better visual understanding of our method, we first compute orthogonal directional responses as shown in Figure 4.6. We then show results of edge detection using PST and PAGE in Figure 4.7. The values for the parameters

strength $S_{u'} = 2.8$, $S_{v'} = 0.5$, $\mu_{u'} = 0$, $\mu_{v'} = 0.4$, $\sigma_{u'} = 0.05$, $\sigma_{v'} = 0.7$, $\sigma_{LPF} = 0.1$ and $Threshold(Min, Max) = (-1, 0.0019)$. The number of filters considered for a 1 degree resolution is equals to 180. Morphological operations used for the result shown in Figure 4.7(C) include edge thinning and isolated pixel removing for each directional response. As evident in Figure 4.7, edges are accurately extracted with our technique. Different colors in the computed edge response indicate the edge directionality.

(A) Original Image   **(B) Semantic information at low frequency using PAGE**   **(C) Semantic information at high frequency using PAGE**
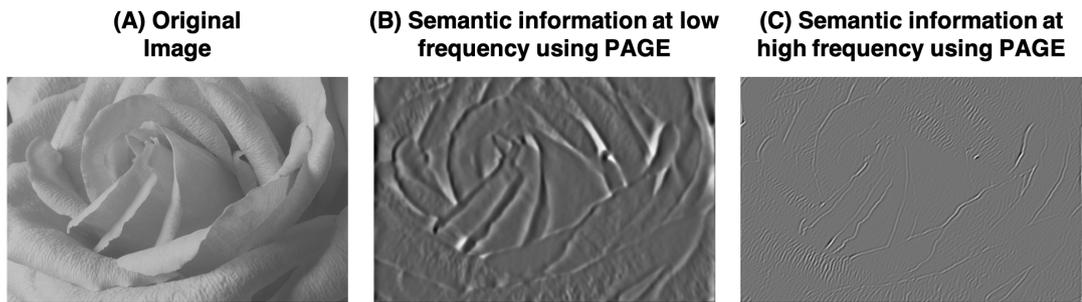
Figure 4.8: **Feature detection using Phase-stretch Gradient Field Extractor (PAGE) at low and high frequency:** Features detected at low frequency are much smoother whereas for high frequency, the features are sharper. This demonstrates the frequency selectivity for feature detection using PAGE.

4.2.2.2 Frequency Selectivity

The PAGE filter banks can also be designed to detect edges at a particular frequency by controlling the spread of log normal distribution. To demonstrate this functionality, we show the features detected at low and high frequency using the rose image as an example in the Figure 4.8. As seen in the figure, the features detected at low frequency are smoother and at high frequency are sharper.

4.3 Discussion

4.3.1 Comparison to Gabor Feature Extractors

We demonstrate the effectiveness of our decomposition method by comparing the directional edge response obtained by applying Gabor filter banks to an optical character image. We design 24 gabor directional filters and augment the response from each of the filters to generate the image in Figure 4.9(B). As seen in the Figure 4.9(C), with PAGE we have a better spatial localization of the edge response. By spatial localization, we mean that inherently PAGE has a sharper edge response, as seen in the figure. This is because, unlike the Gabor filters whose bandwidth is determined by the sigma parameter of the filter, in PAGE, the bandwidth of the response is determined by the input image dimension. Therefore, there is better localization of edge with PAGE. The parameters values are strength $S_{u'} = 2.8$, $S_{v'} = 0.5$, $\mu_{u'} = 0$, $\mu_{v'} = 0.4$, $\sigma_{u'} = 0.05$, $\sigma_{v'} = 0.7$, $\sigma_{LPF} = 0.1$ and $Threshold(Min, Max) = (-1, 0.0019)$. The number of filters considered for a 1 degree resolution is equals to 180.

(A) Original Image   **(B) Embedded into Gabor Feature Map**   **(C) Embedded into PAGE Feature Map**
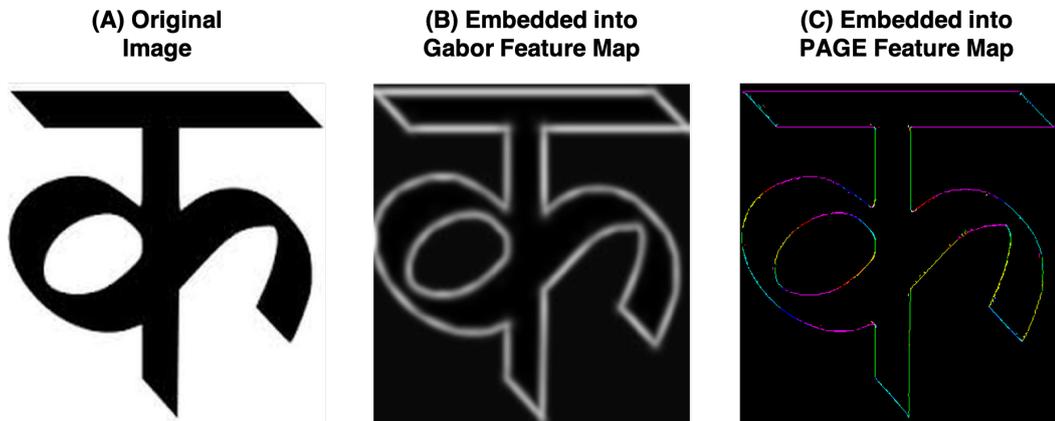
Figure 4.9: **Comparison to Gabor Feature Extractors:** Features detected using Gabor do not have inherent spatial feature localization. With PAGE, the features are more sharper as the bandwidth of the response is determined by the input image dimension.

4.3.2 Comparison to Derivative Feature Extractors

To demonstrate the superiority of our decomposition method, we compare the edge response obtained by applying derivative based operators to a test image shown in Figure 4.10 (A). The response to a derivative based operator is computed by using the edge function of Matlab software (canny) and is shown in Figure 4.10 (B). As seen in Figure 4.10 (C), PAGE outperforms derivative based operators by producing the orientation information and low contrast details. The parameters values are strength $S_{u'} = 2.7$, , $S_{v'} = 0.5$, $\mu_{u'} = 0$, $\mu_{v'} = 0.4$, $\sigma_{u'} = 0.05$, $\sigma_{v'} = 0.7$, $\sigma_{LPF} = 0.1$ and $Threshold(Min, Max) = (-1, 0.0019)$. The number of filters considered for a 1 degree resolution is equals to 180.

(A) Original Image   **(B) Embedded into Canny Edge Map**   **(C) Embedded into PAGE Feature Map**
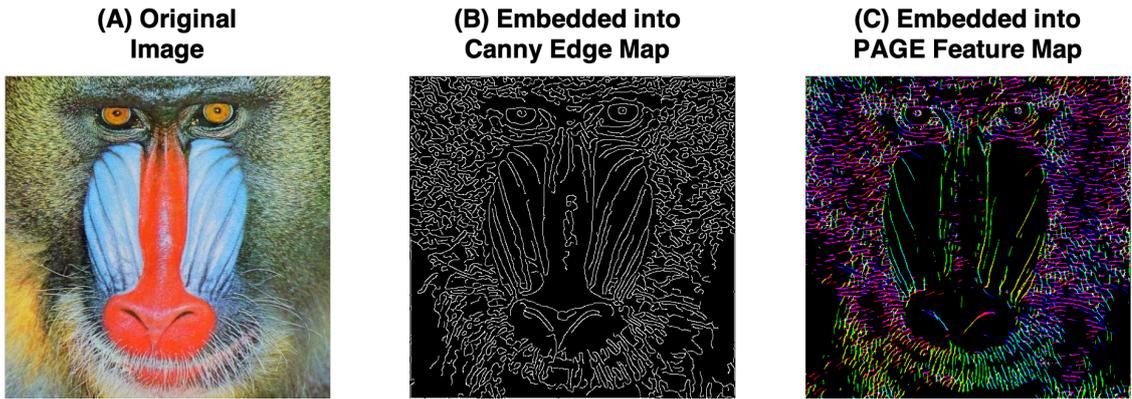
Figure 4.10: **Comparison to Derivative Feature Extractors:** Features detected with derivative based edge operators calculate the directionality based on the horizontal and vertical gradients and do not provide information about the spatial frequency of the edges. PAGE provides both the orientation as well as the spatial frequency selectivity in the output response.

4.3.3 Applications to Various Machine Vision Problems

We apply our decomposition method to different types of images to show that the directional edge response obtained by PAGE can be used for various machine vision applications. For example, in Figure 4.11 , we show application of PAGE to a Single Electron Microscope (SEM) image of an integrated circuit chip. As seen, the PAGE feature response is able to capture the edges corresponding to the chip layout (even the low contrast details). Based on the viewing angle (camera position), the layout edges should appropriately be rendered in the image as well as in the edge map. This can be used to identify any chip artifacts during the fabrication process. The parameters values for generating the feature map shown in Figure 4.11 are strength $S_{u'} = 3.1$, , $S_{v'} = 0.9$, $\mu_{u'} = 0$, $\mu_{v'} = 0.4$, $\sigma_{u'} = 0.05$, $\sigma_{v'} = 0.7$, $\sigma_{LPF} = 0.1$ and $Threshold(Min, Max) = (-1, 0.0042)$. The number of filters considered for a 1 degree resolution is equals to 180.
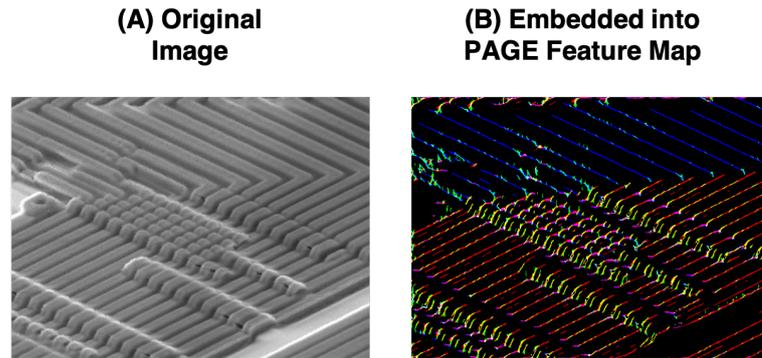
Figure 4.11: **Fabrication artifact detection using Phase-stretch Gradient-field Extractor (PAGE) on a Single Electron Microscope (SEM) image of integrated circuit chip.** The original image is shown in (A). The output edge image obtained using PAGE filter banks that support edge detection at all frequencies is shown in (B). Different color values are used to show the orientation of the edges that correspond to the chip layout and can be used to detect fabrication artifacts.

We also apply PAGE to detect directional edge response to an image of a finger print as shown in Figure 4.12. Not only does PAGE detects a directional edge response, but also has an inherent equalization property to detect low contrast edges. The parameters values are strength $S_{u'} = 1.5$, , $S_{v'} = 0.4$, $\mu_{u'} = 0$, $\mu_{v'} = 0.4$, $\sigma_{u'} = 0.05$, $\sigma_{v'} = 0.7$, $\sigma_{LPF} = 0.08$ and $Threshold(Min, Max) = (-1, 0.0019)$. The number of filters considered for a 1 degree resolution is equals to 180.

(A) Original Image **(B) Embedded into PAGE Feature Map**
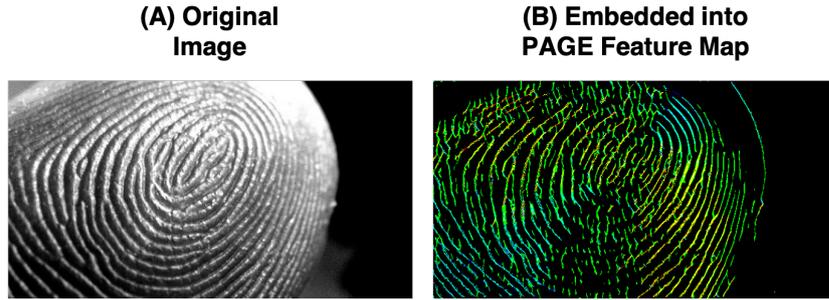
Figure 4.12: **Fingerprint feature map using Phase-stretch Gradient-field Extractor (PAGE).** The original image is shown in (A). The output edge image obtained using PAGE filter banks that support edge detection at all frequencies is shown in (B). As the edges of the fingerprint rotate, the response value changes (shown here with different color value).

Next, we show application of our decomposition method PAGE to extract edges of vessels from a retinal image in Figure 4.13. The distribution of vessels based on the orientation of the edges can be used as an important feature to detect any abnormalities in the eye structure. As seen, the PAGE feature response is able to capture both the low contrast details as well as information about the directionality of the vessel edges which is coded in form of the color value in RGB space. The parameters values are strength $S_{u'} = 2.2$, , $S_{v'} = 1.1$, $\mu_{u'} = 0$, $\mu_{v'} = 0.4$, $\sigma_{u'} = 0.05$, $\sigma_{v'} = 0.7$, $\sigma_{LPF} = 0.1$ and $Threshold(Min, Max) = (-1, 0.0019)$. The number of filters considered for a 1 degree resolution is equals to 180.
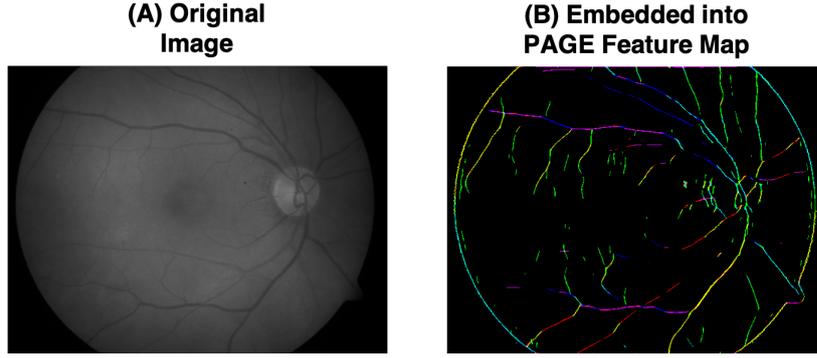
Figure 4.13: **Vessel detection using Phase-stretch Gradient-field Extractor (PAGE) on an image of a retina.** The original image is shown in (A). The output edge image obtained using PAGE filter banks that support edge detection at all frequencies is shown in (B). Different color values are used to show the orientation of the edges. The low contrast vessels are not only detected using PAGE but also information on how the direction of the blood flow changes across the eye based on the vessel distribution is extracted.

4.4 Diffractive Physics

In this section, we introduce a generalized method to compute gradient fields by employing the physical phenomenon of diffraction. To understand our proposed method of computing gradients using diffraction, we first study the physics behind the homogeneous paraxial diffraction here. For this, we start with the general solution to the homogeneous electromagnetic wave equation formed as a weighted superposition of all possible elementary plane wave solutions given by the following equation [111]:

$$E_o(x, y, z) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \widetilde{E}(k_x, k_y) \ e^{+jk_z z} \ e^{j(k_x x + k_y y)} \ dk_x dk_y \qquad (4.14)$$

The output $E_o(x, y, z)$ is a complex quantity defined as,

$$E_o(x, y, z) = |E_o(x, y, z)| \ e^{j\psi(x,y,z)} \qquad (4.15)$$

94

where $|E_o(x, y, z)|$ is the magnitude and $e^{j\psi(x,y,z)}$ is the phase of the diffracted output field as function of transverse spatial coordinates x and y and longitudinal spatial coordinate z after propagation along the $-z$ axis. $\widetilde{E}(k_x, k_y)$ is the spectrum of the incident field obtained by the Fourier transform of the input incident field $E(x, y)$ such that k_x and k_y are the transverse spatial frequency variables. Therefore, the operation of Inverse Fourier transform on the input incident field spectrum $\widetilde{E}(k_x, k_y)$ can be represented as:

$$E(x, y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \widetilde{E}(k_x, k_y)\ e^{j(k_x x + k_y y)}\ dk_x dk_y \qquad (4.16)$$

The diffraction phase kernel $H(k_z) = e^{jk_z z}$ is an isotropic phase kernel. For the case of paraxial plane waves (where the optic axis is assumed along z-axis), the propagation constant along the z axis, k_z, is equal to

$$k_z = \sqrt{k^2 - \left(k_x^2 + k_y^2\right)} \qquad (4.17)$$

where $k_x = 2\pi/\triangle x$ and $k_y = 2\pi/\triangle y$. According to the paraxial approximation, the spatial features $\triangle x$ and $\triangle y$ are large compared to the propagation wavelength $(\triangle x, \triangle y >> \lambda/n)$ and therefore, k_x and k_y are very small compared to k. We can then rewrite propagation constant equations as follows:

$$k_z = k - (k_x^2 + k_y^2)/(2 \cdot k) \qquad (4.18)$$

using the binomial approximation $(1 + d)^\gamma \approx 1 + \gamma d$, valid when $|d| < 1$ and $|\gamma d| \ll 1$. The diffractive phase imparted after propagating a distance z, in case of an isotropic medium can then be written as:

$$\phi(k_x, k_y) = k_z.z = \phi_o - \left\{\phi(k_x^2) + \phi(k_y^2)\right\} \qquad (4.19)$$

where the term $\phi_o = k.z$ is a constant scalar quantity. A constant term in the phase kernel implies a finite delay in transmission and hence, can be safely ignored for our discussion hereafter. The diffractive phase at position z, $\phi(k_x, k_y)$ then becomes a quadratic function of spatial frequency k_x and k_y given by the sum

of phase terms $\phi(k_x^2) = k_x^2/(2.k/z)$ and $\phi(k_y^2) = k_y^2/(2.k/z)$. For a homogenous isotropic medium, the spatial frequency k_x and k_y represents a sphere in the frequency space and therefore, a symmetric quadratic phase shift is imparted to the input field. We insert this expression of diffractive phase in Eq. 4.14:

$$E_o(x, y, z) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \widetilde{E}(k_x, k_y) \; e^{j\phi(k_x, k_y)} \; e^{j(k_x x + k_y y)} dk_x dk_y \qquad (4.20)$$

which can be also expanded as:

$$E_o(x, y, z) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \widetilde{E}(k_x, k_y) \left\{ \cos \left(\phi(k_x, k_y) \right) + \right.$$
$$\left. j \sin \left(\phi(k_x, k_y) \right) \right\} e^{j(k_x x + k_y y)} \; dk_x dk_y \quad (4.21)$$

By ignoring the constant phase term ϕ_o in the expression of the diffractive phase (such that $\phi(k_x, k_y) = -\phi(k_x^2) - \phi(k_y^2) = -(k_x^2 + k_y^2)/(2 \cdot k/z)$) and assuming that the value of this diffractive phase is very small compared to the amplitude of the input field (such that $\cos \theta = 1$ and $\sin \theta = \theta$) we obtain:

$$E_o(x, y, z) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \widetilde{E}(k_x, k_y) \left\{ 1 - j(k_x^2 + k_y^2)/(2 \cdot k/z) \right\} e^{j(k_x x + k_y y)} \; dk_x dk_y$$
$$(4.22)$$

The interesting aspect of our proposed method comes from the complex behavior of the output. The real part of our diffractive output is equal to the input signal while the imaginary component of the diffractive output gives us second order gradients along x and y direction from the differentiation property of Fourier transform:

$$E_o(x, y, z) = E(x, y) \;\; + \frac{j}{(2 \cdot k/z)} \left\{ \frac{d^2 E(x, y)}{dx^2} + \frac{d^2 E(x, y)}{dy^2} \right\} \qquad (4.23)$$

The final step of our proposed method **D**iffractive **G**radient **T**ransform ($\mathbb{DGT}$) is to compute the phase of the diffraction output using a coherent detector:

$$\mathbb{DGT} \left(E(x, y) \right) = \measuredangle \left\{ E_o(x, y, z) \right\} \qquad (4.24)$$

where $\measuredangle \langle \cdot \rangle$ is the phase operator.

As the applied phase value is small, we can approximate the coherent phase detection $\psi = tan^{-1}\left(imaginary/real\right) = imaginary/real$ resulting in:

$$\mathbb{DGT}\left(E(x,y)\right) = \frac{1}{(2 \cdot k/z)} \left\{\frac{d^2 E(x,y)}{dx^2} + \frac{d^2 E(x,y)}{dy^2}\right\} \Big/ E(x,y) \qquad (4.25)$$

The output of $\mathbb{DGT}$ consists of second-order directional derivative terms (one along x-axis and other along y-axis) such that the gain factor defined as $\alpha_x = \alpha_y = 1/(2 \cdot k/z)$ is same along both the directions x and y. There are two important observations to be noted down from the output of our proposed transform $\mathbb{DGT}$. First, the output consists of directional gradients along various dimensions of the input signal (here, x and y). Second, our proposed transform $\mathbb{DGT}$ has an inherent equalization property which comes from the nonlinear phase operation on the complex output as defined in Eq. 4.24. This powerful operation amplifies the gradient response at low input levels and thus, helps in capturing critical gradient features in low intensity regions. We will show applications of this equalization property to 1D as well as 2D signals later in the text.

4.4.1 Parameterization and Generalization of the diffraction theory

In order to understand the impact of various terms in the expression of diffraction phase as derived in Eq. 4.19, we consider various generalizations of this expression some of which are loosely applicable to physical systems but not always. However, these generalization are always true in case of the digital domain.

4.4.1.1 Generalization of symmetry and dimensionality

Let us first consider the case of D - dimensional space $(x_1, x_2, ...x_D)$ or $\mathbb{R}^D$ instead of 3D space (x, y, z) or $\mathbb{R}^3$. Let us also generalize the gain factor α of gradients along different directions. The expression for $\mathbb{DGT}$ can then be represented:

$$\mathbb{DGT}(E) = \sum_{d=1}^{(D)} \alpha_d . \frac{d^2 E}{dx_d^2} \bigg/ E \tag{4.26}$$

The above expression implies an asymmetric phase shift along different coordinate axis. For the case of anisotropic 3D space, the dependence of ϕ on k_x and k_y is governed by the equation of an ellipse as follows:

$$\phi(k_x, k_y) = \alpha_x . (k_y^2) + \alpha_y . (k_y^2) \tag{4.27}$$

where constants α_x, α_y control the ellipticity of the phase in the spatial frequency space formed by k_x and k_y such that $\alpha_x \neq \alpha_y$.

4.4.1.2 Generalization of the order

In the case of conventional homogenous diffraction, we have diffractive phase $\phi(k_x, k_y)$ as a quadratic function of spatial frequency variables k_x and k_y, such that $\phi(k_x, k_y) = (k_x^2 + k_y^2)/(2.k/z)$ and as such we are able to compute second order derivative of the input in the expression of the output of $\mathbb{DGT}$. Here, we generalize the expression of the diffractive phase $\phi(k)$ to a polynomial expression of the order M such that $\phi(k) = \alpha_1 . k^1 + \alpha_2 . k^2 + \alpha_3 . k^3 .... + \alpha_M . k^M$. The diffractive output of $\mathbb{DGT}$ will be a sum of different orders of derivatives given by:

$$\mathbb{DGT}(E) = \sum_{m=1}^{(M)} \sum_{d=1}^{(D)} \alpha_{d,m} . \frac{d^m E}{dx_d^m} \bigg/ E \tag{4.28}$$

The different powers of the spatial frequency k control the curvature of the phase $\phi(k)$ and thereby, contribute to different orders of derivative at the output. The terms $\alpha_{d,m}$ represents the gain factor for the mth order gradient along the d-direction.

4.4.1.3 Generalization of the frequency response

If we consider a physical system with a preferential resonating frequency, the imparted diffraction phase will have a narrow band response in this scenario. Thus,

ϕ on k_x and k_y is no longer a monotonic function but rather a non-monotonic function which is centered around the resonating frequency of the physical diffractive system. This scenario is used to design filters with channelized frequency response in digital domain as shown previously with PAGE.

4.4.1.4 Generalization of nonlinear equalization property

The natural phenomenons have unique properties which we aim to exploit for developing algorithms with unparelled performance. For instance, as shown previously our proposed transform $\mathbb{DGT}$ has an inherent equalization property which play a critical role in capturing gradient information at low intensity regions. Here, we generalize the equalization power of the input field E to an arbitrary order of equalization p as following:

$$\mathbb{DGT}(E) = \sum_{p=1}^{(P)} \sum_{m=1}^{(M)} \sum_{d=1}^{(D)} \alpha_{d,m,p} \cdot \frac{d^m E}{dx_d^m} \bigg/ E^p \qquad (4.29)$$

This is the most general expression of our transform $\mathbb{DGT}(E)$ which consists of sum of **D**iffractive **G**radient **F**ields defined as follows:

$$g_{d,m,p} = \alpha_{d,m,p} \cdot \frac{d^m E}{dx_d^m} \bigg/ E^p \qquad (4.30)$$

where $g_{d,m,p}$ is the mth order derivative of the input along the x_d direction such that it is equalized by the input to the power of p. When $p = 0$, the gradient fields $g_{d,m,0}$ are simply the different orders of derivatives along various directions. For $p > 0$, the derivatives are equalized by the input E^p and therefore, the response of our method is inherently nonlinear. The $\mathbb{DGT}$ response can then be represented in the form of a sum of gradient fields for the generalized case as:

$$\mathbb{DGT}(E) = \sum_{p=1}^{(P)} \sum_{m=1}^{(M)} \sum_{d=1}^{(D)} g_{d,m,p} \qquad (4.31)$$

99

Variable	Continuous	Discrete
Spatial variables	x_d	n_d
Spatial frequencies	k_d	u_d
Input Signal	$E(x_d)$	$E[n_d]$
Input Spectrum	$\widetilde{E}(k_d)$	$\widetilde{E}[u_d]$
Fourier and Inverse Fourier Transform	$\mathcal{F}$, $\mathcal{F}^{-1}$	$FFT, IFFT$

Table 4.2: **Discretization of the physical diffraction:** Discrete equivalents for continuous domain variables

4.4.2   Discretization of the physical diffraction theory

In this section, we discretize the phenomenon of physical diffraction by translating continuous domain variables and operations to their discrete counterparts as shown in Table 4.2. We can then write the discrete expression of **D**iffractive **G**radient **T**ransform ($\mathbb{DGT}$) as follows:

$$\mathbb{DGT}\left(E\right) = \sum_{p=1}^{(P)} \sum_{m=1}^{(M)} \sum_{d=1}^{(D)} \alpha_{d,m,p} . \frac{d^m E}{dn_d^m} \bigg/ E^p \qquad (4.32)$$

and **D**iffractive **G**radient **F**ields (g) as follows

$$g_{d,m,p} = \alpha_{d,m,p} . \frac{d^m E}{dn_d^m} \bigg/ E^p \qquad (4.33)$$
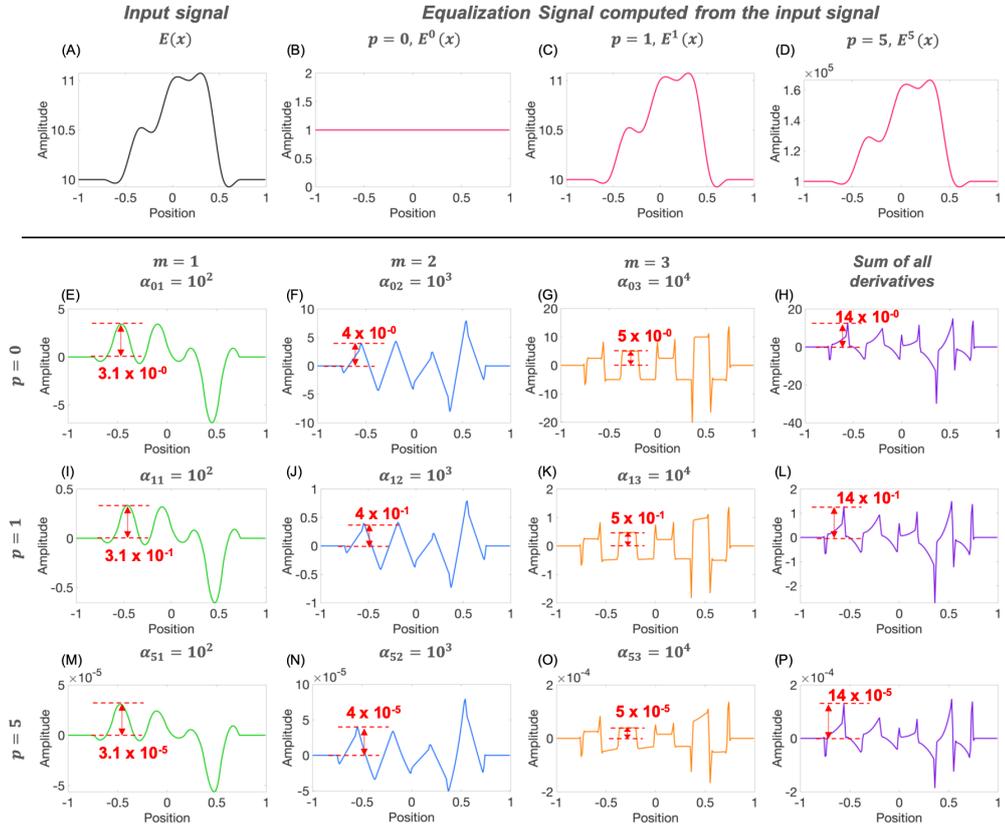
Figure 4.14: **Diffractive Gradient Transform ($\mathbb{DGT}$)**: The arbitrary 1D signal is shown in (A). As the power of equalization increases the amplitude of the gradient field becomes smaller for a higher brightness level (see annotations in Figure (E) vs (M)). The higher order of derivatives contribute to the fine details in the gradient field (see Figure (H), (L) and (P)).

101
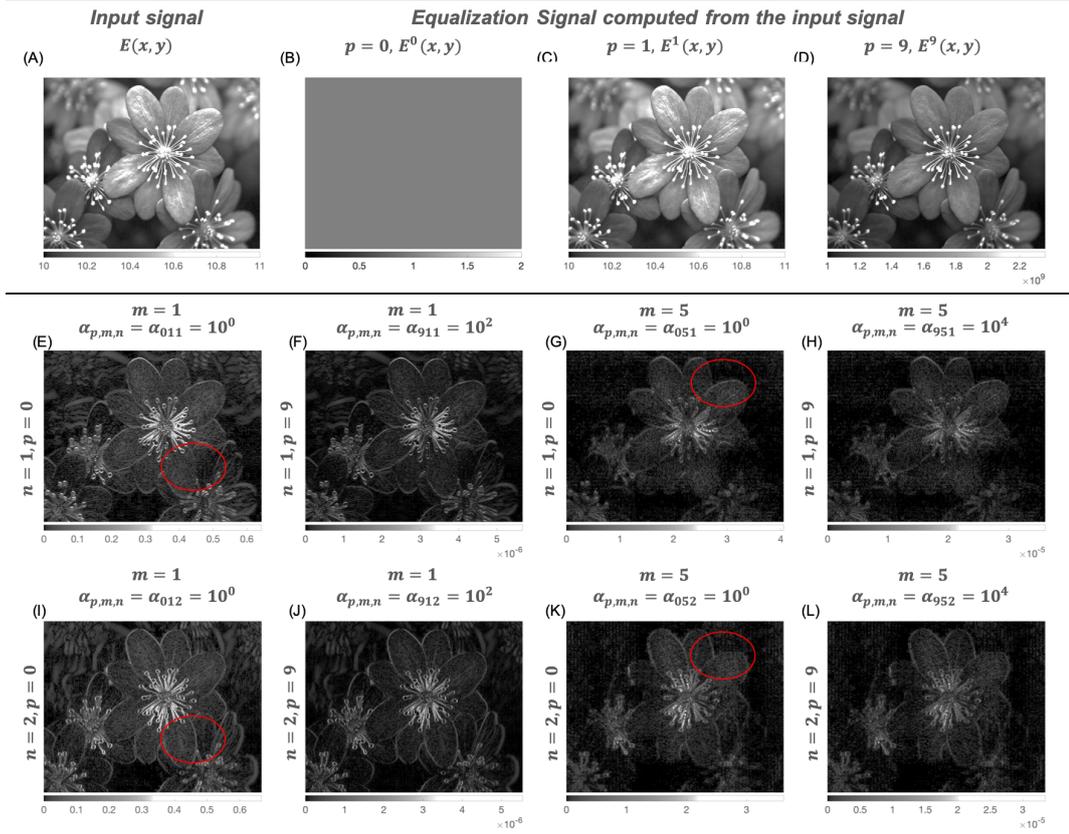
Figure 4.15: **Diffractive Gradient Transform on 2D digital images**: We compute gradient fields along different direction ($n1$ and $n2$) using the definition of $\mathbb{DGT}$. The input signal shows a strong gradient field in one direction and absent in other (see red circle in Figure (E) vs (I) or Figure (G) vs (K)). Also, the power of equalization play a critical role in controlling the amplitude of the gradient field (see for $m = 5$ vs $m = 1$).

To analyze performance of our transform $\mathbb{DGT}$, we consider first an arbitrary 1D signal (hence, D=1) as shown in Figure 4.14(A), such that we compute upto three orders of derivative ($m = 1, m = 2$ and $m = 3$) and show how a higher order of equalization plays critical role in controlling the amplitude of gradient fields especially at high intensity levels and produce a equalized derivative response (see Figure (E) vs (M). While the higher order of derivatives contribute to the fine details in the gradient field (see Figure (H), (L) and (P)). We show the directional

102

dependence of gradient fields by applying our transform $\mathbb{DGT}$ to a 2D digital image (see Figure 4.15). We also show here how the equalization controls the amplitude of gradient at low intensity regions and higher order derivatives contain finer details.

Next we consider the case for a time varying 2-D input and when $p = 1$ such that the $\mathbb{DGT}$ filters compute gradient fields at different orientations from the input image. Additionally, these filters are deigned to have channelized frequency response which controls the frequency distribution of the gradient fields. These computed DGF fields are combined together to produce the final feature image as shown in Figure 4.16. The final feature image contains the directional information of the gradient fields encoded in form of the color pixel values.
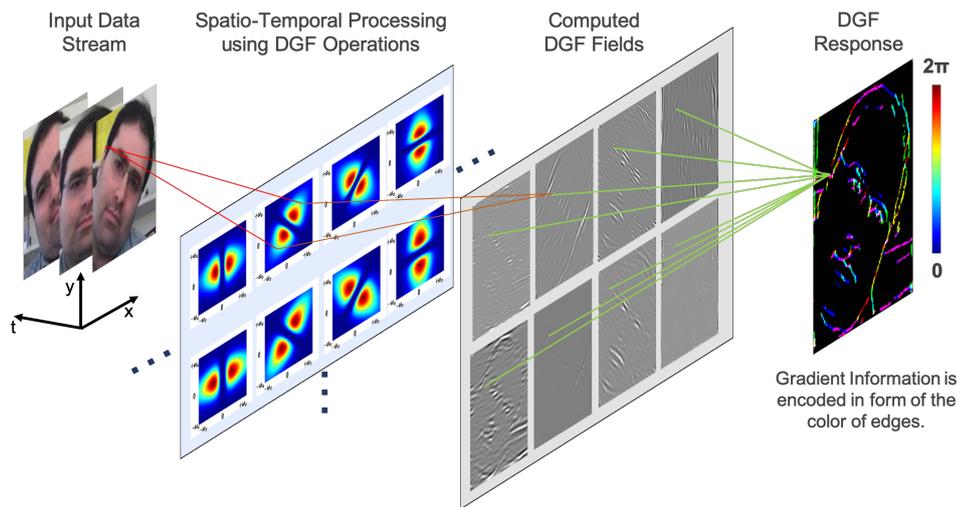
Figure 4.16:   **Application of Diffractive Gradient Transform to a time varying 2D input:** The spectral diffractive phase operation is the main component of the $\mathbb{DGT}$ algorithm. It embeds the original image into a set of gradient fields at different orientation, and spatial frequency. These fields are combined together to produce a feature image which contains the directional information of the features encoded in form of the color pixel values. This is also an example of channelized diffractive phase.

4.4.3 Relation to Birefringence

Birefringence is a phenomenon observed in optically anisotropic materials where the refractive index n is dependent on the polarization as well as the propagation direction of incident light (and therefore $n(d)$). For the case of propagation of a polarized input E such that $E = E_{ox} + E_{oy} = E_x(x,y) + E_y(x,y)$ through a birefringent material, the propagation phase $\phi(k_z, z)$ depends also on the state of polarization (SOP) but is not a function of spatial coordinates (as the medium is uniform). On the other hand, our transform $\mathbb{DGT}$ takes a scalar input (like vector potential) and mimics propagation through a birefringent material such that the optical properties of the medium (refractive index, n) are a function of spatial coordinates $n(x,y)$. The diffractive phase ϕ depends on the spatial coordinate in case of birefringence through a non-uniform medium when the input is scalar ($E(x,y)$) and therefore, there is only 1 SOP.

4.5 Conclusions

In this chapter, a presentation is made on a new feature engineering method that takes inspiration from the physical phenomenon of birefringence in an optical system. The introduced method called Phase-stretch Adaptive Gradient-field Extractor (PAGE) controls the diffractive properties of the simulated medium as a function of spatial location and channelized frequency. This method when applied to 2D digital images extracts semantic information from the input image at different orientation, scale and frequency and embeds this information into a hyper-dimensional feature map. The computed response is compared to other directional filters such as Gabor to demonstrate superior performance of PAGE. Applications of the algorithm for edge detection and extraction of semantic information from medical images, electron microscopy images of semiconductor circuits, optical character and finger print images is also shown.

4.6 Supplementary information

4.6.1 Stationary Phase Approximation

Stationary phase approximation allows us to create a mapping of the input signal between the time domain and spectral domain. We start with the general solution to the homogenous electromagnetic wave equation in physical rectangular coordinates (x, y, z):

$$E_o(x, y, z) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \widetilde{E}(k_x, k_y) \; e^{j(k_x x + k_y y)} \; e^{jzk_z} \; dk_x dk_y \qquad (4.34)$$

where $\widetilde{E}(k_x, k_y)$ is the Fourier transform of the input incident field $E(x, y)$, x and y are the transverse spatial coordinates, z is the longitudinal spatial coordinate and k_x and k_y are the transverse spatial frequency variables. Hence, $\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \widetilde{E}(k_x, k_y) \; e^{j(k_x x + k_y y)} \; dk_x dk_y$ represents the operation of inverse Fourier transform. The diffraction kernel along the z-axis is an isotropic phase kernel represented by the expression e^{jzk_z}. And, $E_o(x, y, z)$ is a complex quantity defined as,

$$E_o(x, y, z) = |E_o(x, y)| \; e^{j\psi(x,y)} \qquad (4.35)$$

where $|E_o(x, y, z)|$ is the magnitude and $e^{j\psi(x,y)}$ is the phase of the diffracted field as function of x and y at position z after propagation.

For the case of paraxial plane waves (where the optic axis is assumed along z-axis), the propagation constant along the z axis, k_z, is equal to

$$k_z = \sqrt{k^2 - \left(k_x^2 + k_y^2\right)} \qquad (4.36)$$

where $k_x = \frac{2\pi}{\triangle x}$ and $k_y = \frac{2\pi}{\triangle y}$. Next, using the paraxial approximation, we approximate that the spatial features $\triangle x$ and $\triangle y$ are large enough such that k_x and k_y are very small compared to k. By assuming $\triangle x, \triangle y << \frac{\lambda}{n}$, we can rewrite above equation as

$$k_z = k - \frac{\left(k_x^2 + k_y^2\right)}{2 \cdot k} \qquad (4.37)$$

using the binomial approximation $(1 + d)^\gamma \approx 1 + \gamma d$ which is valid when $|d| < 1$ and $|\gamma d| \ll 1$. We can then rewrite Eq. (1) as following:

$$E_o(x, y, z) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \widetilde{E}(k_x, k_y) \; e^{j(k_x x + k_y y)} \; e^{jz(k - \frac{k_x^2 + k_y^2}{2 \cdot k})} \; dk_x dk_y \qquad (4.38)$$

where the term e^{jzk} is a constant scalar quantity. A constant term in the phase kernel implies a finite delay in transmission. This term is independent of variable k_x and k_y and hence, can be taken out of the integration which results in:

$$E_o(x, y, z) = e^{jzk} \cdot \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \widetilde{E}(k_x, k_y) \; e^{-j\frac{z}{2k}(k_x^2 + k_y^2 - 2 \cdot \frac{xk}{z} \cdot k_x - 2 \cdot \frac{yk}{z} \cdot k_y)} \; dk_x dk_y \quad (4.39)$$

The terms $\frac{xk}{z}$ and $\frac{yk}{z}$ are independent of variables k_x and k_y and hence, can be used to complete the squares of k_x and k_y variable which results in the following equation:

$$E_o(x, y, z) = e^{j\left(zk + \frac{x^2 k}{2z} + \frac{y^2 k}{2z}\right)} \cdot \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \widetilde{E}(k_x, k_y) \; e^{-j\frac{z}{2k}\left((k_x - \frac{xk}{z})^2 + (k_y - \frac{yk}{z})^2\right)} \; dk_x dk_y$$
$$(4.40)$$

In the event the above integral has an asymptotic behavior which is completely defined by the behavior of the integrand at the critical points, we can use the stationary phase approximation which results in a mapping of spatial frequency to spatial coordinates given by:

$$k_x = \frac{xk}{z} \qquad (4.41)$$

and

$$k_y = \frac{yk}{z} \qquad (4.42)$$

4.6.2   Spatial Domain Operation for Diffraction

We show here our analysis of conventional diffraction theory in spatial domain eliminating the need of any spectrum operation. We start with the diffraction

106

equation derived in previous section with stationery phase approximation as:

$$E_o(x, y, z) = e^{j\left(zk + \frac{x^2 k}{2z} + \frac{y^2 k}{2z}\right)} \cdot \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \widetilde{E}(k_x, k_y) \; e^{-j\frac{z}{2k}\left((k_x - \frac{xk}{z})^2 + (k_y - \frac{yk}{z})^2\right)} \; dk_x dk_y$$

$$(4.43)$$

which can also be written in form of the diffractive phase kernel $H(k_x, k_y)$ as :

$$E_o(x, y, z) = e^{j\left(zk + \frac{x^2 k}{2z} + \frac{y^2 k}{2z}\right)} \cdot \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \widetilde{E}(k_x, k_y) \; H(k_x, k_y) \; dk_x dk_y \qquad (4.44)$$

where

$$H(k_x, k_y) = e^{-j\frac{z}{2k}\left((k_x - \frac{xk}{z})^2 + (k_y - \frac{yk}{z})^2\right)} \qquad (4.45)$$

Here for consistency with our analysis, we will be ignoring the constant phase term $e^{j\left(zk + \frac{x^2 k}{2z} + \frac{y^2 k}{2z}\right)}$ which resulted in a finite delay in transmission. Then, we can rewrite the diffraction equation as:

$$E_o(x, y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \widetilde{E}(k_x, k_y) \; H(k_x, k_y) \; dk_x dk_y \qquad (4.46)$$

Here, we have the phase kernel $H(k_x, k_y)$ as function of spatial frequencies k_x, k_y. Let us define a spatial phase kernel $K(x, y)$ such that

$$K(x, y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} H(k_x, k_y) \; dk_x dk_y \qquad (4.47)$$

By inserting the spatial phase kernel $K(x, y)$ and transforming the multiplication operation in spectrum domain to convolutional operation in spatial domain, we will have the diffraction equation as:

$$E_o(x, y) = E(x, y) * K(x, y) \qquad (4.48)$$

where $*$ represents the spatial convolutional operation between the input field $E(x, y)$ and the spatial domain phase kernel $K(x, y)$. Similarly, discrete equivalent of diffraction in spatial domain can be defined as:

$$E_o[n_1, n_2] = E[n_1, n_2] * K[n_1, n_2] \qquad (4.49)$$

where $E_o[n_1, n_2]$ is the diffractive output of the 2D input $E[n_1, n_2]$ and $K[n_1, n_2]$ is the diffractive phase kernel in spatial domain given by:

$$K[n_1, n_2] = FFT^2 \{ e^{j\phi[u_1, u_2]} \} \qquad (4.50)$$

where u_1, u_2 are discrete domain spatial frequency variables.

CHAPTER 5

Applications of Physics-inspired Computational Imaging Algorithms in Drug Discovery and Cancer Immunotherapy

In this chapter, we demonstrate efficacy of physics-inspired computational imaging algorithms in solving a variety of problems for different real-world applications. Specifically, we demonstrate two software tools, 1) CytoLive, an award-winning real-time live cell tracking tool used to accelerate Drug Discovery by analyzing time-lapse microscopy videos and 2) CytoEye, a high throughput computational pathology tool for analysis of immuno-histochemistry images useful in tailoring personalized cancer immunotherapy treatments.

5.1 CytoLive

With the recent development of fast and highly sensitive microscopy techniques, cell biologists are now capable to extract more information about the anatomic as well as dynamic behavior of cellular structures. This information is critical for advancements in the field of medicine and biology [112]. The live cell imaging experiments generate extensive amount of data which cannot be fully processed by a human observer. In order to completely retrieve the information contained in the massive data, automated computerized image analysis tools are required. Additionally, the images produced in these experiments are noisy which calls for development of advanced image processing algorithms.

For this, we develop CytoLive [113], an award-winning real-time live cell tracking tool utilizing our NLSE-guided algorithms to analyze time-lapse microscopy videos acquired under low light conditions. The tool was developed in Jalali lab and in collaboration with the Institute for Quantitative and Computational Biosciences (QCBio) here at UCLA.

5.1.1 Problem Statement

Using CytoLive, the problem we are addressing is the high cost of drug development by being able to catch the failures at an early stage using computational imaging. The reason behind the high cost of drug development is the high failure rate of almost 96 percent, including a 90 percent failure rate during the clinical development stage [114–116]. This failure rate could be controlled by taking several measures. For example, having a better representation by enlarging the patient cohort. Another means to regulate the failure rate during the early clinical drug development phase is to comprehensively understand the underlying biological mechanisms of cell behavior to a targeted drug [117, 118].

This understanding of underlying biological mechanisms for accelerating drug discovery is achieved through tracking live cells by time-lapse microscopy [119] over a long period of time. Researchers adopt live-cell imaging as they are more focused on the study of cellular functions or behaviors and pathways to gain a deeper understanding of disease mechanisms and responses to specific tailored treatments. Live-cell imaging is a key approach to getting this information. Unlike traditional fixed-endpoint cell assays, which only give a point-in-time snapshot of cellular responses, live-cell imaging provides a fuller, more realistic picture of the effects of perturbations, visualization of cellular behaviors such as division, movement, and apoptosis. And by examining how the cells are responding to a targeted drug over a long period of time, drug development can be achieved efficiently.

Time-lapse microscopy has been successfully used in discovery of self-renewing division mechanism of muscle satellite cells in [120], for proof of the existence of hemogenic endothelium in [121], for investigating the mechanisms of germ layer formation in [122], among various other applications. Extracting accurate information of cellular responses to the targeted drug is critical and must be done in a timely fashion with minimum human intervention.

However, live-cell imaging is technically challenging [123, 124] as cells are mostly water which makes them transparent. At the same time, cell viability is the most important aspect when performing live cell experiments to ensure that the physiological and biological processes that are under investigation are not altered in any way. In order to prevent cells from being killed because of over exposure to light the long duration experiments are done in a low-light-level environment which makes these images inherently low contrast, noisy and have uneven illumination which lowers down the accuracy of any cell detection method [125].
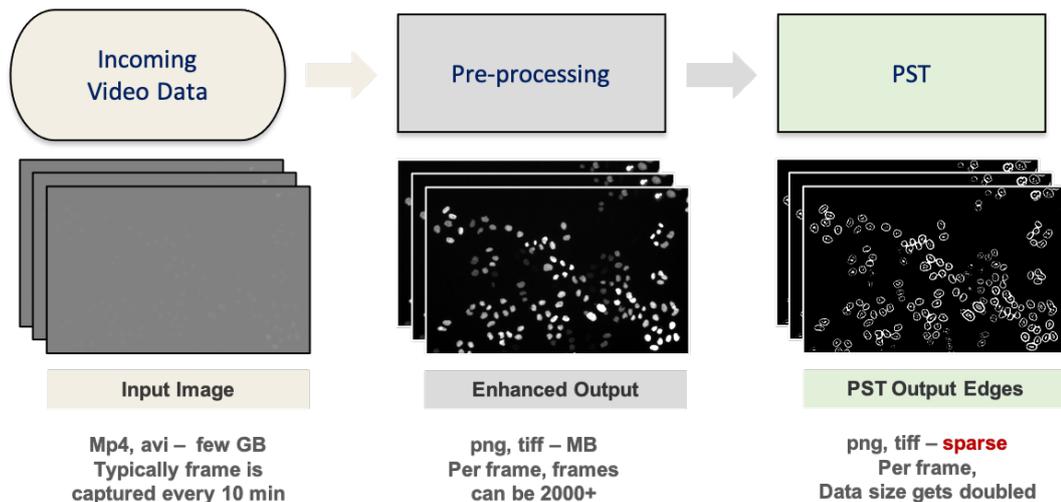
Figure 5.1: **CytoLive:** The software uses Phase Stretch Transform (PST) on the contrast enhanced live cell images and extract low contrast cells, respectively. The PST edge maps and the contrast enhanced live cell images are used to detect cells with high accuracy and subsequently tracked over time and space.

5.1.2   Proposed Solution

CytoLive is a powerful tool in quantitative cell biology and drug discovery as it preserves inherent cell behavior by overcoming photo-toxicity and photo-bleaching and allows live cell experiments at low light levels, as shown in Figure 5.1. The central part of CytoLive exploits the intrinsic equalization ability of physics-inspired edge detection method, Phase Stretch Transform (PST) for efficient localization of low contrast cellular structures. As we are able to capture these low contrast cells, we can track them with high accuracy both over time and space and even in the presence of low light level conditions and thereby, preserve the inherent cell behavior. Therefore, CytoLive offers best in class performance for low contrast cell detection compared to conventional cell detectors that have a very limited dynamic range and also no equalization ability to overcome the limitations imposed by the imaging systems.

In the first step, live cells are detected with high accuracy in individual frames throughout a time-lapse image sequence by using both the PST edge maps as well as the contrast enhanced images. Also, a hyper-dimensional feature data set for each identified cell is produced. In the second step, the tool links the detected cells between consecutive frames by computing over hyper-dimensional data set that comprises of geometrical as well as morphological features derived in the first step. The missing cell events are also included by linking the resulting track segments in a loop to generate complete trajectories. The proposed software tool has also reduced the overall computational time for manual annotating cellular structures and generating the lineage tree for a single live cell experiment (conducted in duration of 5 days) from 30 days to few minutes. Furthermore, our hyper-dimensional live cell tracking software tool, CytoLive, is more accurate than existing live cell tracking software tool for tracking low contrast cells while minimizing the amount of manual intervention.

5.1.3 Analysis using CytoLive

CytoLive generates the cell-parent-daughter information as well as a rich feature dataset that comprises of individual cell area, eccentricity, velocity exported frame by frame and also averaged over a period of frames, see Figure 5.2. This rich dataset can help in understanding the mechanisms that control critical biological events of cell populations, such as proliferation, differentiation, or cell fate decisions, which are crucial for understanding the cell behavior to a drug. CytoLive complements live cell microscopy by providing insights on cell responses to a drug, the population information at the single-cell level as well as stratification of cells based on the overall response. This in turn could increase the drug development success rate, reduce the overall cost and certainly, help in bringing the drug earlier to market.

A	B	C	D	E	F	G	H	I	J	K
Frame Number	Cell ID	Parent ID	Center-X	Center-Y	Area	Intensity	Eccentricty	Velocity-X	Velocity-Y	Velocity
0	1	0	12.8446	184.189	148	85.8378	2.15179	0	0	0
0	2	0	22.2195	213.049	123	101.374	1.09085	0	0	0
0	3	0	95.4696	411.957	115	104.487	1.08859	0	0	0
0	4	0	291.835	121.391	115	103.635	1.04302	0	0	0
0	5	0	310.459	22.0367	109	97.5505	1.02571	0	0	0
0	6	0	426.84	13.352	125	91.632	1.06476	0	0	0
1	7	1	12.8446	184.189	148	85.8378	2.15179	0	0	0
1	8	2	22.2195	213.049	123	101.374	1.09085	0	0	0
1	9	3	95.4696	411.957	115	104.487	1.08859	0	0	0
1	10	4	291.835	121.391	115	103.635	1.04302	0	0	0
1	11	5	310.459	22.0367	109	97.5505	1.02571	0	0	0
1	12	6	426.84	13.352	125	91.632	1.06476	0	0	0
2	13	7	9.71667	185.142	120	86.3417	1.00096	-1.563965	0.4765	1.63494305
2	14	8	21.5081	215.815	124	98.2016	1.02124	-0.3557	1.383	1.42800963
2	15	9	87.3707	410.44	116	102.767	1.05117	-4.04945	-0.7585	4.1198747
2	16	10	301.073	12.5182	110	85.3	1.08115	4.619	-54.4364	54.6320127
2	17	11	301.67	120.635	115	102.078	1.03284	-4.3945	49.29915	49.4946242
2	18	12	423.784	8.72973	111	87.8919	1.17833	-1.528	-2.311135	2.77058279

Figure 5.2: **Hyper-dimensional feature set generated using CytoLive:** CytoLive generates the cell-parent-daughter information as well as a rich feature dataset that comprises of individual cell area, eccentricity, velocity exported frame by frame and also averaged over a period of frames, useful in understanding the mechanisms that control critical biological events of cell populations.
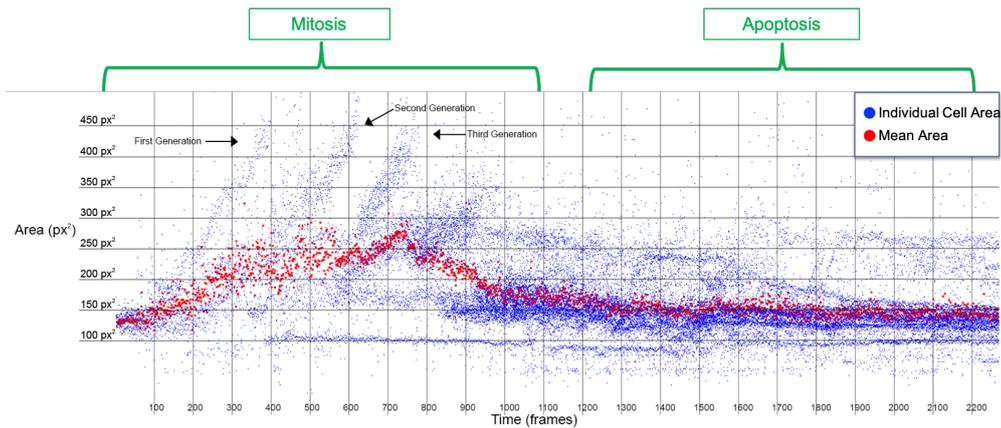
Figure 5.3: **Tracking cellular behaviour through its evolution life cycle using CytoLive:** Cells experience an increase in cell size during the Mitosis phase (proliferation of cells) and subsequently, a decrease in cell size during the Apoptosis phase, initiated because of the photo-toxicity.

We demonstrate in Figure 5.3 that the cells experience an increase in cell size during the Mitosis phase and a decrease in size during the Apoptosis phase. Particularly, we show evolution of each cell through its life cycle in a large population. To show this, we plot the cell area as it changes over time in different phase of the cell life cycle. Each blue dot represents area of an individual cell present in all frames and each red dot is the mean area of all cells present at a particular instant of time. As shown, the mean area first increases with time, which corresponds to the mitosis phase. This is biologically consistent as just before the division cells experience an increase in size and thereby, an increase in the cell area.

Now after a certain point of time, because of the photo-toxicity cells start to die, this is referred as the Apoptosis phase. As cells start dying, they first shrink in size, and hence area decreases and eventually they die or become blur in the background. Another key observation is the sudden peaks of increment in the total number of cells, as indicated in the Figure 5.3. These peaks correspond to various generations of cell-daughter, during the Mitosis phase of the cell life cycle.

114
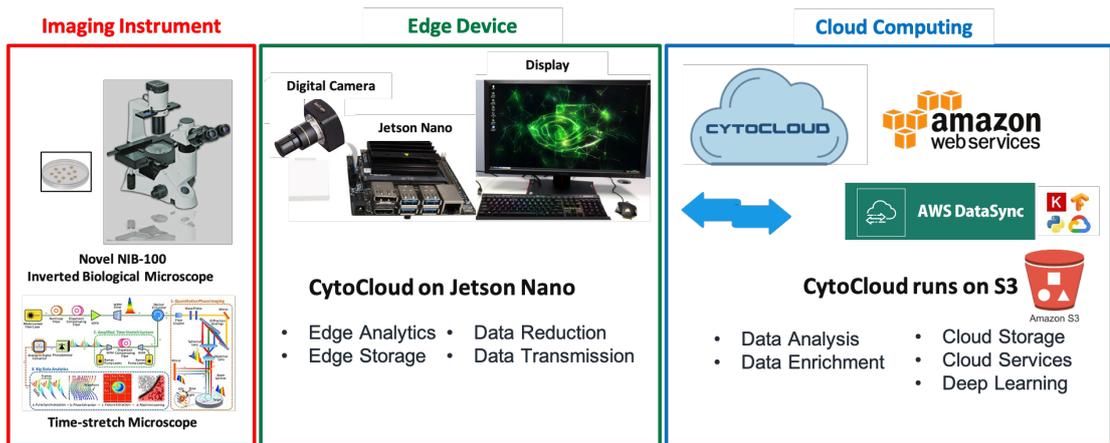
Figure 5.4: **Building a GPU-enabled Cellular Imaging workflow on Amazon Web Services (AWS)**: CytoCloud is cloud-based computing service that will run CytoLive on AWS as well as an edge device. Biologists can use the service to analyze their live cell experiments using their own imaging acquisition systems or customized imaging hardware's developed in our lab such as time-stretch camera [2].

5.1.4 Future Work

As part of our ongoing efforts aimed towards wider adoption of CytoLive for live cell tracking, we are collaborating with Amazon Web Services (AWS), to develop a cloud based modular and scalable architecture, called as CytoCloud that extends right from image acquisition system, followed by low level processing on a edge device (for example an NVIDIA jetson nano development kit) and the computationally expensive computing is done on the cloud service (on AWS). The entire system will be integrated as an online platform allowing biologists to conveniently check their experiment progress online by building a GPU-enabled Cellular Imaging workflow on AWS, as proposed in Figure 5.4.

5.2 CytoEye

Immunotherapy techniques date back at least a century, but recent breakthroughs in the field have sparked hope for their ability to offer less-invasive, durable treatment for a wide variety of cancer types [126–128]. A major current hurdle for several leading therapies, however, lies in their variable patient efficacies as these therapies are influenced by a heterogeneous combination of health, immune, and tumor factors [129]. Additionally, some initial responders eventually develop resistance to these therapies leading to relapse of the disease [130, 131]. For instance, 30 to 55 percent of patients suffering from early-stage non-small cell lung carcinoma (ES-NSCLC) still tend to have either local or distant recurrence even after systemic therapy [132, 133].

Due to this uncertainty, patient stratification for personalized immunotherapy treatments approaches guesswork. In-fact till date, there is no clinically-used well established method that is accurate and reproducible to stratify patients who will benefit from the given personalized therapy versus patients who may have a high risk of disease relapse and therefore, need early intensification of treatment versus low-risk immune active patients who can be treated with surgery alone. As such, understanding the immunosuppressive mechanisms by which tumor cells inhibit immune cells in a tumor micro-environment (TME) is a crucial step towards personalized immunotherapy.

With this objective in mind, we develop, CytoEye [134], an immunotherapy toolbox for discovery of immune cells expressions and quantification of interactions of immune cells with cancer cells by computationally efficient analysis of giga-pixel sized pathology images of TME using physics-inspired computational imaging algorithms. The generated quantitative insights is useful for stratifying patients for personalized immunotherapy treatments. This project was done in collaboration with the Parker Center for Cancer Immunotherapy (PICI) at UCLA.

116

5.2.1 Problem Statement

In recent years, immunotherapy utilizing checkpoint blockade, particularly, therapies that target the programmed death-1 (PD-1) receptor has risen in prominence to become one of the fundamental methodologies of cancer immunotherapy [135–137], leading to durable therapeutic responses not typically seen with traditional cytotoxic anticancer agents. The biological impetus is the co-inhibitory signaling pathway, which serves as a source of immune tolerance for native cellular tissue. For instance, by up-regulation of PD-1 ligand (PD-L1) and its ligation to PD-1 on immune cells, the cancer cells limit the host immune response as shown in Figure 5.5. As these pathways do not delineate between cancerous and non-cancerous tissue, leading to immunotolerance for malignant cells which is termed as adaptive immune resistance. Immune checkpoint blockade therapy works by disrupting this signaling by injection of antibodies (Anti-PD-1 antibodies for immune cells and Anti-PD-L1 antibodies for cancer cells) and this helps in promoting immune response in the tumor micro-environment as shown in Figure 5.5(B).
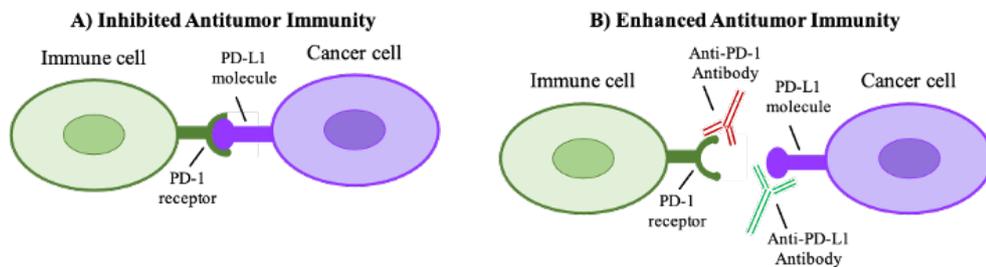
Figure 5.5: **Types of immune cell and cancer cell interactions in a Tumor Micro-Environment (TME).** A) Example of inhibited antitumor immunity where cancer cells mimic benign interaction with immune cells through PD1-PDL1 interaction and B) By uncloaking this PD1-PDL1 interaction of immune cell with Anti-PD-1 anti-body and cancer cell with Anti-PD-L1 antibody, antitumor immunity can be en-hanced. PD-1 : Programmed Death 1 and PD-L1: Programmed Death Ligand 1.

While the availability of advanced whole slide scanners has enabled complete digitization of pathology slides, the development of automated histo-pathology image analysis tools has been challenging. On the other hand, immune checkpoint blockade treatment decisions are often dependent on quantification of PD-1/PD-L1 expression and immune cell infiltration as measured by traditional immuno-histochemistry (IHC) microscopy images, the gold standard for digital pathology.

One of the critical steps in most state-of-the-art automated image analysis is the cell segmentation process, which is often based on color intensity analyses. This requires the pathologist to manually define the intensity distributions for each type of immunostain, cell or tumor type to achieve acceptable accuracy. As such a manual inspection of the gigapixel sized pathology images is an extremely laborious, a time-consuming and potentially error-prone approach that is also susceptible to potentially erroneous diagnosis, high variability and limited reproducibility (and in many cases is performed without any computational assistance).

Additionally, while some computational methods exist for the analysis of these images for research purposes of patients who do not respond to treatment or who develop relapsed disease on therapy, these also require extensive expert training to operate, have limited accuracy and can have varying levels of subjectivity in the analysis. This calls for development of computationally efficient algorithms that are capable of quickly and accurately completing annotation and feature extraction.

Also, due to the large variety of imaging modalities (conventional brightfield and fluorescence whole-slide scanning as well as multispectral imaging), it is crucial to tailor pathology image analysis methods that can generalize to any modality and capable of quickly and accurately completing cell annotation-feature extraction and compute spatial tumor–host interactions to answer specifics aims of pathology in question.

5.2.2 Proposed Solution

In this section, we discuss CytoEye [134], a novel computational pathology pipeline tailor-made for the quantitative analysis of pathology images with the ultimate aim of inferring patient response that also mitigates the computational overload of analyzing giga-pixel sized pathology images. Quantitative features extracted by this tool have the capability to predict whether or not patients respond to therapy – an important step toward personalized cancer immunotherapy.
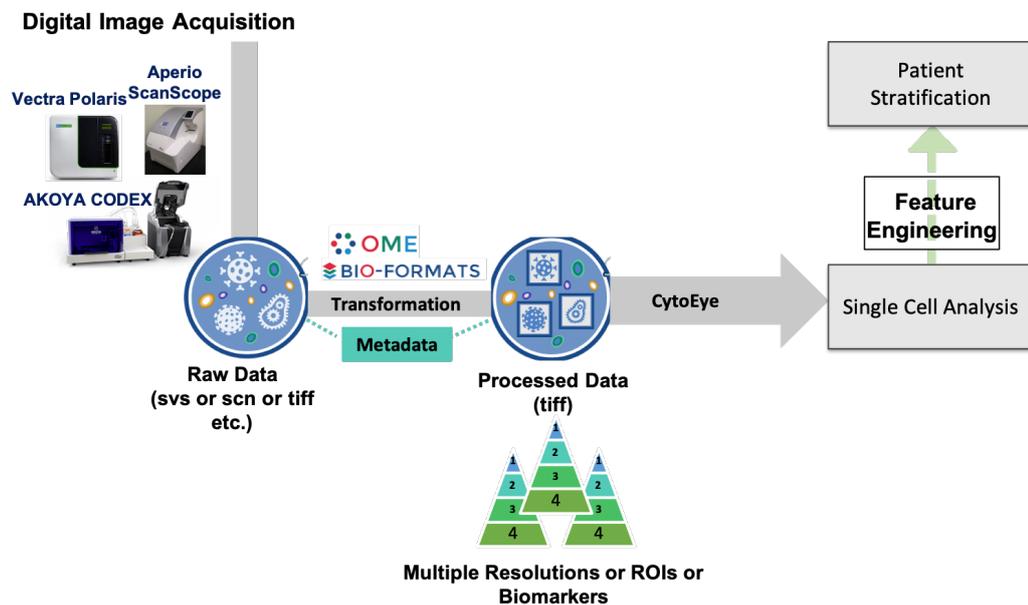
Figure 5.6: **Overview of workflow of CytoEye:** CytoEye consists of single cell analysis module that produces quantitative features for each patient IHC image data useful in patient stratification.

A high level overview of our workflow is shown in Figure 5.6. We start with raw digital images acquired from bright-field instruments such as Aperio ScanScope or multi spectral imaging systems such as Vectra or CODEX. These images are often stored in the proprietary format of the instrument, for instance in case of Aperio Scan Scope, images are saved as ScanScope Virtual Slide (SVS) or SCN for Vectra. The first step of our workflow is to convert these images in a more

generic or standardized and open format such as tiff. This is done using the open-source Bio-Formats software tool [138] developed by the OME Open Microscopy Environment consortium that reads the proprietary microscopy image data. It also reads in the metadata information consisting of levels of resolution at which the images were captured or information on the bio-marker panel considered in the given study while the images were acquired etc. Once all this information is read, the image data is then converted and stored in a suitable format which acts as an input to CytoEye module for single cell analysis. This produces cell segmentation results and single cell quantitative features. These features are then aggregated together to generate analytics for each patient IHC image data useful in patient stratification.

We also demonstrate that CytoEye features extracted from digitized pathology images capture the interactions between cell types (immune-tumor interactions) and devise a new way of constructing graphs that may be predictive of the patient response to PD1-PDLI immunotherapy. We also show that unified framework of CytoEye also takes care of the limitations imposed by the method in which the images are taken. To explain the usefulness of this feature, we consider the case of a whole slide scanner such as Aperio ScanScope system that scans a single microscope slide at a given time. As such we have a set of images for a single patient which are acquired by staining various slices of tissue with a given bio-marker panel used in the study. Additionally, through this understanding of the mechanisms by which immune cells and cancer cells interact in TME has the potential to aid the development of new therapeutics using CytoEye.

As an example study, we analyzed IHC data collected from patients with advanced melanoma receiving immunotherapy directed against PD-1/PD-L1 pathway such that diagnostic stratification can be done by quantification of PD-1/PD-L1 expression and immune cell infiltration from IHC data. The data used in this analysis consists of 20 patients that were used in the study [135]. Slides were

stained with haematoxylin and eosin, S100, CD8, PD1 and PDL1, shown in Figure 5.7. All patients in the study cohorts underwent mandatory biopsy of a metastatic tumour within 30 days of starting the checkpoint blockade treatment. Clinical tumour responses to pembrolizumab were evaluated using (RECIST) version 1.1 criteria.
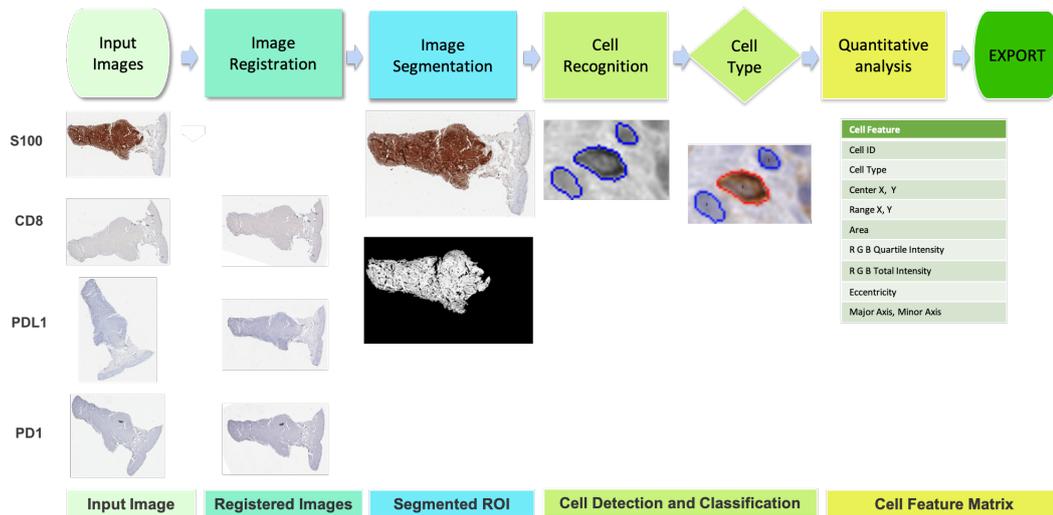
Figure 5.7: **CytoEye workflow for single-cell analysis of images acquired from the whole slide scanner:** The bio-markers images are first registered for spatial localization followed by region-of-interest detection, cell segmentation and feature extraction.

5.2.3 Single-cell Analysis using CytoEye

The first step is to register all bio-marker images with a given bio-marker. In our case, we use S100 as the fixed image and register the other images with S100. S100 expression is used to define the invasive margin and the tumor area. The next step is to apply cell segmentation (watershed and otsu's segmentation) to each of the bio-marker images and identify cells both in the invasive margin and the tumor area. Then based on the presence or absence of the bio-marker expression, the segmented cells are classified as + or -. Finally, a hyper-dimensional feature

matrix is generated corresponding to each cells, as shown in Figure 5.7.
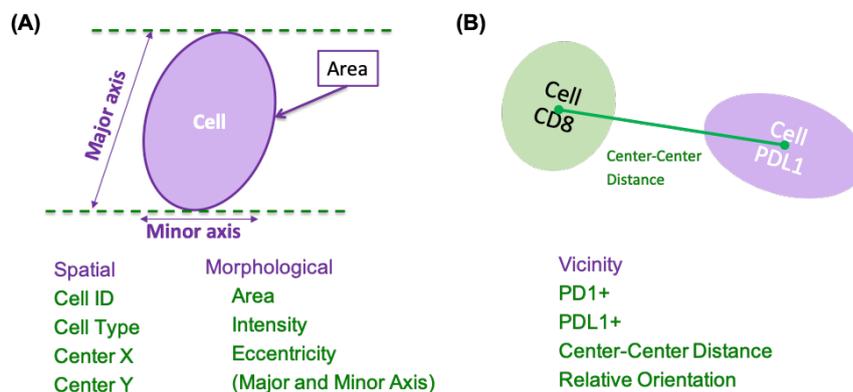
Figure 5.8: **Hyper-dimensional feature computed using CytoEye:** The features are broadly categorized as spatial, morphological and vicinity and are critical to study the immune cell- cancer cell interactions.

The hyper-dimensional feature are broadly categorized as spatial, morphological and vicinity, as shown in Figure 5.8. While the spatial features help in localization of bio-marker expression and present us pictorial representation of distribution of bio-marker across the tissue slice, morphological features are calculated after the cell segmentation is done. The morphological features comprise of features such as area, intensity and eccentricity. All these features are important to study the cell-cell interaction, however eccentricity feature is very useful in defining the PD1-PDL1 and immune cell interaction along a preferred direction.

After, the localization of each bio-marker expression in each of the bio-marker images, cells in the vicinity across different bio-markers are analyzed which leads us to the definition of vicinity characteristics, as shown in Figure 5.8 (B). For examples, given that there is a CD8+ cell, find if there is a PD1+ or PDL1+ cell in the cell neighbourhood and compute, the center to center distance between these two cells. All this is used to define candidate cells with expressions supporting either the case of the enhanced antitumour immunity or inhibited antitumour

immunity, as shown previously in Figure 5.5.

A) Aggregated Features

1. CD8+ / Total Tissue Region
2. S100+ CD8+ / Total S100 Region
3. CD8+ Area / Total Tissue Region
4. S100+ CD8+ Area / Total S100 Region
5. CD8+ / Total Cells
6. S100+ CD8+ / S100+
7. PD1+ / Total Tissue Region
8. S100+ PD1+ / Total S100 Region
9. PD1+ Area / Total Tissue Region
10. S100+ PD1+ Area / Total S100 Region
11. PD1+ / Total Cells
12. S100+ PD1+ / S100+
13. PDL1+ / Total Tissue Region
14. S100+ PDL1+ / Total S100 Region
15. PDL1+ Area / Total Tissue Region
16. S100+ PDL1+ Area / Total S100 Region
17. PDL1+ / Total Cells
18. S100+ PDL1+ / S100+
19. CD8+ PD1+/ CD8+
20. CD8+ PDL1+/ CD8+
21. PD1+ CD8+ / PD1+
22. PD1+ PDL1+ / PD1+
23. PDL1+ CD8+ / PDL1+
24. PDL1+ PD1+ / PDL1+

Figure 5.9: **Patient stratification using CytoEye:** Quantitative features for each patient IHC data computed from single-cell analysis are aggregated into 24 key feature shown in (A). Patient stratification can be potentially done using CytoEye as evident from the high correlation of these features to patient outcome.

5.2.4 Patient Stratification using CytoEye

Once, all the cells in all the bio-marker images of a patient are analyzed, we aggregate the features into 24 key features as shown in Figure 5.9 (A), for instance the total number of CD8+ cells in the S100+ region. This is done to understand what kind of features have a strong correlation with the patient response. The 24 quantitative features are correlated with the patient outcome using Principal Component Analysis (PCA) and the computed correlation is shown using a heat map in Figure 5.9 (B).

As expected from the underlying biology, patients responding to treatment

showed proliferation of intratumoral CD8+ T cells which means a higher CD8+ density indicating immune cell infiltration and hence probability of patient responding increases. On the other hand, one mechanism by which cancer tissues limit the host immune response is via upregulation of PD-1 ligand (PD-L1) and its ligation to PD-1 on CD8+ T cells (termed adaptive immune resistance) which means that the density of PDL1+ should be negatively correlated as shown here. Therefore, quantitative features extracted by CytoEye have the capability to predict whether or not patients respond to therapy – an important step toward personalized cancer immunotherapy.

5.2.5   Future work

We also demonstrate that CytoEye features extracted from pathology images capture the interactions between cell types (immune-tumor interactions) and devise a new way of constructing graphs that may be predictive of the patient response to PD1-PDLI therapy. However, limitations of this study include the limit of access to more extensive patient cohort that results in a low accuracy for patient stratification using Graph Convolutional Networks [139,140]. Nevertheless, we have been able to identify a novel way to create patient graphs computed using quantitative cell features from CytoEye that may potentially enable patient stratification as shown in Figure 5.10. Furthermore, the concept of creating such patient graphs could easily be adapted for the analysis of other cell–cell interactions in TME.

5.3   Conclusions

We present here software tools utilizing NLSE-guided algorithms to solve problems associated with imaging modalities for time lapse microscopy and immunohisto-chemistry useful in drug discovery and cancer immunotherapy.

124
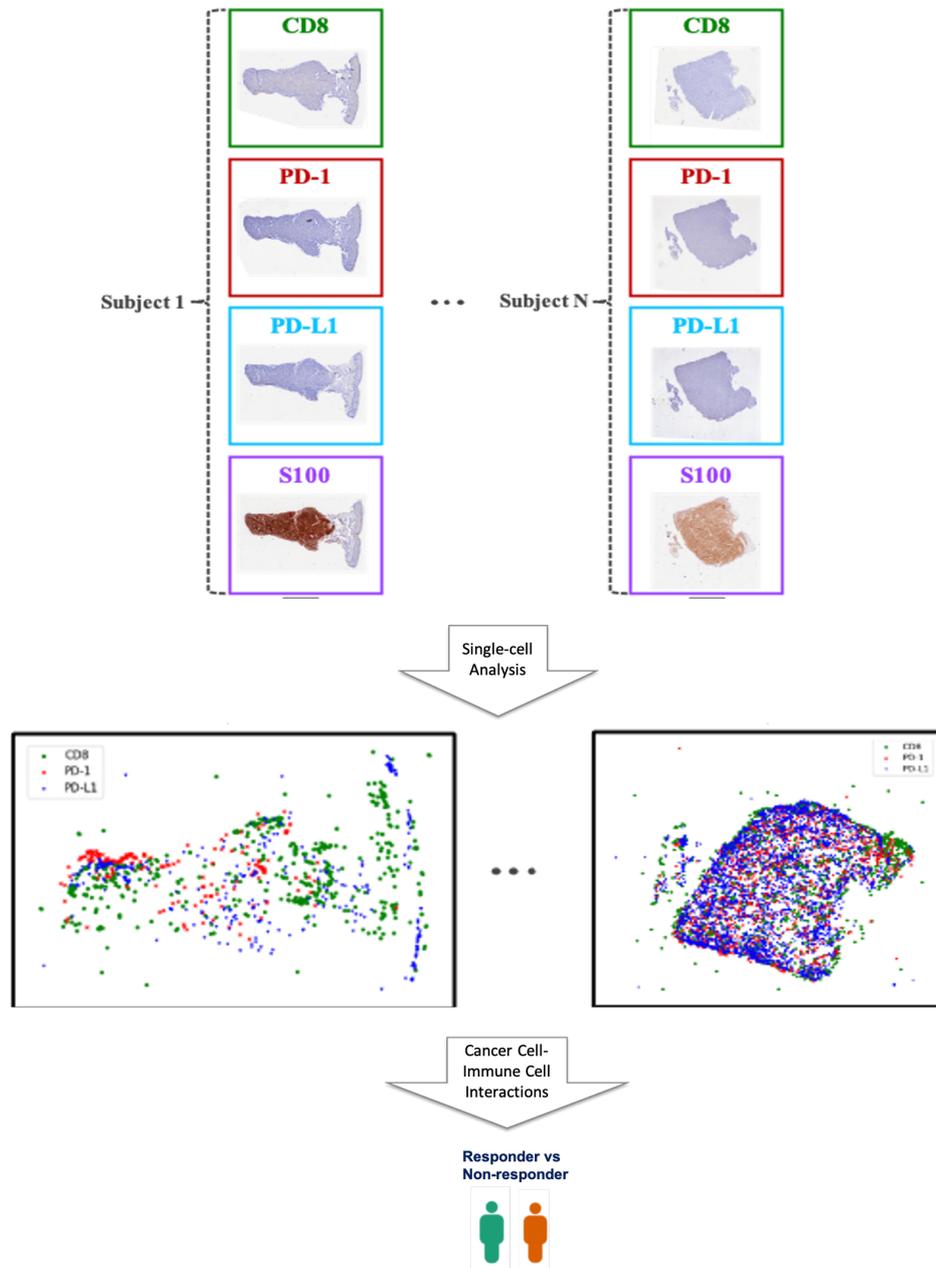
Figure 5.10: **Constructing patient graphs using CytoEye features:** Quantitative features for each patient IHC data computed from single-cell analysis are used to create graphs. The graphs are then fed to a Graph Convolutional Networks for learning underlying bio-marker interactions for predicting patient outcome.

CHAPTER 6

Conclusion and Future Works

6.1   Conclusion

The thesis demonstrated the use of Non-Linear Schrodinger Equation (NLSE),
the master equation in optical physics, to craft fundamentally and qualitatively
different and novel physics-inspired computational imaging algorithms. Specifi-
cally, we described a unified nonlinear $\underline{S}$chrodinger framework for $\underline{I}$mage and $\underline{V}$ideo
$\underline{E}$nhancement (SIVE) by subjecting the image to virtual propagation through a
two-dimensional optical medium such that the optical properties of the medium
can be engineered to obtain desired response. The propagation can either imparts
a brightness-dependent pixelwise phase in case of SIVE as shown in Chapter 2) or
a frequency-dependent pixelwise phase in case of Phase Stretch Transform (PST)
and Phase-stretch Adaptive Gradient-field Extractor (PAGE), as shown in Chap-
ter 3 and 4 respectively. Due to this transformation, in every case, the real valued
input image is turned into a complex-valued quantity containing both real and
imaginary components.

Specifically, we showed that SIVE, a new computationally efficient contrast
enhancement algorithm, achieves superior color gamut performance, and is able to
support real-time video enhancement at 4K and 8K resolutions. We also show how
the NLSE operation leads us to PST, a physics-inspired computational imaging
algorithm with exceptional performance in low light levels. PST emulates 2-
dimensional diffraction followed by coherent detection and has applications for

feature detection in visually impaired images. We also presented PAGE, a new feature decomposition method that emulates birefringent diffraction followed by coherent detection and extracts semantic information at different orientation, scale and frequency.

Lastly, we also demonstrated efficacy of these algorithms in solving a variety of problems for different real-world applications in Chapter 5. Specifically, we discussed CytoLive - an award-winning clinical research tool to complement live cell microscopy for accelerating drug development. We also presented CytoEye – an immunotherapy toolbox for quantifying cancer cells and immune cells interaction in a tumor microenvironment. Quantitative features extracted by this tool have the capability to predict whether or not patients respond to therapy – an important step toward personalized cancer immunotherapy.

6.2 Future works

The physics-inspired computational imaging algorithms, as discussed in the text, have exceptional and unprecedented performance compared to the conventionally designed counterparts. Most importantly as these algorithms emulate a physical phenomenon, in certain cases, these algorithms have the potential to be implemented in physical optics. Future work involves implementing these tools on a physical system. We discuss possible experiments in this section, that can be investigated for studying the efficacy of proposed algorithms for computing at the speed of light.

Metaphoric physical interpretation of the diffractive algorithm:

First, we describe a metaphorical interpretation of our diffractive algorithm (from Chapter 4) on a physical system potentially using metamaterials. Metamaterials are composite electromagnetic materials that have the ability to transform an input wave to not just a converging wave on the other side but a completely

arbitrarily complex wavefront by imparting an abrupt, spatially varying phase profile on the incident light which enables us to perform powerful wave processing using metamaterials. These sub-wavelength scale scatterers and have been used to for a large number of applications ranging from achieving negative refraction in lens [141], to function as a real-time tunable, spectrally sensitive spatial masks for terahertz imaging [142] and also to perform diffractive optics by designing meta-materials with a specific phase profile [143]. These capabilities of metamaterials allows us to formulate a metaphorical interpretation of our method to compute gradient fields using the principle of diffraction followed by coherent detection, shown in Figure 6.1.
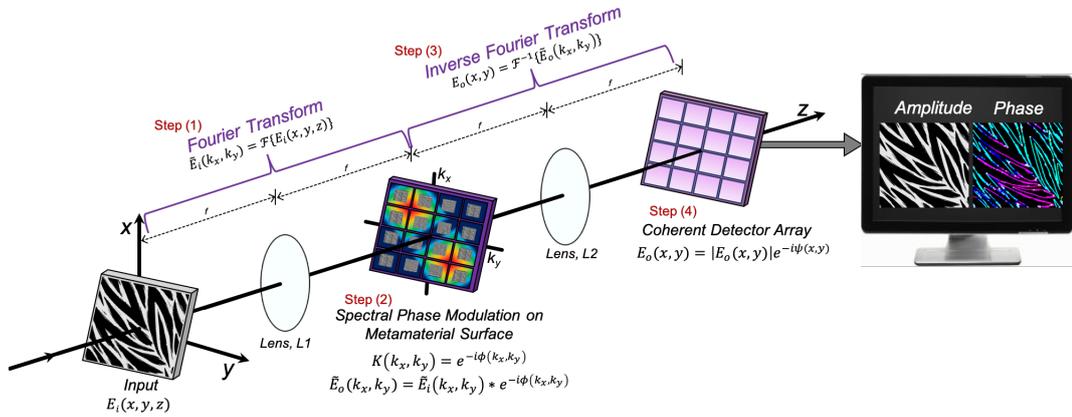
Figure 6.1: **Metaphoric interpretation of our diffractive algorithm in terms of physical optics in 3D space:** It shows a metaphoric interpretation of our D-dimensional diffraction algorithm in 3D space. As shown, the first step is to convert the object image from spatial domain to Fourier domain using a lens (L_1). The Fourier spectra of the object image is modulated using a metamaterial which is designed to impart specific phase as function of frequency (spatial). Next step is to convert the modulated spectra back to spatial domain using another lens (L_2). The diffracted output is captured through an array of single pixel coherent detector array. The phase of the output contains the gradient information of the input signal.

The first step of our proposed diffractive method, from Chapter 4, is to translate the spatial input to frequency or spectral domain. This could be achieved by constructing an optical 4F such that the distance between the object plane and the first lens L_1 is F, followed by a metamaterial surface placed on the Fourier plane and therefore, at distance F from the lens L_1. As shown, the first step is to convert the object image from spatial domain to Fourier domain using a lens (L_1).

We propose to use a metamaterial surface which is designed to have specific diffractive property such that an engineered phase could be imparted to the spectrum of our input as earlier demonstrated in [143, 144]. This modulated input spectrum then propagates through another lens L_2 at distance F which is the operation of inverse Fourier transform. Finally, we can have a single pixel phase detector array that captures the magnitude as well as the phase of modulated spatial field such that the amplitude of the output is same as the input while the phase captures features of interest from the input.

Optical Differentiation using Dispersion for Signal Classification:
The inherent reconfigurability of the differentiation method using group delay dispersion, discussed in Chapter 3, can be examined for signal classification. The group delay dispersion property of the optical medium can be configured as required [145]. Optical experiments can be performed with tunable group delay dispersion to measure optical differentiation of signals. Extracted orders of differentiation of the input signal can be used for subsequent analysis of signals example for optical signal classification.

References

[1] Asuri S Bhushan, F Coppinger, and Bahram Jalali. Time-stretched analogue-to-digital conversion. *Electronics Letters*, 34(9):839–841, 1998. ix, 2, 3, 46

[2] K Goda, KK Tsia, and B Jalali. Serial time-encoded amplified imaging for real-time observation of fast dynamic phenomena. *Nature*, 458(7242):1145–1149, 2009. ix, xxviii, 2, 3, 115

[3] Daniel R Solli, Claus Ropers, Prakash Koonath, and Bahram Jalali. Optical rogue waves. *Nature*, 450(7172):1054–1057, 2007. ix, 2, 3, 46, 77

[4] Christophe Szwaj, Clément Evain, Marc Le Parquier, Serge Bielawski, Eléonore Roussel, Laurent Manceron, Jean-Blaise Brubach, Marie-Agnès Tordeux, Jean-Paul Ricaud, Lodovico Cassinari, et al. Unveiling the complex shapes of relativistic electrons bunches, using photonic time-stretch electro-optic sampling. In *2016 IEEE Photonics Society Summer Topical Meeting Series (SUM)*, pages 136–137. IEEE, 2016. ix, 2, 3, 46

[5] Georg Herink, Felix Kurtz, Bahram Jalali, Daniel R Solli, and Claus Ropers. Real-time spectral interferometry probes the internal dynamics of femtosecond soliton molecules. *Science*, 356(6333):50–54, 2017. ix, 2, 3, 77

[6] Eric D Diebold, Brandon W Buckley, Daniel R Gossett, and Bahram Jalali. Digitally synthesized beat frequency multiplexing for sub-millisecond fluorescence microscopy. *Nature Photonics*, 7(10):806–810, 2013. ix, 2

[7] Georg Herink, B Jalali, Claus Ropers, and Daniel R Solli. Resolving the build-up of femtosecond mode-locking with single-shot spectroscopy at 90 mhz frame rate. *Nature Photonics*, 10(5):321–326, 2016. ix, 2, 3, 77

[8] Claire Lifan Chen, Ata Mahjoubfar, Li-Chia Tai, Ian K Blaby, Allen Huang, Kayvan Reza Niazi, and Bahram Jalali. Deep learning in label-free cell classification. *Scientific reports*, 6(1):1–16, 2016. ix, 2, 3, 46, 77

[9] Ata Mahjoubfar, Claire Lifan Chen, and Bahram Jalali. *Artificial Intelligence in Label-free Microscopy*. Springer, 2017. ix, 2, 3, 46, 77

[10] Yueqin Li, Ata Mahjoubfar, Claire Lifan Chen, Kayvan Reza Niazi, Li Pei, and Bahram Jalali. Deep cytometry: deep learning with real-time inference in cell sorting and flow cytometry. *Scientific reports*, 9(1):1–12, 2019. ix, 2

[11] Robert Hummel. Image enhancement by histogram transformation. *Unknown*, 1975. x, xi, xiii, xiv, xv, xvi, xxxi, xxxii, 10, 18, 19, 28, 31, 32, 33, 37, 39, 40, 69

[12] Karel Zuiderveld. Contrast limited adaptive histogram equalization. *Graphics gems*, pages 474–485, 1994. x, xi, xiii, xiv, xv, xvi, xxxi, xxxii, 10, 18, 19, 28, 31, 32, 33, 37, 39, 40

[13] Anil Singh Parihar, Om Prakash Verma, and Chintan Khanna. Fuzzy-contextual contrast enhancement. *IEEE Transactions on Image Processing*, 26(4):1810–1819, 2017. x, xi, xiii, xiv, xv, xvi, xxxi, xxxii, 10, 11, 18, 19, 28, 31, 32, 33, 37, 39, 40

[14] Xueyang Fu, Delu Zeng, Yue Huang, Xiao-Ping Zhang, and Xinghao Ding. A weighted variational model for simultaneous reflectance and illumination estimation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 2782–2790, 2016. x, xi, xiii, xiv, xv, xvi, xxxi, xxxii, 11, 18, 19, 28, 31, 32, 33, 37, 39, 40

[15] Xiaojie Guo, Yu Li, and Haibin Ling. Lime: Low-light image enhancement via illumination map estimation. *IEEE Transactions on image processing*, 26(2):982–993, 2016. x, xi, xiii, xiv, xv, xvi, xxxi, xxxii, 11, 18, 19, 25, 28, 31, 32, 33, 37, 39, 40

[16] Feifan Lv, Feng Lu, Jianhua Wu, and Chongsoon Lim. Mbllen: Low-light image/video enhancement using cnns. In *BMVC*, page 220, 2018. x, xi, xiii, xiv, xv, xvi, xxxi, xxxii, 11, 18, 19, 28, 31, 32, 33, 37, 39, 40

[17] Wenjing Wang, Chen Wei, Wenhan Yang, and Jiaying Liu. Gladnet: Low-light enhancement network with global awareness. In *2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018)*, pages 751–755. IEEE, 2018. x, xi, xiii, xiv, xv, xvi, xxxi, xxxii, 11, 18, 19, 28, 31, 32, 33, 37, 39, 40

[18] Chunle Guo, Chongyi Li, Jichang Guo, Chen Change Loy, Junhui Hou, Sam Kwong, and Runmin Cong. Zero-reference deep curve estimation for low-light image enhancement. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 1780–1789, 2020. x, xi, xiii, xiv, xv, xvi, xxxi, xxxii, 11, 18, 19, 23, 28, 30, 31, 32, 33, 37, 39, 40

[19] David Hasler and Sabine E Suesstrunk. Measuring colorfulness in natural images. In *Human vision and electronic imaging VIII*, volume 5007, pages 87–95. International Society for Optics and Photonics, 2003. xi, xxx, xxxi, xxxii, 19, 20, 27, 30, 32, 34, 36, 37

[20] Anish Mittal, Rajiv Soundararajan, and Alan C Bovik. Making a "completely blind" image quality analyzer. *IEEE Signal processing letters*, 20(3):209–212, 2012. xi, xxxi, xxxii, 19, 20, 30, 32, 34, 37, 38

131

[21] Chulwoo Lee, Chul Lee, and Chang-Su Kim. Contrast enhancement based on layered difference representation of 2d histograms. *IEEE transactions on image processing*, 22(12):5372–5384, 2013. xiii, 26, 28

[22] Bahram Jalali, Madhuri Suthar, Mohamad Asghari, and Ata Mahjoubfar. Optics-inspired computing. In *Proceedings of the 5th International Conference on Photonics, Optics and Laser Technology*, volume 1, pages 340–345, 2017. xvi, xviii, 15, 41, 42, 43, 46, 52, 53, 63, 71, 76

[23] Madhuri Suthar, Ata Mahjoubfar, Kevin Seals, Edward W Lee, and Bahram Jalaii. Diagnostic tool for pneumothorax. In *2016 IEEE Photonics Society Summer Topical Meeting Series (SUM)*, pages 218–219. IEEE, 2016. xvi, xix, 15, 41, 42, 43, 47, 59, 77

[24] Madhuri Suthar. *Decision support systems for radiologists based on phase stretch transform.* PhD thesis, UCLA, 2016. xvi, xix, 15, 41, 42, 43, 47, 59, 77

[25] Madhuri Suthar, Hossein Asghari, and Bahram Jalali. Feature enhancement in visually impaired images. *IEEE Access*, 6:1407–1415, 2017. xvi, xvii, 8, 15, 41, 42, 43, 46, 48, 57, 71, 76, 84

[26] Tali Ilovitsh, Bahram Jalali, Mohammad H Asghari, and Zeev Zalevsky. Phase stretch transform for super-resolution localization microscopy. *Biomedical optics express*, 7(10):4198–4209, 2016. xvi, 15, 41, 42, 43, 47, 77

[27] Mohammad H Asghari and Bahram Jalali. Edge detection in digital images using dispersive phase stretch transform. *International journal of biomedical imaging*, 2015, 2015. xvi, xvii, 15, 41, 42, 43, 46, 47, 50, 51, 54, 76

[28] Mohammad H Asghari and Bahrain Jalali. Physics-inspired image edge detection. In *Signal and Information Processing (GlobalSIP), 2014 IEEE Global Conference on*, pages 293–296. IEEE, 2014. xvii, 46, 47, 50, 51, 54

[29] Bahram Jalali, Madhuri Suthar, Mohammad Asghari, and Ata Mahjoubfar. Physics-based feature engineering. In *Optics, Photonics and Laser Technology 2017*, pages 255–275. Springer, 2019. xviii, 8, 46, 52, 53, 63, 76

[30] Bahram Jalali, Madhuri Suthar, Mohamad Asghari, and Ata Mahjoubfar. Time stretch inspired computational imaging. *arXiv preprint arXiv:1706.07841*, 2017. xviii, 46, 52, 53, 57, 63

[31] Ibm (2016). bringing big data to the enterprise. https://www-01.ibm.com/software/in/data/bigdata/. 1

[32] F Coppinger, AS Bhushan, and B Jalali. Photonic time stretch and its application to analog-to-digital conversion. *IEEE Transactions on microwave theory and techniques*, 47(7):1309–1314, 1999. 3, 46

[33] Yan Han and Bahram Jalali. Photonic time-stretched analog-to-digital converter: fundamental concepts and practical considerations. *Journal of Lightwave Technology*, 21(12):3085, 2003. 3, 46

[34] W Ng, T Rockwood, and A Reamon. Demonstration of channel-stitched photonic time-stretch analog-to-digital converter with enob 8 for a 10 ghz signal bandwidth. In *Proceedings of the Government Microcircuit Applications and Critical Technology Conference (GOMACTech'14)*, 2014. 3, 46

[35] Keisuke Goda, Ali Ayazi, Daniel R Gossett, Jagannath Sadasivam, Cejo K Lonappan, Elodie Sollier, Ali M Fard, Soojung Claire Hur, Jost Adam, Coleman Murray, et al. High-throughput single-microparticle imaging flow analyzer. *Proceedings of the National Academy of Sciences*, 109(29):11630–11635, 2012. 3

[36] Claire Lifan Chen, Ata Mahjoubfar, and Bahram Jalali. Optical data compression in time stretch imaging. *PloS one*, 10(4):e0125106, 2015. 3, 46, 72

[37] Govind P Agrawal. Nonlinear science at the dawn of the 21st century. *Lecture Notes in Physics*, 542:195–211, 2000. 3, 12

[38] Jalalilabucla. https://github.com/JalaliLabUCLA/. 5

[39] How many photos will be taken in 2021? https://focus.mylio.com/tech-today/how-many-photos-will-be-taken-in-2021. 6

[40] Mark Gamadia, Nasser Kehtarnavaz, and Katie Roberts-Hoffman. Low-light auto-focus enhancement for digital and cell-phone camera image pipelines. *IEEE Transactions on Consumer Electronics*, 53(2):249–257, 2007. 7

[41] Neal Wadhwa, Rahul Garg, David E Jacobs, Bryan E Feldman, Nori Kanazawa, Robert Carroll, Yair Movshovitz-Attias, Jonathan T Barron, Yael Pritch, and Marc Levoy. Synthetic depth-of-field with a single-camera mobile phone. *ACM Transactions on Graphics (ToG)*, 37(4):1–13, 2018. 7

[42] Samuel W Hasinoff, Dillon Sharlet, Ryan Geiss, Andrew Adams, Jonathan T Barron, Florian Kainz, Jiawen Chen, and Marc Levoy. Burst photography for high dynamic range and low-light imaging on mobile cameras. *ACM Transactions on Graphics (TOG)*, 35(6):1–12, 2016. 8

[43] Tarik Arici, Salih Dikbas, and Yucel Altunbasak. A histogram modification framework and its application for image contrast enhancement. *IEEE Transactions on image processing*, 18(9):1921–1935, 2009. 10

[44] Edwin H Land. The retinex theory of color vision. *Scientific american*, 237(6):108–129, 1977. 10

[45] Edwin H Land and John J McCann. Lightness and retinex theory. *Josa*, 61(1):1–11, 1971. 11

[46] Mading Li, Jiaying Liu, Wenhan Yang, Xiaoyan Sun, and Zongming Guo. Structure-revealing low-light image enhancement via robust retinex model. *IEEE Transactions on Image Processing*, 27(6):2828–2841, 2018. 11

[47] Huanjing Yue, Jingyu Yang, Xiaoyan Sun, Feng Wu, and Chunping Hou. Contrast enhancement based on intrinsic image decomposition. *IEEE Transactions on Image Processing*, 26(8):3981–3994, 2017. 11

[48] Eli Schwartz, Raja Giryes, and Alex M Bronstein. Deepisp: Toward learning an end-to-end image processing pipeline. *IEEE Transactions on Image Processing*, 28(2):912–923, 2018. 11

[49] Chen Chen, Qifeng Chen, Jia Xu, and Vladlen Koltun. Learning to see in the dark. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 3291–3300, 2018. 11

[50] Zhetong Liang, Jianrui Cai, Zisheng Cao, and Lei Zhang. Cameranet: A two-stage framework for effective camera isp learning. *IEEE Transactions on Image Processing*, 30:2248–2262, 2021. 11

[51] Andrey Ignatov, Nikolay Kobyshev, Radu Timofte, Kenneth Vanhoey, and Luc Van Gool. Dslr-quality photos on mobile devices with deep convolutional networks. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 3277–3285, 2017. 11

[52] Bin Xiao, Yunqiu Xu, Han Tang, Xiuli Bi, and Weisheng Li. Histogram learning in image contrast enhancement. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, pages 0–0, 2019. 11

[53] Peter E Powers and Joseph W Haus. *Fundamentals of nonlinear optics*. CRC press, 2017. 16

[54] John Kerr. Xl. a new relation between electricity and light: Dielectrified media birefringent. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 50(332):337–348, 1875. 16

[55] Jianrui Cai, Shuhang Gu, and Lei Zhang. Learning a deep single image contrast enhancer from multi-exposure images. *IEEE Transactions on Image Processing*, 27(4):2049–2062, 2018. 17, 24

[56] Jinggang Huang and David Mumford. Statistics of natural images and models. In *Proceedings. 1999 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (Cat. No PR00149)*, volume 1, pages 541–547. IEEE, 1999. 22

[57] Gershon Buchsbaum. A spatial processor model for object colour perception. *Journal of the Franklin institute*, 310(1):1–26, 1980. 23

[58] Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014. 25

[59] Kodak image dataset. http://www.cs.albany.edu/~xypan/research/snr/Kodak.html. 25

[60] Kede Ma, Kai Zeng, and Zhou Wang. Perceptual quality assessment for multi-exposure image fusion. *IEEE Transactions on Image Processing*, 24(11):3345–3356, 2015. 26

[61] Hepeng Jia. Who will win the future of display technologies? *National Science Review*, 5(3):427–431, 2018. 36

[62] Matthew Anderson, Ricardo Motta, Srinivasan Chandrasekar, and Michael Stokes. Proposal for a standard default color space for the internet—srgb. In *Color and imaging conference*, volume 1996, pages 238–245. Society for Imaging Science and Technology, 1996. 36

[63] NY PleasonMlle. Society of motion picture and television engineers. 1959. 36

[64] BS Manjunath, Chandra Shekhar, and Rama Chellappa. A new approach to image feature detection with applications. *Pattern Recognition*, 29(4):627–640, 1996. 45

[65] Barbara Zitova and Jan Flusser. Image registration methods: a survey. *Image and vision computing*, 21(11):977–1000, 2003. 45

[66] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. Deep learning. nature (2015). *May; 521 (7553): 436 10.1038/nature14539*, 2015. 45

[67] Gerard Medioni and Ramakant Nevatia. Matching images using linear features. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, (6):675–685, 1984. 46

[68] Harro Stokman and Theo Gevers. Selection and fusion of color models for image feature detection. *IEEE transactions on pattern analysis and machine intelligence*, 29(3):371–381, 2007. 46

[69] Joost Van de Weijer, Theo Gevers, and Andrew D Bagdanov. Boosting color saliency in image feature detection. *IEEE transactions on pattern analysis and machine intelligence*, 28(1):150–156, 2005. 46

[70] Tony Lindeberg. Feature detection with automatic scale selection. *International journal of computer vision*, 30(2):79–116, 1998. 46

[71] Mario Pavlić, Heidrun Belzner, Gerhard Rigoll, and Slobodan Ilić. Image based fog detection in vehicles. In *2012 IEEE Intelligent Vehicles Symposium*, pages 1132–1137. IEEE, 2012. 46

[72] Sayanan Sivaraman and Mohan Manubhai Trivedi. Looking at vehicles on the road: A survey of vision-based vehicle detection, tracking, and behavior analysis. *IEEE transactions on intelligent transportation systems*, 14(4):1773–1795, 2013. 46

[73] Bahram Jalali and Ata Mahjoubfar. Tailoring wideband signals with a photonic hardware accelerator. *Proceedings of the IEEE*, 103(7):1071–1086, 2015. 46, 49, 72

[74] Christos V Ilioudis, Carmine Clemente, Mohammad H Asghari, Bahram Jalali, and John J Soraghan. Edge detection in sar images using phase stretch transform. In *2nd IET International Conference on Intelligent Signal Processing 2015 (ISP)*, pages 1–5. IET, 2015. 47

[75] Jalalilabucla/image-feature-detection-using-phase-stretch-transform. https://github.com/JalaliLabUCLA/Image-feature-detection-using-Phase-Stretch-Transform/. 47, 76

[76] Govind P Agrawal. Nonlinear fiber optics (academic press, new york, usa), 2007. 49

[77] Ata Mahjoubfar, Dmitry V Churkin, Stéphane Barland, Neil Broderick, Sergei K Turitsyn, and Bahram Jalali. Time stretch and its applications. *Nature Photonics*, 11(6):341, 2017. 49, 77

[78] Zohaib Hameed and Chunyan Wang. Edge detection using histogram equalization and multi-filtering process. In *2011 IEEE International Symposium of Circuits and Systems (ISCAS)*, pages 1077–1080. IEEE, 2011. 57

[79] James J Rankine, Antony N Thomas, and Dorothee Fluechter. Diagnosis of pneumothorax in critically ill adults. *Postgraduate medical journal*, 76(897):399–404, 2000. 59

[80] Daniel R Solli and Bahram Jalali. Analog optical computing. *Nature Photonics*, 9(11):704–706, 2015. 72

[81] Marcello Ferrera, Yongwoo Park, Luca Razzari, Brent E Little, Sai T Chu, Roberto Morandotti, David J Moss, and José Azaña. On-chip cmos-compatible all-optical integrator. *Nature communications*, 1(1):1–5, 2010. 72

[82] Bahram Jalali, Jacky Chan, and Mohammad H Asghari. Time–bandwidth engineering. *Optica*, 1(1):23–31, 2014. 72

[83] Yunshan Jiang, Peter TS DeVore, and Bahram Jalali. Analog optical computing primitives in silicon photonics. *Optics letters*, 41(6):1273–1276, 2016. 72

[84] Yunshan Jiang, Saili Zhao, and Bahram Jalali. Invited article: Optical dynamic range compression. *APL Photonics*, 3(11):110806, 2018. 72

[85] Fangfei Liu, Tao Wang, Li Qiang, Tong Ye, Ziyang Zhang, Min Qiu, and Yikai Su. Compact optical temporal differentiator based on silicon microring resonator. *Optics Express*, 16(20):15880–15886, 2008. 72

[86] Pantelis Velanas, Adonis Bogris, Apostolos Argyris, and Dimitris Syvridis. High-speed all-optical first-and second-order differentiators based on cross-phase modulation in fibers. *Journal of Lightwave Technology*, 26(18):3269–3276, 2008. 72

[87] Madhuri Suthar and Bahram Jalali. Optical computing of mathematical derivatives using dispersion and coherent detection. In *2019 Conference on Lasers and Electro-Optics Europe & European Quantum Electronics Conference (CLEO/Europe-EQEC)*, pages 1–1. IEEE, 2019. 73

[88] Pietro Perona and Jitendra Malik. Scale-space and edge detection using anisotropic diffusion. *IEEE Transactions on pattern analysis and machine intelligence*, 12(7):629–639, 1990. 76

[89] Joachim Weickert. *Anisotropic diffusion in image processing*, volume 1. Teubner Stuttgart, 1998. 76

[90] K Niklas Nordström. Biased anisotropic diffusion: a unified regularization and diffusion approach to edge detection. *Image and vision computing*, 8(4):318–327, 1990. 76

[91] Francine Catté, Pierre-Louis Lions, Jean-Michel Morel, and Tomeu Coll. Image selective smoothing and edge detection by nonlinear diffusion. *SIAM Journal on Numerical analysis*, 29(1):182–193, 1992. 76

[92] Luis Alvarez, Pierre-Louis Lions, and Jean-Michel Morel. Image selective smoothing and edge detection by nonlinear diffusion. ii. *SIAM Journal on numerical analysis*, 29(3):845–866, 1992. 76

[93] Hao Zhao, Ming Lu, Anbang Yao, Yiwen Guo, Yurong Chen, and Li Zhang. Physics inspired optimization on semantic transfer features: An alternative method for room layout estimation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 10–18, 2017. 76

[94] Shan Yang, Zherong Pan, Tanya Amert, Ke Wang, Licheng Yu, Tamara Berg, and Ming C Lin. Physics-inspired garment recovery from a single-view image. *ACM Transactions on Graphics (TOG)*, 37(5):1–14, 2018. 76

[95] Trung Quy Phan, Palaiahnakote Shivakumara, and Chew Lim Tan. Detecting text in the real world. In *Proceedings of the 20th ACM international conference on Multimedia*, pages 765–768, 2012. 76

[96] Vitaliy Fadeyev and Carl Haber. A novel application of high energy physics technology to the problem of audio preservation. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 518(1-2):456–462, 2004. 76

[97] Sifeng He and Bahram Jalali. Brain mri image super resolution using phase stretch transform and transfer learning. *arXiv preprint arXiv:1807.11643*, 2018. 77

[98] Raymond Bing Quan Ang, Humaira Nisar, Muhammad Burhan Khan, and Chi-Yi Tsai. Image segmentation of activated sludge phase contrast images using phase stretch transform. *Microscopy*, 68(2):144–158, 2019. 77

[99] Madhuri Suthar and Bahram Jalali. Phase-stretch adaptive gradient-field extractor (page). In *Coding Theory*. IntechOpen, 2020. 78

[100] Fumihito Yasuma, Tomoo Mitsunaga, Daisuke Iso, and Shree K Nayar. Generalized assorted pixel camera: postcapture control of resolution, dynamic range, and spectrum. *IEEE transactions on image processing*, 19(9):2241–2253, 2010. 79

[101] Mats GL Gustafsson, Lin Shao, Peter M Carlton, CJ Rachel Wang, Inna N Golubovskaya, W Zacheus Cande, David A Agard, and John W Sedat. Three-dimensional resolution doubling in wide-field fluorescence microscopy by structured illumination. *Biophysical journal*, 94(12):4957–4970, 2008. 79

[102] Mats GL Gustafsson. Nonlinear structured-illumination microscopy: wide-field fluorescence imaging with theoretically unlimited resolution. *Proceedings of the National Academy of Sciences*, 102(37):13081–13086, 2005. 79

[103] Stefan Hell and Ernst HK Stelzer. Fundamental improvement of resolution with a 4pi-confocal fluorescence microscope using two-photon excitation. *Optics Communications*, 93(5-6):277–282, 1992. 79

[104] Samuel T Hess, Thanu PK Girirajan, and Michael D Mason. Ultra-high resolution imaging by fluorescence photoactivation localization microscopy. *Biophysical journal*, 91(11):4258–4272, 2006. 79

[105] Mark Bates, Bo Huang, and Xiaowei Zhuang. Super-resolution microscopy by nanoscale localization of photo-switchable fluorescent probes. *Current opinion in chemical biology*, 12(5):505–514, 2008. 79

[106] Alptekin Temizel and Theo Vlachos. Wavelet domain image resolution enhancement using cycle-spinning. *Electronics Letters*, 41(3):119–121, 2005. 79

[107] Yinji Piao, HyunWook Park, et al. Image resolution enhancement using inter-subband correlation in wavelet domain. In *2007 IEEE International Conference on Image Processing*, volume 1, pages I–445. IEEE, 2007. 79

[108] Miguel Granados, Tunç Ozan Aydın, J Rafael Tena, Jean-François Lalonde, and Christian Theobalt. Hdr image noise estimation for denoising tone mapped images. In *Proceedings of the 12th European Conference on Visual Media Production*, pages 1–8, 2015. 83

[109] Stuart Perry. Image and video noise: An industry perspective. In *Denoising of Photographic Images and Video*, pages 207–234. Springer, 2018. 83

[110] Peyman Milanfar. A tour of modern image filtering: New insights and methods, both practical and theoretical. *IEEE signal processing magazine*, 30(1):106–128, 2012. 83

[111] Pierre Michel Duffieux. The fourier transform and its applications to optics. *The Fourier Transform and Its Applications to Optics by PM Duffieux New York*, 1983. 94

[112] Beverley Isherwood, Paul Timpson, Ewan J McGhee, Kurt I Anderson, Marta Canel, Alan Serrels, Valerie G Brunton, and Neil O Carragher. Live cell in vitro and in vivo imaging applications: accelerating drug discovery. *Pharmaceutics*, 3(2):141–170, 2011. 109

[113] Madhuri Suthar, Jonathan Mendelson, Ata Mahjoubfar, Cejo Lonappan, Koushik Roy, Simon Mitchell, Alexander Hoffmann, and Bahram Jalali. Automated hyper-dimensional live cell tracking. *Imaging, Manipulation, and Analysis of Biomolecules, Cells, and Tissues XVI*, (10497-21), 2018. 110

[114] Michael D Rawlins. Cutting the cost of drug development? *Nature reviews Drug discovery*, 3(4):360–364, 2004. 110

[115] Christopher P Adams and Van V Brantner. Estimating the cost of new drug development: is it really dollars 802 million? *Health affairs*, 25(2):420–428, 2006. 110

[116] Steve Morgan, Paul Grootendorst, Joel Lexchin, Colleen Cunningham, and Devon Greyson. The cost of drug development: a systematic review. *Health policy*, 100(1):4–17, 2011. 110

[117] Michael Dickson and Jean Paul Gagnon. The cost of new drug discovery and development. *Discovery medicine*, 4(22):172–179, 2009. 110

[118] Leon Aarons, Mats O Karlsson, France Mentré, Ferdinand Rombout, Jean-Louis Steimer, Achiel van Peer, et al. Role of modelling and simulation in phase i drug development. *European journal of pharmaceutical sciences*, 13(2):115–122, 2001. 110

[119] Xiaobo Zhou and Stephen TC Wong. High content cellular imaging for drug development. *IEEE Signal Processing Magazine*, 23(2):170–174, 2006. 110

[120] Shihuan Kuang, Kazuki Kuroda, Fabien Le Grand, and Michael A Rudnicki. Asymmetric self-renewal and commitment of satellite stem cells in muscle. *Cell*, 129(5):999–1010, 2007. 111

[121] Karima Kissa and Philippe Herbomel. Blood stem cells emerge from aortic endothelium by a novel type of cell transition. *Nature*, 464(7285):112–115, 2010. 111

[122] Ingo Burtscher and Heiko Lickert. Foxa2 regulates polarity and epithelialization in the endoderm germ layer of the mouse embryo. *Development*, 136(6):1029–1038, 2009. 111

[123] Daniel L Coutu and Timm Schroeder. Probing cellular processes by long-term live imaging–historic problems and current solutions. *Journal of cell science*, 126(17):3805–3815, 2013. 111

[124] William C Lemon and Katie McDole. Live-cell imaging in the era of too many microscopes. *Current Opinion in Cell Biology*, 66:34–42, 2020. 111

[125] Jean-Yves Tinevez, Joe Dragavon, Lamya Baba-Aissa, Pascal Roux, Emmanuelle Perret, Astrid Canivet, Vincent Galy, and Spencer Shorte. A quantitative method for measuring phototoxicity of a live cell imaging microscope. *Methods in enzymology*, 506:291–309, 2012. 111

[126] Ira Mellman, George Coukos, and Glenn Dranoff. Cancer immunotherapy comes of age. *Nature*, 480(7378):480–489, 2011. 116

[127] Jennifer Couzin-Frankel. Cancer immunotherapy, 2013. 116

[128] Sofia Farkona, Eleftherios P Diamandis, and Ivan M Blasutig. Cancer immunotherapy: the beginning of the end of cancer? *BMC medicine*, 14(1):1–18, 2016. 116

[129] Niels Halama. The next age of immunotherapy: optimisation, stratification and therapeutic synergies, 2019. 116

[130] Adam J Schoenfeld and Matthew D Hellmann. Acquired resistance to immune checkpoint inhibitors. *Cancer Cell*, 37(4):443–455, 2020. 116

[131] Russell W Jenkins, David A Barbie, and Keith T Flaherty. Mechanisms of resistance to immune checkpoint inhibitors. *British journal of cancer*, 118(1):9–16, 2018. 116

[132] Mathieu Spaas and Yolande Lievens. Is the combination of immunotherapy and radiotherapy in non-small cell lung cancer a feasible and effective approach? *Frontiers in medicine*, 6:244, 2019. 116

[133] Cheng Lu, Can Koyuncu, German Corredor, Prateek Prasanna, Patrick Leo, XiangXue Wang, Andrew Janowczyk, Kaustav Bera, James Lewis Jr, Vamsidhar Velcheti, et al. Feature-driven local cell graph (flock): New computational pathology-based descriptors for prognosis of lung cancer and hpv status of oropharyngeal cancers. *Medical Image Analysis*, 68:101903, 2021. 116

[134] Madhuri Suthar, Zhouya Bai, Scott Nowicki, Antoni Ribas, Cejo Lonappan, and Bahram Jalali. A i curated digital pathology for cancer immunotherapy. In *2019 SPIE Medical Imaging Conference*, pages 1–1. SPIE, 2019. 116, 119

[135] Paul C Tumeh, Christina L Harview, Jennifer H Yearley, I Peter Shintaku, Emma JM Taylor, Lidia Robert, Bartosz Chmielowski, Marko Spasic, Gina Henry, Voicu Ciobanu, et al. Pd-1 blockade induces responses by inhibiting adaptive immune resistance. *Nature*, 515(7528):568–571, 2014. 117, 120

[136] Theodore S Nowicki, Siwen Hu-Lieskovan, and Antoni Ribas. Mechanisms of resistance to pd-1 and pd-l1 blockade. *Cancer journal (Sudbury, Mass.)*, 24(1):47, 2018. 117

[137] Theodore S Nowicki, Ryan Akiyama, Rong Rong Huang, I Peter Shintaku, Xiaoyan Wang, Paul C Tumeh, Arun Singh, Bartosz Chmielowski, Christopher Denny, Noah Federman, et al. Infiltration of cd8 t cells and expression of pd-1 and pd-l1 in synovial sarcoma. *Cancer immunology research*, 5(2):118–126, 2017. 117

[138] Josh Moore, Melissa Linkert, Colin Blackburn, Mark Carroll, Richard K Ferguson, Helen Flynn, Kenneth Gillen, Roger Leigh, Simon Li, Dominik Lindner, et al. Omero and bio-formats 5: flexible access to large bioimaging datasets at scale. In *Medical Imaging 2015: Image Processing*, volume 9413, page 941307. International Society for Optics and Photonics, 2015. 120

[139] Si Zhang, Hanghang Tong, Jiejun Xu, and Ross Maciejewski. Graph convolutional networks: a comprehensive review. *Computational Social Networks*, 6(1):1–23, 2019. 124

[140] Thomas N Kipf and Max Welling. Semi-supervised classification with graph convolutional networks. *arXiv preprint arXiv:1609.02907*, 2016. 124

[141] John Brian Pendry. Negative refraction makes a perfect lens. *Physical review letters*, 85(18):3966, 2000. 128

[142] Claire M Watts, David Shrekenhamer, John Montoya, Guy Lipworth, John Hunt, Timothy Sleasman, Sanjay Krishna, David R Smith, and Willie J Padilla. Terahertz compressive imaging with metamaterial spatial light modulators. *Nature Photonics*, 8(8):605–609, 2014. 128

[143] Alan Zhan, Shane Colburn, Christopher M Dodson, and Arka Majumdar. Metasurface freeform nanophotonics. *Scientific reports*, 7(1):1–9, 2017. 128, 129

[144] Shane Colburn, Alan Zhan, and Arka Majumdar. Metasurface optics for full-color computational imaging. *Science advances*, 4(2):eaar2114, 2018. 129

[145] BL Heffner. Accurate, automated measurement of differential group delay dispersion and principal state variation using jones matrix eigenanalysis. *IEEE photonics technology letters*, 5(7):814–817, 1993. 129