RICE UNIVERSITY

Compressive Hyperspectral Structured Illumination and Classification via Neural Networks

By

Yibo Xu

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE

Master of Science

APPROVED, THESIS COMMITTEE

Kevin F. Kelly, Chair
Associate Professor of Electrical and Computer Engineering

Jacob T. Robinson
Assistant Professor of Electrical and Computer Engineering

Tomasz S. Tkaczyk
Associate Professor of Bioengineering

HOUSTON, TEXAS
April 2016
ABSTRACT

Compressive Hyperspectral Structured Illumination and Classification via Neural Networks

By

We demonstrate two complementary applications based on compressive imaging: hyperspectral compressive structured illumination for three-dimensional imaging and compressive classification of objects using neural networks. The structured light method usually uses structured patterns generated from a commercial digital projector which contain very limited spectral content, using white light or RGB-based giving very little material content and not exploiting possible wavelength-dependent scattering. Therefore we designed and implemented a hyperspectral projector system that is able to generate structured patterns consisting of arbitrarily defined spectrum instead. We used the system to recover the unique spectrum-dependent 3-D volume density of the colored targets of participating media. For the image classification problem, it is known that a set of images of a fixed scene under varying articulation parameters forms a low-dimensional, nonlinear manifold that random projections can stably embed using far fewer measurements. Thus random projections in compressive sampling can be regarded as a dimension-reducing process. We demonstrate a method using compressive measurements of images to train a neural network that has a relatively simple architecture for object classification. As a proof of concept, simulations were performed on infrared vehicle images that demonstrated the utility of this approach.
over previous compressive matched filtering. The success of both these projects bodes well for their overall integration into a single infrared compressive hyperspectral machine-vision instrument.
Acknowledgements

I would like to thank my advisor Professor Kevin F. Kelly of Electrical and Computer Engineering Department at Rice University for his guidance and support throughout my study and research for this thesis in his group. Professor Kelly has been kind and patient in helping me with any difficulty I have encountered during research. He has provided me with not only professional academic guidance but also freedom to explore the research area. Without his help I couldn’t have finished this thesis. I have learned a lot and grown a lot in his group. I would also like to thank my fellow students Liyang Lu and Jianbo Chen in our group for helping me while I am in the group, especially Liyang Lu who provided inspiring suggestions on the optical design of the hyperspectral projector system. Last but not least I want to thank my parents whom I love so much. They have been giving me support, encouragement, guidance, and care for as long as I can remember.
# Contents

Acknowledgements ......................................................................................... iii

Contents ........................................................................................................ iv

List of Figures .................................................................................................. vi

1. **Introduction** ............................................................................................ 10
   1.1. Structured Light .................................................................................... 10
   1.2. Compressive Structured Light ............................................................. 11
   1.3. Compressive Sensing Classification using a Neural Network ............... 13
   1.4. Thesis Outline ..................................................................................... 15

2. **Compressive Imaging** .............................................................................. 16
   2.1. Sampling and Nyquist Rate ................................................................ 16
   2.2. Compressive Sensing .......................................................................... 18
      2.2.1. CS measurements ......................................................................... 19
      2.2.2. CS reconstruction ......................................................................... 20
   2.3. Single-Pixel Camera ........................................................................... 21

3. **Hyperspectral Projector System** ............................................................. 24
   3.1. Hyperspectral Project System Design ................................................. 25
      3.1.1. Optical Design ............................................................................. 27
      3.1.2. Spectral Modulation .................................................................... 30
      3.1.3. DMD Control ............................................................................. 34
   3.2. Compressive Structured Light for Recovering Volume Density of Participating Medium ................................................................. 34
      3.2.1. Image Formation Model ................................................................ 35
      3.2.2. Coding and Formulation ............................................................... 39
      3.2.3. Measurement Data Reconstruction .............................................. 41

4. **Hyperspectral Compressive Structured Light** ......................................... 44
   4.1. Black and White Compressive Structured Light .................................. 45
      4.1.1. Experiment Design ........................................................................ 45
      4.1.2. Reconstruction Results ................................................................. 48
4.2. Hyperspectral Compressive Structured light ............................................. 54
  4.2.1. Experiment Design.................................................................................. 55
  4.2.2. Hyperspectral 3-D Reconstruction ......................................................... 57

5.  Compressive Sensing Classification using a Neural Network ...................... 67
  5.1. Compressive Classification ....................................................................... 67
  5.2. Neural Network Architecture .................................................................. 68
  5.3. Results .................................................................................................... 69
    5.3.1. Classification on Video Chips ............................................................... 69
    5.3.2. Classification on Video Patches ............................................................ 73
    5.3.3. Classification under Noise ................................................................. 76

6.  Conclusion and Future Work ........................................................................ 82

References ...................................................................................................... 85
List of Figures

Figure 1 Operation principle of the SPC. Each measurement is the inner product between the binary mirror orientation patterns on the DMD and the scene to be acquired. .................................................................22

Figure 2 Schematic layout of the hyperspectral projector (top view). ...........26

Figure 3 Illustration of two point light sources $a$ and $b$ along the slit being focused at different $x$ positions on the DMD. $a'$ and $b'$ are the dispersed spectral lines spanning the $y$ direction formed from $a$ and $b$, respectively (side view of the hyperspectral projector). .................................................................28

Figure 4 Illustration of the spectrum focused on the surface of the DMD ........29

Figure 5 (a) DMD Diamond Pixel Geometry. (b) DMD Diamond Pixel Array Configuration [37]. ........................................................................................................31

Figure 6 Spectral modulation. (a) Illustration of an example DMD pattern. Mirrors in the white area are on and in the black area are off. (b) Spectrum on the DMD surface. (c) Spectrum on the white area where the mirrors are on is selected. (d) Image of the projected hyperspectral stripes on the screen when DMD displays the pattern in (a). ........................................................................................................32

Figure 7 The spectrum measured by a spectrometer of the top stripe which is white and the bottom stripe which composes of eight spectral bands. ........33

Figure 8 Example hyperspectral stripes projected on a toy car. .................33

Figure 9 (a) Compressive structured light for recovering participating media. Coded light is emitted along the $z$-axis to the volume while the camera acquires images as line-integrated measurements of the volume density along the $x$-axis. Volume density is reconstructed from the acquired measurements by using compressive sensing techniques [32]. (b) Image formation model for participating medium under single scattering. The image intensity at one pixel, $I(x, y, z)$, depends on the integral along the $x$-axis of the projector’s radiance, $L(x, y)$, and the medium density, $\rho(x, y, z)$, along a ray through the camera center [2]. ........................................................................................................36

Figure 10 Temporal coding of the volume using compressive structured light ........................................................................................................39
Figure 11 Reconstruction results of two planes. (a) A photograph of the object consisting of two glass slabs with powder. The letters “EC” are on the back slab and “CV” on the front slab. (b) One of the images captured by the camera. (c) Reconstructed volume at different views without attenuation correction [2].

Figure 12 Experimental setup of compressive structured light using the proposed hyperspectral projector system.

Figure 13 (a) Target used for the experiment. The letter “C” is carved manually on each of the front and back planes by removing the plane material. The “C” on different planes curls in opposite directions. (b) Example images of the coded volume captured by the camera.

Figure 14 Reconstruction results of the 3-D volume density of the target of the two planes at resolution of $32 \times 32 \times 32$ using 24 compressive measurements. (a) 3-D views of the reconstruction from two perspectives. (b)(c)(d) Example 2D slices of the reconstructed 3-D volume density in y-x, z-x, x-y views, respectively. The number on the corner of each image is coordinate index of the image in the dimension of slicing. The two planes are distinctive in the 2D slices and locations of the “C” appear as holes in the two planes. The plane with higher intensity is the front plane.

Figure 15 Reconstruction results of the 3-D volume density of the target of the two planes at resolution of $128 \times 128 \times 128$ using 64 compressive measurements. (a) 3-D views of the reconstruction from two perspectives. (b)(c)(d) Example 2D slices of the reconstructed 3-D volume density in y-x, z-x, x-y views, respectively. The number on the corner of each image is coordinate index of the image in the dimension of slicing. The two planes are distinctive in the 2D slices and locations of the “C” appear as holes in the two planes. The plane with higher intensity is the front plane.

Figure 16 The target and its spectrum. (a) Photo of the target for reconstruction which contains two objects placed close together: one object comprises of two red translucent planes with letter “C” carved on each of the front and back planes, the other consists of two cyan translucent planes with letter “V” carved on each of the front and back planes. (b) Image of the target taken from the perspective of the camera using in the experiment under white illumination. (c) Reflectance spectra of the red and cyan planes. Red has strong reflectance between 590 nm and 750 nm, while cyan is strongly reflective between 390 nm and 590 nm.
Figure 17 (a) Image of the camera of the target under an example structured light pattern of wavelength longer than 610 nm, where the red object is encoded and the cyan object is invisible. (b) Spectrum of the first set of structured patterns. (c) Image of the camera of the target under an example structured light pattern of wavelength shorter than 570 nm, where the cyan object is encoded and the red object is invisible. (d) Spectrum of the second set of structured pattern.

Figure 18 Reconstruction results of the 3-D volume density of the red object of $32 \times 32 \times 32$ using 24 compressive measurements. (a) 3-D views of the reconstruction from two perspectives. (b)(c)(d): Example 2D slices of the reconstructed 3-D volume density in y-z, x-y, x-z views, respectively. The number on the upper right corner of (b) (d) and lower corner of (c) of each image is coordinate index of the image in the dimension of slicing. The two planes are distinctive in the 2D slices and locations of the “C” appear as holes in the two planes. The plane with higher intensity is the front plane.

Figure 19 Reconstruction results of the 3-D volume density of the red object of $32 \times 32 \times 32$ using 24 compressive measurements. (a) 3-D views of the reconstruction from two perspectives. (b)(c)(d): Example 2D slices of the reconstructed 3-D volume density in y-z, x-y, x-z views, respectively. The number on the upper right corner of each image is coordinate index of the image in the dimension of slicing. The two planes are distinctive in the 2D slices and locations of the “V” appear as holes in the two planes. The plane with higher intensity is the front plane.

Figure 20 Neural Network Architecture

Figure 21 Example chips for each class of vehicles used for training and testing. The resolution of the chips is 64*64.

Figure 22 Confusion matrices and neural network architectures of test results. All the classification results achieve an excellent error rate of zero percent.

Figure 23 Example video patches for the three classes.

Figure 24 Confusion matrices and neural network architectures of test results. All the classification results achieve an excellent error rate of zero percent.

Figure 25 Synthesized images of the three classes of vehicles.
Figure 26 First row: an image before and after adding Gaussian noise of 10 dB. Second row: an image before and after adding Gaussian noise of 20 dB. ....... 77

Figure 27 Confusion matrices and neural network architectures of test results. The result shows that the neural network is robust to noise in the test image data. 80
1.1. Structured Light

Structured light is considered one of the most reliable techniques for recovering the 3-D shape of objects. A variety of applications of 3-D shape measurement include control for intelligent robots, obstacle detection for vehicle guidance, dimension measurement for die development, stamping panel geometry checking, and accurate stress/strain and vibration measurement. Moreover, automatic on line inspection and recognition issues can be converted to the 3-D shape measurement of an object under inspection, for example, body panel paint defect and dent inspection [1]. Conventional structured light methods project coded light patterns onto the surface of an opaque object and observe it using a camera so the correspondences between image points and points of the projected pattern can be established and the 3-D structure of the scene can be recovered by triangulation.
Over the years, researchers have developed various types of coding strategies, such as binary codes, phase shifting, spatial neighborhood coding, etc. However, many real-world phenomena can only be described by volume densities rather than boundary surfaces. Such phenomena are often referred to as participating media [2]. Examples include translucent objects, smoke, clouds, mixing fluids, and biological tissues. It is an intriguing and fast-growing area to develop methods that recover the 3-D volume densities of these dynamic phenomena.

Many solutions have been proposed to address the problem of recovering the volume density of a participating medium. Hawkins et al. [3] used a high-powered laser sheet and a high-speed camera (5,000 fps) to measure thin slices of a smoke density field via scanning. Fuchs et al. [4] proposed the idea of shooting a set of static laser rays into the volume and using spatial interpolation to reconstruct the volume. However, both methods are straightforward sequential scanning of a volume and, in this case, the measurements are inherently sparse and hence the recovered information is low in resolution.

1.2. Compressive Structured Light

Compressive sensing (CS) [5-11] is a new concept in signal processing where one seeks to minimize the number of measurements to be taken from signals while still retaining the information necessary to approximate them well. CS puts forward
a paradigm that surpasses the traditional Nyquist rate for sampling and has since been used successfully in applications as discussed in Chapter 2. I propose two applications based on compressive sensing theory in this thesis.

Gu J, Nayar S K, et al. [2] proposed a more efficient method, named compressive structured light, for recovering participating medium which combines structured light method and compressive sensing theory. This method projects patterns into a volume of participating medium to produce images which are integral measurements of the volume density along the line of sight. The compressive structured light method makes the measurement of a participating medium highly efficient in terms of acquisition time as well as illumination power.

A drawback in all structured light methods, including the compressive structured light technique, is that a commercialized digital projector is used to project coded structured light patterns on the scene. These projectors usually contain as their light source the red, green and blue LEDs which have very narrow emission spectrums around their peak emission wavelength, or else a broad-spectra lamp and a spinning color filter wheel. Because of this, the projected patterns on the scene can be regarded as containing limited spectral content in both cases. On the other hand, the atoms and molecules, upon which our world is built, possess very complex spectral responses as a part of their innate characteristics, e.g. emission, absorption and scattering properties that are wavelength dependent. This spectrally
dependent information imbedded in all materials, if well employed, is able to reveal and reflect deeper and more meaningful nature of a wide variety of materials and phenomena of scientific interest. Therefore, if the coded light patterns consisted of arbitrarily desired spectrum instead of the single wavelengths or wavebands, spectrum-dependent information of the phenomenon could be revealed in addition to volume density distribution. In Chapter 3, I propose a novel hyperspectral projector system based on a single digital micromirror device (DMD) that exactly meets such a demand and demonstrate its utility to perform hyperspectral compressive structured light for recovering 3-D volume density.

1.3. Compressive Sensing Classification using a Neural Network

Vehicle classification is of great importance in a wide variety of real-world applications such as motorway surveillance for monitoring traffic conditions, reducing congestion and enhancing mobility, fare collection, toll collection, booth gate operator, break-down roadside services, traffic offence detection and so on [12]. The convolutional neural network [13] has been shown to be a powerful tool for doing image classification with very large dataset, but its model complexity not only incurs the need of a large amount of computational power due to the immense size of the network during training, but also leads to overfitting issues when used for tasks with very limited training data. Compressive sensing produces a condensed representation of the image, which give promise to do image
classification via simpler neural networks instead of convolutional neural networks. Studies have shown that image transforms such as the Discrete Cosine Transform (DCT) can be used for reducing redundant information in images and the compressed DCT coefficients can be effectively used for image classification through multilayer perceptron [14], [15]. To my best knowledge, there has been no research on using compressive sensing coefficients for image classification through neural network. In addition to enabling sub-Nyquist measurement, CS enjoys a number of attractive properties [16]. CS measurements are universal in that the same random matrix works simultaneously for exponentially many sparsifying bases with high probability; no knowledge is required of the nuances of the data being acquired. Whereas with DCT the compression process is image-dependent in that the complete set of DCT coefficients needs to be computed first and sorted and then smaller coefficients are dropped keeping only the large coefficients. Moreover DCT requires the same number of measurements as the number of pixels in the image while compressive sensing requires much smaller number of measurements. Due to the incoherent nature of the measurements, CS is robust in that the measurements have equal priority, unlike the DCT, Fourier or wavelet coefficients in a transform coder. In Chapter 4, I propose and implement a two layer, feed-forward neural network architecture and use it to do vehicle classification with compressive samplings of shortwave-infrared (SWIR) images of three types of vehicles. This method gives promise to building a single-pixel camera that can do vehicle detection and classification in SWIR without reconstructing the original image.
1.4. Thesis Outline

In Chapter 2, I will review the theory of compressive sensing and introduce the single-pixel camera, a unique hardware implementation of compressive imaging system with a single-element photon detector. In Chapter 3, the proposed hyperspectral projector system will be described in detail. Then a series of experiments will be presented on using this system to perform hyperspectral compressive structured light for recovering 3-D volume density of a static translucent object. In Chapter 4, I will present a two layer, feed-forward neural network architecture and use it to classify short wave infrared (SWIR) vehicle images with compressive measurements. In Chapter 5, I will give a summary and discuss future directions.
Chapter 2

2. Compressive Imaging

2.1. Sampling and Nyquist Rate

In the modern world, nearly all data begins as an analog signal. But in order to manipulate and analyze such data, it needs to be converted to the digital domain, so that the microprocessor will be able to read, understand, store and manipulate the data. Sampling is the reduction of a continuous analog signal to a discrete digital signal. Sampling can be represented mathematically such that given a continuous signal \( s(t) \) to be sampled and the sampling interval \( T \), the sampled version of \( s \) is given by the sequence:

\[
s_k = s(kT)
\]  \hspace{1cm} (1)
where $k$ is an integer. We notice that the information between samples that originally existed in the continuous analog signal is lost in the digital sampling process. According to the Shannon-Nyquist sampling theorem, for a band-limited signal, the sampling rate $1/T$ needs to be at least twice of the signal bandwidth of interest in order to avoid any loss of relevant information for the original signal after sampling. This principle generally underlies all signal acquisition techniques, such as consumer electronics, medical imaging, and so on.

However, making such measurements is expensive. In many applications, the Nyquist rate may be so high that it poses great challenges in data acquisition, storage, transmission and processing in spite of the tremendous progress in storage capability and computing power. Examples are provided by virtually any domain of science or technology where amounts of data are very large and costs of measurement are nontrivial. As such, the conventional Shannon-Nyquist sampling method is not sufficient to address the dilemma caused between the limited resources and the level of detail one would like to capture.
2.2. Compressive Sensing

Compressive sensing, (CS) [5-11], also known as compressive sampling or compressed sensing, is a relatively recent concept in signal processing where one seeks to minimize the number of measurements to be taken from signals while still retaining the information necessary to produce a nearly complete recovery. The compressive sensing theory beats the Nyquist limit by showing that it is possible to reconstruct sparse or compressible signals almost exactly from a number of nonadaptive linear measurements which is far smaller than required by the Shannon-Nyquist theorem. Compressive sensing puts forward a novel sampling paradigm that replaces the notion of band-limited signals with that of sub-sampling sparse or compressible signals and recovery by optimization instead of by invertible transform.

An N × 1 vector is called K-sparse if only K of its transformation coefficients under a certain basis are nonzero where K≪N. An N × 1 vector is called compressible if only K of its transformation coefficients under a certain basis are significantly non-zero where K≪N and can be well-approximately with those K large coefficients. Images of natural scenes are usually compressible under various transformations, e.g. Wavelet transform, Discrete Cosine Transform (DCT) and Fourier transform. Thus the compressive sensing framework can be well applied to their acquisition and recovery.
2.2.1. CS measurements

Suppose $x$ is an unknown vector in $\mathbb{R}^N$ (a digital image or signal) which is sparse or compressible. In compressive sensing, we plan to sample $x$ using $M$ nonadaptive linear measurements of $x$ and then reconstruct. We are interested in the case $M \ll N$, when we have many fewer measurements than the dimension of the signal space. Every measurement encodes the signal vector $x$ by projecting it onto one of a series of specially designed measurement vectors $\{ \varphi_k \}$, for $k=1,\ldots,M$, producing the measurement value $y_k = \langle x, \varphi_k \rangle$. Then the original signal vector is reconstructed from these measurement data using certain reconstruction algorithm. The process can be mathematically expressed as:

$$y = \Phi x = \Phi \Psi \alpha \quad (2)$$

where $x$ is the $N\times1$ signal vector, $\Phi$ is the $M\times N$ measurement matrix with each row being a measurement vector $\varphi_k$, thus having a total of $M$ measurement vectors where $M \ll N$, and $y$ is the $M\times1$ measurement data vector. $\Psi$ is the $N\times N$ matrix representing the transformation basis under which the signal $x$ is sparse, e.g. wavelet basis or DCT basis, with each column of $\Psi$ being a basis vector of the transformation. $\alpha$ is the $N\times1$ vector, representing the transformation coefficients of the signal $x$ under the transformation $\Psi$. While the design of $\Phi$ is beyond the scope of this thesis, an intriguing choice that works with high probability is a random
matrix. For example, we can draw the elements of Φ as i.i.d. ±1 random variables from a uniform Bernoulli distribution.

2.2.2. CS reconstruction

The measurement scheme in equation (1) leads us to arrive at an underdetermined system of linear equations, which, as is well known, in general to be infinitely many possible solutions, commonly referred to as ill-posed. Also the transformation from x to y is a dimensionality reduction and so necessarily loses information. The magic of CS is that Φ can be designed such that x can be recovered exactly (in the case of true sparse) or approximately (in the case of compressible) from the measurement y, that is, if x depends only on a small number of degrees of freedom, thus α has only K≪N non-zero elements for a sparse signal, or K≪N significantly non-zero elements for a compressible signal.

To recover the image x from the random measurement y, the traditional favorite method of least squares can be shown to fail with high probability. Instead, it has been shown that using the $l_1$ optimization [5], [10], [17]:

$$\hat{\alpha} = \text{arg min } \|\alpha\|_1 \quad \text{such that } \|y - \Phi\Psi\alpha\|_2 < \epsilon$$  \hspace{1cm} (3)
we can closely approximate \( K \)-sparse vectors and compressible vectors stably with high probability using just \( M \geq O(K \log(N/K)) \) random measurements. In real world experiments, the measurement \( y \) is usually corrupted by noise and \( \epsilon \) is an upper bound on the noise magnitude. This optimization can be solved using standard convex programming algorithms.

In the field of CS image reconstruction, total variation (TV) regularization is another well-known method for its ability to recover the edges or boundaries more accurately than \( l_1 \) method. TV minimization suggests that the gradient of the 2D image signal is sparse, so it can be considered as a generalized \( l_1 \) minimization problem on the image gradient map. It can be expressed as [18]:

\[
\hat{x} = \arg\min \sum_i \|D_i x\| \quad \text{such that} \quad \|y - \Phi x\|_2 < \epsilon
\]  

(4)

where \( \|D_i x\| \) is the discrete gradient magnitude at pixel \( i \) of the image \( x \).

### 2.3. Single-Pixel Camera

Compressive sensing has a variety of successful applications including optical imaging [16], [19], medical visualization [20], and radar [21]. Recently, compressive sensing has also been widely used to solve many computer vision and computer graphics problems, such as high-speed imaging [22], [23], [24], image restoration and denoising [25], [26], [27] and light transport measurement [28], [29].
Our group at Rice University previously developed a unique imaging hardware platform, named the single-pixel camera (SPC) [16], which incorporates a spatial light modulator and a single detector, as shown in Figure 1. Our group has exploited SPC to construct infrared [19], hyperspectral [30], [31] and low-light imaging systems that have greatly reduced cost in power, space, and expense compared to their traditional counterparts.

![Figure 1 Operation principle of the SPC. Each measurement is the inner product between the binary mirror orientation patterns on the DMD and the scene to be acquired.](image)

In the SPC, a 2D image serves as the original sparse signal $x$, which can be regarded as the $N$ pixels of the 2D image stretched into an $N \times 1$ vector. To encode the signal, the DMD is programmed to displays a sequence of
measurement vectors consisting of binary elements {0, 1} reshaped into a 2D configuration to modulate the intensities of image pixels. When the 2D image is projected onto the DMD, the reflected lights from pixels that are encoded by +1 come out from the DMD in one direction and those encoded by 0 come in an opposing direction. Then lenses are used to sum up the lights encoded by +1 and the final resulting intensity is detected by a single detector as measurement data. Typically the SPC employs pseudo-random Hadamard matrices as measurement vectors on the DMD because randomized measurement basis are generally incoherent with the sparse representation basis and that a DMD can be programmed to display any sequence of patterns including random ones.
Chapter 3

3. Hyperspectral Projector System

In this Chapter, the proposed hyperspectral projector system is described in detail. The hyperspectral projector features a simple and low-cost design based on a single DMD. It is able to generate coded light patterns consisting of arbitrarily desired spectrum of single/multiple wavelength/wavebands, and, when combined with a Dove prism to rotate the stripes, is sufficient to produce the necessary structured patterns for most structured light applications. This hyperspectral projector system could be very useful in applications such as calibration and testing of hyperspectral imagers, 3-D recovery for machine visions and multicolor bio-imaging. Then the compressive structured light method proposed by Gu J, Nayar S K, et al. [2] is explained in detail. As a proof of principle, the hyperspectral projector system is used to perform hyperspectral compressive structured light for recovering 3-D volume density of static translucent objects as a function of color, and this experiment is explained in Chapter 4.
3.1. Hyperspectral Project System Design

This section details the design of the DMD-based hyperspectral projector system. This projector gives complete independence of one spatial and one spectral dimension and when combined with a rotating Dove prism achieves programmable control in all three dimensions. It is realized by exploiting the DMD to serve as a light modulator in the spectral domain, in contrast to the SPC where the DMD performs light modulation in the spatial domain. As shown in Figure 2, a diffraction grating disperses light into a spectrum on the DMD and the DMD modulates the intensities of the spectral lines to keep the desired portion of the spectrum and leave out the rest. Then the selected spectrum is recombined by the same diffraction grating. In addition to these two key components, an achromatic lens is used to focus and collimate the dispersed spectrum. A Dove prism is used to rotate the projected images if needed. A cylindrical lens is then used to stretch the modulated light in one dimension to generate stripe patterns. Details of the optical design and the spectral modulation by DMD are described in following sections.
While this is not the first DMD-based spectral illumination system developed, this new design has distinct advantages over previous work. One of the most complete systems built is NIST's Hyperspectral Image Projector [36]. However the two drawbacks of this system is that it requires two DMDs to separate the unique spectra across both x and y dimensions and it acquires a very intense light source or very sensitive imagers to make up for the optical losses in the system. Meanwhile, the proposed hyperspectral projector here uses a single DMD to produce hyperspectral stripes and, when combined with a Dove prism to rotate the

Figure 2 Schematic layout of the hyperspectral projector (top view).
stripes, is sufficient to produce the necessary structured patterns for most structured light applications. The projector design exploits the light source efficiently in that it does not have any optical loss except for the reflection/absorption loss of light caused its optical elements, e.g. lenses, mirrors, and the loss can only be reduced by upgrading these hardware to have optimized properties.

3.1.1. Optical Design

Figure 2 shows the optical design of the hyperspectral projector. Light coming out of a halogen lamp is guided through an optical fiber and focused on an adjustable vertical slit in the $x$-direction. The slit can be regarded as a line of point light sources. Each point light source is collimated into a parallel light beam by the convex lens 1. The light beams travel into a transmission diffraction grating (Thorlabs, Visible Transmission Grating, 300 Grooves/mm). The grooves on the grating are in $x$ direction. Light is dispersed into its spectral components after the grating which travel in different wavelength-dependent angles. The grating is designed such that most of the incoming light power is concentrated in one of the two symmetric directions of its first order diffracted light, minimizing the light loss in zero order and higher order diffraction. The first order diffracted light then goes into an achromatic lens which focuses the different spectral components onto different $y$ positions on the surface of the DMD. The distance between the grating and the achromatic lens and the distance between the achromatic lens and the DMD
are equal to the focal length $f$ of the achromatic lens. Then the DMD performs spectral modulation, keeping the desired part of the spectrum and abandoning the rest. Details of modulation are described in next section. The spectrum to be kept is reflected by the micro-mirrors back into the achromatic lens, recombines into the diffraction grating, focused by lens 2 and forms the image of a line. Due to the symmetric configuration of the grating, the achromatic lens and the DMD, the image formed by lens 2 is in fact the image of the slit light source, except that it only has a portion of the original spectrum of the slit. Then a cylindrical lens stretches the thin line into a stripe on a screen to be displayed or onto an object for scanning. A Dove prism can be placed between lens 2 and the cylindrical lens to enable rotation of the stripes in all angles, allowing two-dimensional hyperspectral illumination.

Figure 3 Illustration of two point light sources $a$ and $b$ along the slit being focused at different $x$ positions on the DMD. $a'$ and $b'$ are the dispersed
spectral lines spanning the $y$ direction formed from $a$ and $b$, respectively (side view of the hyperspectral projector).

Because a slit light source is used, every spectral component forms a line on the DMD. To demonstrate this, as shown in Figure 3, consider two point light sources $a$ and $b$ along the slit, $a$ forms a spectral line $a'$ spanning in $y$ direction on the DMD, and similarly $b$ forms a spectral line $b'$. Yet $a'$ and $b'$ are focused in different $x$ positions, and likewise for all points along the slit. Therefore on the surface of the DMD, every line in $x$ direction is of the same wavelength formed from all points along the slit, and every line in the $y$ direction is the dispersed spectral line formed from one point on the slit. Figure 4 illustrates the spectrum distribution on the surface of the DMD.

![Figure 4 Illustration of the spectrum focused on the surface of the DMD](image)
3.1.2. Spectral Modulation

To realize spectral modulation, a DMD chip (Texas Instrument DLP LightCrafter 4500) is incorporated at the focal plane of the achromatic lens and orthogonal to the optical axis of the system. The functional part of the DMD is a 912x1140 interlaced array of electrostatically controlled micro-mirrors of size 7.6 × 7.6 μm each (Figure 5 (a)). Every micro-mirror can be independently actuated by an individual SRAM cell, and rotate about a hinge to be at one of two states, +12° (tilting right) and -12° (tilting left) with respect to the DMD surface. In this DMD chip, the micro-mirrors are interlaced in a diamond pixel geometry as demonstrated in Figure 5 (b), so the hinges are all in x direction. The system is designed such that all the micro-mirrors oriented at +12° reflect the spectrum on themselves back into the achromatic lens and finally reach the screen, and the spectrum on the micro-mirrors oriented at −12° does not reach the achromatic lens and gets lost in the space (Figure 2). I will denote the mirror state of +12° as mirror being ON and −12° as mirror being OFF. Therefore spectral modulation is achieved by programming each of the micro-mirrors to be on/off to keep/discard the light focused this micro-mirror.
On the DMD, if a line in $y$-direction of micro-mirrors are turned on, the light focused on this line of mirrors will form the image of a white, thin stripe on the screen. Therefore the spatial resolution of stripes of the projector is up to the number of micro-mirrors on the DMD along $x$-direction. If some of the mirrors on this line are off, the spectrum content focused on these mirrors will be discarded, and the image on the screen will be a thin stripe with specific wavelengths. Therefore the spectral resolution of the projector is up to the bandwidth of spectrum divided by the number of micro-mirrors on the DMD along $x$-direction. In applications where smaller spatial resolution is sufficient, neighboring stripes can be combined to form wider stripes. As demonstrated in Figure 6, the DMD displays a pattern with five stripes (Figure 6 (a)) and the light focused on the white area where the mirrors are on is selected (Figure 6 (c)). The selected light, after recombined by the grating and stretched by the cylindrical lens, forms five hyperspectral stripes on the screen (Figure 6 (d)). Each of the hyperspectral stripes has the spectral content selected by corresponding stripe pattern in Figure 6 (c). Figure 7 shows the spectrum measured by a spectrometer for the top stripe and the bottom stripe. Note that the top stripe is white because the full spectrum is selected as in Figure 6 (c), and the

Figure 5 (a) DMD Diamond Pixel Geometry. (b) DMD Diamond Pixel Array Configuration [37].
bottom stripe composes of eight spectral bands because the eight bands are selected.

Figure 8 displays some example hyperspectral stripes projected on a toy car.

Figure 6 Spectral modulation. (a) Illustration of an example DMD pattern. Mirrors in the white area are on and in the black area are off. (b) Spectrum on the DMD surface. (c) Spectrum on the white area where the mirrors are on is selected. (d) Image of the projected hyperspectral stripes on the screen when DMD displays the pattern in (a).
Figure 7 The spectrum measured by a spectrometer of the top stripe which is white and the bottom stripe which composes of eight spectral bands.

Figure 8 Example hyperspectral stripes projected on a toy car.
3.1.3. DMD Control

The control of the DMD can be achieved in two approaches. One is to use the control software GUI of DLP LightCrafter 4500 that preloads a set of patterns into the memory on the DMD chip board. But because the memory size is not large enough, the DMD can only continuously display a very limited number of patterns before stopping and manually reloading the next set of patterns into the memory. The second approach, which is the method used in my project, is to set the DMD as a second monitor of the PC with the same resolution as the pixel resolution of the DMD. Then create the patterns to be displayed on the DMD in the form of images or videos. Set the images or videos to play in full screen mode on the second monitor and the patterns will be displayed on the DMD. The DMD is set to operate in binary mode for this project. If required, the DMD can operate in up to 8-bit mode, providing 256 levels of intensity for every spectral component. The hyperspectral projector can also operate in IR with an IR light source and optical elements that operate in IR.

3.2. Compressive Structured Light for Recovering Volume Density of Participating Medium

Conventional structured light approaches for recovery of 3-D shape of opaque objects are based on a common assumption: each point in the camera image receives light reflected from a single surface point in the scene. Meanwhile the light
transport model is vastly different in the case of a participating medium such as translucent objects, smoke, clouds and mixing fluids [2]. Consider an image acquired by photographing a volume of a participating medium. Unlike the case of an opaque object, here each pixel receives scattered light from all points along the line of sight within the volume.

Shree Nayar and co-workers [2] proposed the compressive structured light method for recovering the volume density of participating media. By using coded patterns the measurement of a participating medium is highly efficient in terms of acquisition time as well as illumination power. It exploits the fact that the brightness measurements made at image pixels correspond to true line-integrals through the medium (Figure 9) [2]. They target low-density inhomogeneous media, for which the density function is sparse in an appropriately chosen basis; this allows the use of compressive sensing techniques that accurately reconstruct a signal from only a few measurements. In this section I will explain their model and experiments in more detail.

3.2.1. Image Formation Model

In their compressive structured light system [2] (Figure 9), the projector displays coded patterns of binary black and white stripes into the volume of participating medium in the direction of z-axis, and the camera faces orthogonally in
the direction of the $x$-axis and captures the image of the scattered light from volume. The medium density is denoted by $\rho(x, y, z)$, the image intensity received by the camera is $I(x, y)$, and the projector radiance is $L(x, y)$. Because the direction of projection and the camera gaze are perpendicular and that the target volume is nonemissive and low-density, the light captured by the camera can be regarded as only composing of single-scattered light of the projection by the medium. Multiple scattering is assumed to be negligible. As shown in Figure 9(b), each camera pixel receives light scattered from a row of voxels along the line of sight in the volume (i.e., the red line in Figure 1b). For simplicity, we assume the camera and the projector are placed sufficiently far from the working volume, and thus they form an orthographic projection. The distortion caused by perspective projection can be corrected with a calibration step, if needed.

![Figure 9](image.png)

**Figure 9** (a) Compressive structured light for recovering participating media. Coded light is emitted along the $z$-axis to the volume while the camera
acquires images as line-integrated measurements of the volume density along the $x$-axis. Volume density is reconstructed from the acquired measurements by using compressive sensing techniques [32]. (b) Image formation model for participating medium under single scattering. The image intensity at one pixel, $I(y, z)$, depends on the integral along the $x$-axis of the projector's radiance, $L(x, y)$, and the medium density, $\rho(x, y, z)$, along a ray through the camera center [2].

Consider one voxel in the row $\rho(x, y, z)$. Light emitted from the projector, $L(x, y)$ is first attenuated as it travels from the projector to the voxel, scattered at the voxel, and then attenuated as it travels from the voxel to the camera. Assuming single scattering, the radiance sensed by the camera from this particular voxel is [33]

$$L(x, y) \cdot \exp(-\tau_1) \cdot \sigma_s \cdot \rho(x, y, z) \cdot p(\theta) \cdot \exp(-\tau_2)$$

where $\rho(x, y, z)$ is the volume density (i.e., density of particles) at the voxel, $p(\theta)$ is the phase function ($\theta = \pi/2$ since the camera and the projector are perpendicularly placed), and $\tau_1$ and $\tau_2$ are the optical thicknesses from the projector to the voxel and from the voxel to the camera; $\sigma_s$ is the scattering cross
section of the participating medium. Since $\sigma_s$ and $p(\theta = \pi/2)$ are the same for all voxels, the above formula can be simplified to

$$L(x, y) \cdot \exp(-\tau_1 - \tau_2) \cdot \rho(x, y, z)$$ \hspace{1cm} (6)

The image intensity, $L(x, y)$ which is the integral of the scattered light from all the voxels along the line, is therefore

$$I(y, z) = \int_x L(x, y) \cdot \exp(-\tau_1 - \tau_2) \cdot \rho(x, y, z) \, dx$$ \hspace{1cm} (7)

For highly diluted media (i.e. $\rho \to 0$), because the optical thicknesses $\tau_1$ and $\tau_2$ which are proportional to the density $\rho$ are close to 0, the attenuation term usually can also be ignored (i.e., $\exp(-\tau_1 - \tau_2) \approx 1$) for the recovery of volume densities [34], [35]. In this case, equation (7) is reduced to a linear projection of the illumination and the volume density

$$I(y, z) \approx \int_x \rho(x, y, z) \cdot L(x, y) \, dx$$ \hspace{1cm} (8)
3.2.2. Coding and Formulation

Unlike the conventional structured light methods for surface recovery where each camera pixel receives light reflected from one point, for participating media each camera pixel receives light from all points along the line of sight within the volume. Thus each camera pixel is an integral measurement of one row of the volume density. The compressive structured light seeks to reconstruct the 1D density signal from a few measured integrals of this signal.

Figure 10 Temporal coding of the volume using compressive structured light

Suppose we want to reconstruct a volume at the resolution $N \times N \times N$. The measurement vectors used for compressive sampling are $\{\varphi_k\}$, for $k = 1, \ldots, M$, where $\varphi_k$ is a $N \times 1$ random binary vector and each entry is drawn from i.i.d. Bernoulli distribution with a value of 0 or 1. As shown in Figure 10, the projector faces the $z$-direction and projects a sequence of patterns of binary black
and white stripes. Each pattern corresponds to a measurement vector $\varphi_k$. At each pattern, if attenuation is not considered, all rows of the volume in the $x$ direction with varying $y$ and $z$ coordinate values are encoded with the same $\varphi_k$. Therefore, every row in the $x$ direction of the volume can be regarded as an independent $N \times 1$ signal $x$ and the volume composes of $N^2$ such $N \times 1$ signals which are encoded with the same $\{\varphi_k\}$, for $k = 1, \ldots, M$. The camera faces the $x$-direction and takes an image at each pattern. Suppose the camera sensor has a resolution of $c \times c$ pixels where $c \geq N$. Group the neighboring $\frac{c}{N} \times \frac{c}{N}$ pixels to form a superpixel and sum up the intensities of these pixels to be the intensity of the superpixel. Thus, every image can be regarded as having $N \times N$ superpixels. Assuming no attenuation for now, the intensity for each of these $N \times N$ superpixels is a linear projection of the light and the voxels’ density from equation (8). Let $x = [\rho_1, \ldots, \rho_N]^T$ be the vector of the voxel densities along a fixed row of the volume. The intensity values of the a fixed superpixel of all the $M$ images form the measurement vector $y$ and each entry of $y$ is

$$y_k = \langle x, \varphi_k \rangle, \quad k = 1, \ldots, M$$

(9)

Rewriting these $M$ equations in matrix form, we have

$$y = \Phi x$$

(10)

Thus, the problem of recovering the volume is formulated as the problem of reconstructing a set of $N^2$ of 1-D signals along $x$ - axis from a few integral
measurements. Compared to sequential laser scanning, compressive structured light enjoys the advantages of compressive sensing and utilizes the light more efficiently, thus making the measurement process highly efficient both in acquisition time and illumination power.

3.2.3. Measurement Data Reconstruction

In [2], the authors used the compressive structured light system to recover several types of participating media, including multiple translucent layers (Figure 11) [2], a 3-D point cloud of a face etched in a glass cube, and the dynamic process of milk mixing with water. Here I use their reconstruction of the static volume of multiple translucent layers as an example for explaining reconstruction. Figure 11 shows their reconstruction results of an object consisting of two glass slabs with powder on both [2]. The letters “EC” are drawn manually on the back plane and “CV” on the front plane by removing the powder. Thus in the volume only two planes have nonzero density.
Figure 11 Reconstruction results of two planes. (a) A photograph of the object consisting of two glass slabs with powder. The letters “EC” are on the back slab and “CV” on the front slab. (b) One of the images captured by the camera. (c) Reconstructed volume at different views without attenuation correction [2].

In compressive sensing, we can have far fewer measurements than the number of unknowns, which means the above equation is an underdetermined linear system and optimization is required to solve for the best $\mathbf{x}$ according to certain prior structure of the signal. In Chapter 2, the compressive measurement is formulated as

$$\mathbf{y} = \Phi \mathbf{x} = \Phi \Psi \alpha$$

where the signal $\mathbf{x}$ is generally assumed to be sparse or compressible under some transformation $\Psi$. In the case of recovering the multiple translucent layers (Figure 11), in the volume only two planes have nonzero density. This suggests that the signal value itself is sparse, or, put in another way, $\Psi = \mathbf{I}$ where $\mathbf{I}$ is the identity matrix. So the $l_1$-norm of the signal value could be used as the objective function for
minimization. Therefore the reconstruction problem for the transparent layers is formulated as

$$\hat{x} = \arg\min_{x} \|x\|_1 \quad \text{such that} \quad \|\Phi x - y\|_2 < \epsilon \quad \text{and} \quad x \geq 0 \quad (11)$$

There are total of $N^2$ such reconstruction problems to solve to get the density distribution of the whole volume.

In their experiments, the structured patterns they used are in black and white and the targets used for recovery are all white. Also they focused solely on the visible portion of the spectrum. In the next chapter, I will demonstrate using hyperspectral structured illumination for recovering spectrum-dependent 3-D volume density of colored targets.
4. Hyperspectral Compressive Structured Light

In this Chapter, I present the results of using our hyperspectral projector system performing compressive structured light for 3-D volume density recovery. Initially, the conventional black and white structured light patterns are implemented as described in the previous section for encoding the volume of a static translucent object and the 3-D volume density of the object is recovered. This experiment shows the feasibility and performance of the system in recovering volume density. Subsequently, hyperspectral structured light patterns are used to recover spectrum-dependent 3-D volume density of colored static translucent objects to demonstrate the unique advantage of using the hyperspectral projector in recovery of 3-D volume density of colored objects.
4.1. Black and White Compressive Structured Light

In this experiment, the 3-D volume density of a colorless static translucent object is captured using the proposed hyperspectral projector system through the compressive structured light method developed by Gu J, Nayar S K, et al [2]. Black and white structured light patterns are used for encoding the volume and reconstruction results are demonstrated. This experiment shows the feasibility and performance of the hyperspectral projector system in recovering 3-D volume density of participating media.

4.1.1. Experiment Design

Figure 12 is a photograph of our experimental. The camera faces the $z$-direction, which in our case is horizontal, and pattern projection is along the $x$-direction which is vertically from the top. With this configuration, we will reconstruct the data as described in the previous chapter. The camera used is the Mightex USB2.0 Monochrome 1.3MP CMOS Camera.
Figure 12 Experimental setup of compressive structured light using the proposed hyperspectral projector system.

The target for reconstruction (Figure 13 (a)) is a static volume of two white translucent planes. The planes are made by roughening a sheet of transparency with sand paper so it scatters white light. The letter “C” is carved manually on each of the two planes by manually removing the plane material. The letter “C” curves upward on the front plane, and downward on the back plane to differentiate between the front and back plane. Thus in the volume only two planes have nonzero density.
Figure 13 (a) Target used for the experiment. The letter “C” is carved manually on each of the front and back planes by removing the plane material. The “C” on different planes curls in opposite directions. (b) Example images of the coded volume captured by the camera.

The binary stripe patterns are used as the coded light patterns and projected downward. Each pattern has 32 stripes (Figure 13 (b)) so that a volume of resolution $32 \times 32 \times 32$ is recovered. The stripes are randomly assigned to be 0 (black) or 1 (white) according to Bernoulli distribution (with $p=0.5$). The coded images are captured by the camera and the area of interest that will be recovered are cropped from full images. The cropped images, which correspond to the $x$-$y$ plane of the volume, are turned into images with resolution of $32 \times 32$ by summing
up neighboring pixels. In the data reconstruction, for the simple one-dimensional $L_1$ norm optimization, the Matlab function `linprog` is sufficient. The Matlab code for reconstruction is adapted from the code downloaded from [32].

### 4.1.2. Reconstruction Results

Figure 14 shows the reconstruction results of the 3-D volume density of the target at resolution of $32 \times 32 \times 32$ using 24 compressive measurements. In Figure 14(a), the 3-D views of the reconstruction from two perspectives are displayed. The reconstructed 3-D volume density data is first normalized by a threshold to remove noisy points and then plotted in a 3-D scatter plot where the color of the points indicates the density value at that point. It is clearly seen that the two planes and the two letter “C”s are reconstructed. The “C” that curves upwards on the front plan is fully reconstructed and distinctly visible. The “C” that curves downwards on the back plane is almost fully reconstructed except that parts of the backplane are missing. It is due to attenuation of the light coming from the back plane so the reconstructed volume density of the back plane has smaller values and some of the points are lost in the thresholding. The ridge connecting the two planes are in red with much larger density values because of its proximity to the light. Figure 14 (b), (c) and (d) demonstrate some example 2D slices of the reconstructed 3-D volume density in $y$-$x$, $z$-$x$, $x$-$y$ views, respectively. The two planes are distinct in the 2D slices and location of the “C” appears as holes in the two planes. The plane with higher intensity is the front plane.
Figure 14 Reconstruction results of the 3-D volume density of the target of the two planes at resolution of $32 \times 32 \times 32$ using 24 compressive measurements. (a) 3-D views of the reconstruction from two perspectives. (b)(c)(d) Example 2D slices of the reconstructed 3-D volume density in $y$-$x$, $z$-$x$, $x$-$y$ views, respectively. The number on the corner of each image is coordinate index of the image in the dimension of slicing. The two planes are distinctive in the 2D slices and locations of the “C” appear as holes in the two planes. The plane with higher intensity is the front plane.

Figure 15 shows the reconstruction results of the 3-D volume density of the target at resolution $128 \times 128 \times 128$ using 64 compressive measurements. Compared to a raster scan using single stripe patterns that requires 128 measurements, only half number of measurements are needed. In Figure 15(a), the 3-D views of the reconstruction from two perspectives are displayed. The reconstructed 3-D volume
density data is first normalized by a threshold to remove noisy points and then plotted in a 3-D scatter plot where the color of the points indicates the density value at that point. Same as with $32 \times 32 \times 32$ reconstruction, the two planes and the two letter “C”s are reconstructed. The “C” on the front plan is distinctly visible. The “C” on the back plane is almost fully reconstructed except that part of the backplane is missing due to attenuation of light. The ridge connecting the two planes has much larger density values. Figure 15 (b), (c) and (d) demonstrate some example 2D slices of the reconstructed 3-D volume density in y-x, z-x, x-y views, respectively. The two planes are distinct in the 2D slices and location of the “C” appears as holes in the two planes. The plane with higher intensity is the front plane.
Figure 15 Reconstruction results of the 3-D volume density of the target of the two planes at resolution of $128 \times 128 \times 128$ using 64 compressive measurements. (a) 3-D views of the reconstruction from two perspectives. (b)(c)(d) Example 2D slices of the reconstructed 3-D volume density in y-x, z-x, x-y views, respectively. The number on the corner of each image is coordinate index of the image in the dimension of slicing. The two planes are distinctive in the 2D slices and locations of the “C” appear as holes in the two planes. The plane with higher intensity is the front plane.

4.2. Hyperspectral Compressive Structured light

Following our initial success with broadband illumination, hyperspectral structured light patterns are used to recover spectrum-dependent 3-D volume
density of colored static translucent objects. Objects with different colors in the same scene are reconstructed individually demonstrating the unique advantage of using the hyperspectral projector in spectrum-dependent recovery of 3-D volume density of colored objects.

4.2.1. Experiment Design

In order to demonstrate the advantage of the spectral dimension of the hyperspectral projector system, color transparencies are used as targets here instead of white ones in the previous experiment. The target contains two objects placed close together as shown in Figure 16. One object consists of two translucent planes of red color with letter “C” carved on each of the front and back planes, where the letter “C” curves in opposite directions to differentiate between front plane and back plane. Similarly, the other object consists of two translucent planes of cyan color with letter “V” carved on each of the front and back planes, where the letter “V” curves in opposite directions to differentiate between front plane and back plane. Instead of roughening the transparency to make it white as in the previous experiment, these objects are made by printing color toners on the transparencies. Red and cyan are specifically selected because these two colors have almost non-overlapping responses of reflectance spectra in the visible region. Figure 16 (c) shows the reflectance spectra of the two colors printed on the transparency. Between 390 nm and 590 nm, cyan has strong reflectance and red has very weak reflectance. Between 590 nm and 750 nm the situation is reversed where red is
strongly reflective and cyan is fairly weak. The spectra are plotted using the measured spectra of the two colors after they have been normalized with respect to the illumination spectrum.

Figure 16 The target and its spectrum. (a) Photo of the target for reconstruction which contains two objects placed close together: one object comprises of two red translucent planes with letter “C” carved on each of the front and back planes, the other consists of two cyan translucent planes with letter “V” carved on each of the front and back planes. (b) Image of the target taken from the perspective of the camera using in the experiment under white illumination. (c) Reflectance spectra of the red and cyan planes. Red has strong reflectance between 590 nm and 750 nm, while cyan is strongly reflective between 390 nm and 590 nm.
4.2.2. Hyperspectral 3-D Reconstruction

The experiment uses two sets of structured patterns that have the same binary stripe coding scheme but different spectral content. The spectrum of each set of patterns is designed to match to the reflectance spectrum of each of the two colors to selectively recover the volume density of the object that we want. As shown in Figure 17, the first set of structured patterns contain spectral content of greater than 610 nm, under which the red object is illuminated but the cyan object is almost invisible. The second set of structured patterns contains spectral content of less than 570 nm, under which cyan red object is illuminated but the red object is almost invisible. In the volume coding process, the two sets of patterns are projected on the target in sequence, and in reconstruction, the two sets are used separately to generate two reconstruction results. The first reconstruction contains the volume density of the red object and second reconstruction contains the cyan object.
Figure 17 (a) Image of the camera of the target under an example structured light pattern of wavelength longer than 610 nm, where the red object is encoded and the cyan object is invisible. (b) Spectrum of the first set of structured patterns. (c) Image of the camera of the target under an example structured light pattern of wavelength shorter than 570 nm, where the cyan object is encoded and the red object is invisible. (d) Spectrum of the second set of structured pattern.
The red and cyan objects in the same scene can be reconstruction separately using hyperspectral structured patterns as described above. Figure 18 shows the reconstruction results of the 3-D volume density of the red object at resolution of $32 \times 32 \times 32$ using 24 compressive measurements. In Figure 18(a), the 3-D views of the reconstruction from two perspectives are displayed. The reconstructed 3-D volume density data is first filtered by a threshold to remove noisy points and then plotted in a 3-D scatter plot where the color of the points indicates the density value at that point. It is clearly seen that the two planes and the two letter “C”s are reconstructed. The “C” that curves upwards on the front plan is fully reconstructed and distinctly visible. The “C” that curves downwards on the back plane is almost fully reconstructed except that part of the backplane is missing due to attenuation. The ridge connecting the two planes has larger density values. Figure 18 (b), (c) and (d) demonstrate some example 2D slices of the reconstructed 3-D volume density of red object in y-z, x-y, x-z views, respectively. The two planes are distinct in the 2D slices and location of the “C” appears as holes in the two planes. The plane with higher intensity is the front plane.
Figure 18 Reconstruction results of the 3-D volume density of the red object of 32 × 32 × 32 using 24 compressive measurements. (a) 3-D views of the reconstruction from two perspectives. (b)(c)(d): Example 2D slices of the reconstructed 3-D volume density in y-z, x-y, x-z views, respectively. The number on the upper right corner of (b) (d) and lower corner of (c) of each image is coordinate index of the image in the dimension of slicing. The two planes are distinctive in the 2D slices and locations of the “C” appear as holes in the two planes. The plane with higher intensity is the front plane.

Figure 19 shows the reconstruction results of the 3-D volume density of the cyan object at resolution of 32 × 32 × 32 using 24 compressive measurements. In Figure 19(a), the 3-D views of the reconstruction from two perspectives are displayed. The reconstructed 3-D volume density data is first filtered by a threshold to remove noisy points and then plotted in a 3-D scatter plot where the color of the
63

points indicates the density value at that point. It can be seen that the two planes and the two letter “V”s are reconstructed. The “V” that curves upwards on the front plan is fully reconstructed and distinctly visible. The “V” that curves downwards on the back plane is almost fully reconstructed except that part of the backplane is missing due to attenuation. The ridge connecting the two planes has larger density values. Figure 19 (b), (c) and (d) demonstrate some example 2D slices of the reconstructed 3-D volume density of red object in in y-z, x-y, x-z views, respectively. The two planes are distinct in the 2D slices and location of the “V” appears as holes in the two planes. The plane with higher intensity is the front plane.

(a)
Figure 19 Reconstruction results of the 3-D volume density of the red object of $32 \times 32 \times 32$ using 24 compressive measurements. (a) 3-D views of the reconstruction from two perspectives. (b)(c)(d): Example 2D slices of the reconstructed 3-D volume density in y-z, x-y, x-z views, respectively. The number on the upper right corner of each image is coordinate index of the image in the dimension of slicing. The two planes are distinctive in the 2D slices and locations of the “V” appear as holes in the two planes. The plane with higher intensity is the front plane.

The reconstruction results using hyperspectral compressive structured light demonstrate that the red and cyan objects in the same scene can be reconstructed separately. This experiment serves as an example that the hyperspectral projector system can be used for revealing spectrum-dependent information of the target. This feature could be very useful for a lot of applications. For example, in imaging the 3-D volume density of the dynamic process of mixing fluids of different colors, the development of density distribution of each type of fluid can be separately reconstructed. Another example is imaging biological tissues where more than one
type of fluorescence markers is present for labeling different molecules/locations/cells. Different fluorescence markers have unique spectral responses to the illumination and the hyperspectral compressive light method can be used to reconstruct each type of markers separately.
Chapter 5

5. Compressive Sensing Classification using a Neural Network

5.1. Compressive Classification

Classification is of great importance in a wide variety of real-world camera applications. Accurate and fast classification on vehicles could be beneficial in monitoring traffic conditions, reducing congestion, fare collection, fare and toll collection, roadside services, traffic offence ticketing and so on [12]. Meanwhile, vehicle images and videos in the infrared region are able to reveal different details of the scene than in the visible region which could be useful and desirable in many situations. However, high resolution imaging and video in infrared (IR) is more expensive compared to the silicon-based consumer digital cameras. As described in Chapter 2, the SPC is a simpler, smaller, and cheaper camera architecture that can operate efficiently in IR. Yet in many data acquisition/processing applications, we
are not interested in obtaining a precise reconstruction, but rather are only interested in making some kind of detection or classification decision. For instance, in vehicle classification, we simply wish to identify the class to which the vehicle belongs out of several possibilities. We know that a set of images of a fixed scene under varying articulation parameters forms a low-dimensional, nonlinear manifold, and it has been shown that random projections stably embed a smooth manifold in a lower-dimensional space [38]. Thus random projections in compressive sampling can be regarded as a dimension-reducing process and can be used as input to the neural network for classification. In this Chapter, I present a two layer, feed-forward neural network architecture and use it to classify IR vehicle images with compressive measurements. This framework gives promise to building a single-pixel camera that can do vehicle detection and classification in IR without reconstructing the original image.

### 5.2. Neural Network Architecture

This section details a two-layer feed-forward network for compressive vehicle classification. The neural network receives random projections of vehicle images as inputs. In the model example, IR images of three types of vehicles are classified into three categories: Ram, Corolla and Frontier. I use the Matlab R2014a Neural Network Pattern Recognition application for building, training and testing the neural network. The network architecture is shown in Figure 20.
Figure 20 Neural Network Architecture

It is a two-layer feed-forward network, with sigmoid hidden and softmax output neurons. Objective function used is cross-entropy. Such architecture can classify vectors arbitrarily well, given enough neurons in its hidden layer. There are three output neurons to represent three classes. The label/target of each class is assigned as follows: [1,0,0] for Ram, [0,1,0] for Corolla, [0,0,1] for Frontier. The classification criteria is winner take all. The network is trained with scaled conjugate gradient backpropagation.

5.3. Results

5.3.1. Classification on Video Chips

The IR image data used in this project are extracted from IR videos of three vehicles provided by the United States Air Force\(^1\). Also provided along with the

\(^1\) "Distribution A. Approved for public release, distribution unlimited. (96TW-2015-0103)"
videos are the 64*64 chips containing solely the vehicles with background subtracted. The chips are extracted from long wave infrared (LWIR) videos of three vehicle classes: Ram, Corolla, and Frontier. There are a total of 196 chips for Ram, 86 for Corolla, and 82 for Frontier. The images for each class contain the vehicle placed in all rotation angles. Figure 21 shows some example chips:

![Example chips for each class of vehicles used for training and testing. The resolution of the chips is 64*64.](image)

In the simulation, the measurement matrix used to generate compressive samplings of these chips is the 4096*4096 double-permuted Walsh-Hadamard matrix where the same measurement matrix is used for taking compressive measurements of all chips of the same resolution. The algorithm randomly splits the whole dataset into three sets: 70% of all chips are for training, 15% for validation, and 15% for testing. Various proportions of the full compressive measurements and different numbers of hidden neurons in the hidden layer are
tried for optimal performance. Figure 22 shows the neural network architectures and confusion matrices of test results. All the classification results achieve an excellent error rate of zero percent.
Figure 22 Confusion matrices and neural network architectures of test results.

All the classification results achieve an excellent error rate of zero percent.
5.3.2. Classification on Video Patches

The short wave infrared (SWIR) videos of each of the Ram, Corolla and Frontier models are used. In each video, the vehicle moves around in an elliptical route on the background. I select an area of size 64*256 from all frames of the videos such that the moving vehicle is fully contained in this area in each frame. The images contain the vehicles in all rotation angles. There are a total of 2752 images for Ram, 3598 for Corolla, and 2155 for Frontier. Figure 23 shows some example patches:

Figure 23 Example video patches for the three classes.
In the simulation, the measurement matrix used to generate compressive
samplings of these images is the 16384*16384 double-permuted Walsh-Hadamard
matrix. The same measurement matrix is used for taking compressive
measurements of images of the same resolution. Same as with video chips, there are
three output neurons to represent three classes. The label/target of each class is
assigned as follows: [1,0,0] for Ram, [0,1,0] for Corolla, [0,0,1] for Frontier. The
algorithm randomly splits the whole dataset into three sets: 70% of all chips are for
training, 15% for validation, and 15% for testing. Various proportions of the full
compressive measurements are used as inputs to the neural network. And different
numbers of hidden neurons in the hidden layer are tried for optimal performance.
Figure 24 shows the neural network architectures and confusion matrices of test
results. All the classification results achieve an excellent error rate of zero percent.
Figure 24 Confusion matrices and neural network architectures of test results.

All the classification results achieve an excellent error rate of zero percent.
5.3.3. Classification under Noise

The robustness of the neural network under noise is tested. The images used are synthesize images generated by inserting the 64*64 vehicle chips into a 256*256 background extracted from the video. Figure 25 are example images.

![Figure 25 Synthesized images of the three classes of vehicles.](image)

The training data are clean images without adding noise. And test data are the clean images with different levels of Gaussian noise added. Gaussian noise is added in the fashion of SNR. Below are examples of image data before and after adding noise. Figure 26 show example images before and after adding noise.
Figure 26 First row: an image before and after adding Gaussian noise of 10 dB.
Second row: an image before and after adding Gaussian noise of 20 dB.

In the simulation, the measurement matrix used to generate compressive samplings of these chips is the 65536*65536 double-permuted Walsh-Hadamard matrix. The same measurement matrix is used for taking compressive measurements of all chips of the same resolution. Figure 27 shows the neural network architectures and confusion matrix of testing results. The result shows that
the neural network is robust to noise in the test image data. 10 hidden neurons gives the best results. More hidden neurons cause overfitting problem.
(a) Noise level = 20 dB, 1/16 measurements, 30 hidden neurons.

(b) Neural network architecture used in (a).

(c) Noise level = 10 dB, 1/16 measurements, 5 hidden neurons.

(d) Neural network architecture used in (c).

(e) Noise level = 10 dB, 1/16 measurements, 10 hidden neurons

(f) Neural network architecture used in (c).
Figure 27 Confusion matrices and neural network architectures of test results.

The result shows that the neural network is robust to noise in the test image data.
6. Conclusion and Future Work

Two projects are demonstrated in this thesis. Initially, I illustrate the design of a hyperspectral projector system based on a single DMD that is able to generate hyperspectral structured illumination of arbitrarily desired spectral content of multiple/single wavelengths/wavebands, and implement black and white/hyperspectral compressive structured light method to recover spectrum-dependent 3-D volume density of translucent objects. The experimental results show correct reconstructions of colorless objects and spectrum-dependent reconstructions of colored objects. Subsequently, I demonstrate the effectiveness of compressive sensing classification method using a proposed two layer feed-forward neural network on the example model of vehicle classification. Zero classification error rate is achieved with clean image data and very small error rate is achieved for noisy images.

A future application of our hyperspectral projector system is to build a public hyperspectral image library spanning the spectrum from ultraviolet to the infrared to advance the development of analysis in the machine vision community by coupling this
projector system with standard, broadband visible and infrared cameras. By doing so, we hope to better understand how significant are the benefit and what spectral resolution is necessary in object identification and human motion inference. Also the hyperspectral projector system could be used for the calibration and testing of hyperspectral imagers. With the compressive sensing classification method, a future direction is to test the robustness of the neural network on vehicle translations by generating datasets that include vehicles in different translated locations on the background. Also more work could be done on the feasibility of identifying the angle of the vehicle and on the classification on vehicles that are partially blocked.
References


[37] DLP® LightCrafter™ 4500 Evaluation Module User’s Guide (Rev. E)