DEVELOPMENT OF PLENOPTIC INFRARED CAMERA USING LOW DIMENSIONAL MATERIAL BASED PHOTODETECTORS

By

Liangliang Chen

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Electrical Engineering - Doctor of Philosophy

2016

ABSTRACT

DEVELOPMENT OF PLENOPTIC INFRARED CAMERA USING LOW DIMENSIONAL MATERIAL BASED PHOTODETECTORS

By

Liangliang Chen

Infrared (IR) sensor has extended imaging from submicron visible spectrum to tens of microns wavelength, which has been widely used for military and civilian application. The conventional bulk semiconductor materials based IR cameras suffer from low frame rate, low resolution, temperature dependent and highly cost, while the unusual Carbon Nanotube (CNT), low dimensional material based nanotechnology has been made much progress in research and industry. The unique properties of CNT lead to investigate CNT based IR photodetectors and imaging system, resolving the sensitivity, speed and cooling difficulties in state of the art IR imagings.

The reliability and stability is critical to the transition from nano science to nano engineering especially for infrared sensing. It is not only for the fundamental understanding of CNT photoresponse induced processes, but also for the development of a novel infrared sensitive material with unique optical and electrical features. In the proposed research, the sandwich-structured sensor was fabricated within two polymer layers. The substrate polyimide provided sensor with isolation to background noise, and top parylene packing blocked humid environmental factors. At the same time, the fabrication process was optimized by real time electrical detection dielectrophoresis and multiple annealing to improve fabrication yield and sensor performance. The nanoscale infrared photodetector was characterized by digital microscopy and precise linear stage in order for fully understanding it. Besides, the low noise, high gain readout system was designed together with CNT photodetector to make the nano sensor IR camera available.

To explore more of infrared light, we employ compressive sensing algorithm into light field

sampling, 3-D camera and compressive video sensing. The redundant of whole light field, including angular images for light field, binocular images for 3-D camera and temporal information of video streams, are extracted and expressed in compressive approach. The following computational algorithms are applied to reconstruct images beyond 2D static information. The super resolution signal processing was then used to enhance and improve the image spatial resolution. The whole camera system brings a deeply detailed content for infrared spectrum sensing.

ACKNOWLEDGMENTS

I feel tremendously lucky to have had the opportunity to work with Dr. Ning Xi, Dr. Lixin Dong on the ideas in this dissertation, and I would like to thank them for their guidance and support. Dr. Xi instilled in me a love for designing nanosensor based infrared camera, agreed to take me on as a graduate student, and encouraged me to immerse myself in something I had a passion for. Dr. Dong inspired me much on nanophotonics enhancement and nano manipulation. I have never met a professor more generous and friendly with his time and experience.

I am grateful to Dr. Timothy Grotjohn, Dr. Fathi M. Salem, Dr. Donnie Reinhard and Dr. Zhengfang Zhou for serving on my thesis committee. They offered me timely help and unfailing support that improve the technical soundness and the presentation of this dissertation.

I would like to acknowledge the fine work of the other individuals who have contributed to this camera research. Dr. Baokang Bi and Dr. Reza Loloee in Department of physics and astronomy, helped me on nanosensor fabrication and assembly device. Dr. Loloee generously donated his time and expertise to help verify fabrication process. Dr. Ming Yan in Department of mathematics contributed the most to explaining how the nonconvex works, and many of the optimization problem in these pages are due to his artistry. Dr. Weihong Guo in case reserve university, supported me on single image super resolution. Mr. David Smith in wintechdigital, gave much guidance on digital micromirror setup and applications.

In addition, I would like to thank Dr. King Wai Chiu Lai, Dr. Carmen Kar Man Fung, Dr. Hongzhi Chen, Dr. Ruiguo Yang, Mr. Bo Song, Dr. Zhanxin Zhou, Dr. Yongliang Yang, Mr. Zhiyong Sun, Dr. Chi Zhang, Dr. Erick Nieves, Dr. Jianguo Zhao, Dr. Yunyi Jia, Mr. Yu Cheng, Mr. Mustaffa Alfatlawi, Mr. Emad Alsaedi, Mr. Lai Wei and Mr. Xiao Zeng for their support in the experiments and discussion at the MSU Robotics and Automation Lab. Thanks also to Dr.

Tong Jia, Dr. Jianyong Lou for helpful discussions related to this work.

Finally, I would like to thank my family, my wife, Qiaozhi Sun, for their love and support. This dissertation would not have been possible without their years of encouragement and continuous support. My wife have made countless sacrifices for me, and have provided me with steady guidance and encouragement.This dissertation is dedicated to them.

To my parents, my wife Qiaozhi

for their love and support

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1

Introduction

1.1 Infrared Everywhere

1.1.1 Fundamental of Infrared

The first discovery of electromagnetic radiation other than visible light came in 1800, when William Herschel discovered infrared (IR) radiation [1]. It is widely used in civilian application from industrial, agricultural, scientific, night vision, building inspection, medical thermography and meteorology, medical diagnosis due to that the IR camera can explore more information than visible light camera. Most importantly, it operates in night and long distance compared to visible light camera so that it is becoming one of most popular non-destructive diagnostic technology in industry. There are thousands of commercialized applications, including hyperspectral imaging in biological and mineralogical measurements, target acquisition and tracking, night-vision [2], IR data communications by standards published by IrDA, Infrared telescope in astronomy, environment monitoring in meteorology, etc. Besides, infrared photography is not only applicable in industry, but also for scientific research. The most popular application is Fourier Transform InfraRed spectroscopy (FTIR) [3], which determines what fraction of the incident radiation is absorbed by passing infrared radiation through a sample. Another counterpart application is that Stimulated Raman Scattering (SRS) benefits from the use of pulsed near-infrared lasers, which generates high signal levels at a moderate average power in biomedical cutting edge research [4]. Meanwhile, the infrared photog-

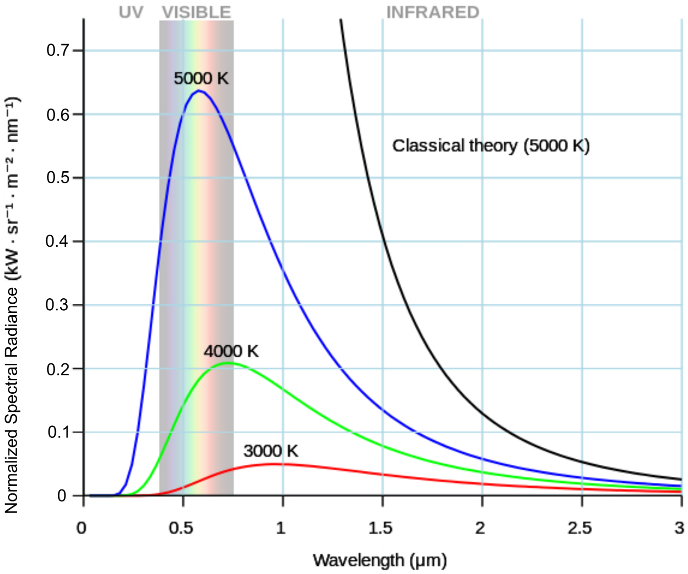


Figure 1.1 Planck's law (colored curves) and classical theory (black curve). For interpretation of the references to color in this and all other figures, the reader is referred to the electronic version of this thesis.

raphy discovers beneath a paintings surface and views detail that would otherwise remain unseen in art science. It is also applied to detect disease, insect infestation in plants science [5]. In the field of medicine science, medical infrared thermography is a non-invasive, non-radiating low cost detection method for analyzing physiological functions in sports medicine as traumatic knee injuries [6], cancer diagnostics [7]. The most widely applications come from military, where infrared light is extensively employed for target acquisition in wide IR bandwidth.

Infrared thermography detects electromagnetic spectrum from 720 nm up to 14 μm. Since IR radiation is emitted by all objects above absolute zero kelvin and the irradiated wavelength depends on its temperature according to the blackbody radiation law [8], thermography allows one to observe variations in temperature. This technology is especially related to human body because humans at ambient room temperature can radiate around 12 μm wavelength infrared light based on Wiens displacement law [8].

The planck's law describes the black body electromagnetic radiation in thermal equilibrium.

As shown in Eq. (1.1), B_v represents the spectral radiance which tells the amount of energy at different wavelength in absolute temperature T, where k_B is the Boltzmann constant, h is the planck constant, c is the speed of light in the medium. In Figure 1.1, the colored curves show 5000 K, 4000 K and 3000 K energy radiance according to wavelength respectively. In long wavelength range, the Planck's law tends to be Rayleigh Jeans law, while it is close to Wien approximation in short wavelength [9]. The peak wavelength (λ_{max}) can be numerically evaluated by solving mathematical equation Eq. (1.1) and concluded as Eq. (1.2).

$$B_v(v,T) = \frac{2hv^3}{c^2} \frac{1}{e^{\frac{hv}{k_B T}} - 1} \tag{1.1}$$

$$\lambda_{max} = \frac{hc}{x} \frac{1}{kT} = \frac{2.89776829 \times 10^6 nm \cdot K}{T} \tag{1.2}$$

The IR images measure the infrared irradiation and its distributions. There are two IR sources for imaging, one is internal emissivity as Planck's law depicted and another is external reflection similar as visible light. The reflected light, e.g. sunlight, indoor lighting, is not only the predominate element for visible image sensors, but also contributes on near IR, SWIR imaging. In naturally, both sunlight and airglow at night generate near IR lightwave and radiate on object so that the camera could capture the reflective light to form a near IR image. The wavelength of man-made IR illuminator also locates within near IR and SWIR spectrum, e.g. light bulbs and solid state Light Emitting Diode (LED). The incandescent light bulbs heat a tungsten filament to high temperature and produce visible light but together with infrared radiation. However the solid state LED are more efficient with near monochromatic infrared energy, which depends on the spontaneous and stimulated emission. When an electron orbits the nucleus of an atom [10] in high energy state, it has chance spontaneously decay path to low energy. The electron decay in such a manner

Table 1.1 Infrared sub-division scheme.

Division Name	Abbreviation	Wavelength (μm)	Frequency (THz)	Photon Energy(meV)	Characteristics
Near-infrared	NIR, IR-A DIN	0.72-1.4	214-400	886-1653	Passive night vision devices
Short-wavelength infrared	SWIR, IR-B DIN	1.4-3	100-214	413-886	Water absorption and long-distance telecommunications
Mid-wavelength infrared	MWIR, IR-C DIN	3-8	37-100	155-413	Atmospheric window and thermal infrared above body temperature
Long-wavelength infrared	LWIR, IR-C DIN	8-15	20-37	83-155	The thermal imaging region, requiring no illumination
Far-infrared	FIR	15-1000	0.3-20	1.2-83	Far-infrared laser

will introduce a photon emitted at exactly the same wavelength and phase. It becomes man-made infrared light source when photon energy is within infrared spectrum as shown in Table 1.1. It lists the infrared classification by wavelength, frequency, photon energy and application.

1.1.2 Conventional Infrared Sensor

IR sensor can be broadly divided into two categories: cooled and uncooled detector. It also can be classified by detection mechanism: a thermal type that has no wavelength dependence (thermal detector) and a quantum type that is wavelength-dependent (photon detector) [11]. Infrared radiation energy is equal to the vibrational or rotational energy of molecules ranging from 1.24 eV at 1 μm down to 0.12 eV at 10 μm.

Many specialized IR sensor structures and materials have been manufactured and compared as thermistor for IR detection [12]. It senses the heat irradiation, through which the resistance will be changed, and a readout circuit monitors resistance variation to determine photocurrent response of

IR sensor. The whole process will cost too much time and make the sensor relatively slow response because the detector element is suspended on lags which are connected to the heat sink, though it works in room temperature and independent of wavelength.

The quantum detectors, including indium antimonide, indium arsenide, MCT [13], lead sulfide and lead selenide [14], usually need a cryogenically dewar for the operation of semiconductor materials. The fore infrared photodetectors are also known as quantum detectors which depend on the bandgap of materials. They offer higher detection performance and a faster response speed. However, they always work in very low temperature to keep high responsibility, so that a cryogenic cooled system is needed to maintain IR sensor work stable. Even though some other photoelectric effect quantum dot (QD) and quantum well (QW) IR sensors improve working temperature and it could detect 10 μm wavelength at room temperature [15], the sensitivity deteriorates much. At the same time, the bulky cooling system is against with portable design of imaging system.

1.1.3 Infrared Detector Market

Infrared camera (imager) is also named focal plane array (FPA) or IR sensors in infrared industry and scientific research. It is a thermal system which converts infrared radiation into a visible image and the core of camera includes electronics, IR lenses and sensor. In overall, the military demand dominates infrared camera market in last 50 years. However, since the military market for uncooled infrared imaging technologies declines, it turns toward commercial businesses such as personal vision and smartphones by Yole report [16].

Although the widely application of IR imaging covers consumer electronics, surveillance, aerospace and defense, automotive, industrial, medical, and firefighting, etc., the IR cameras market growth is being driven by the ultra-low-end market which consists of low-resolution cameras for basic radiometric purposes. It is estimated that the global infrared imaging market will reach

$8,450 million by 2020 [17]. The global infrared market is divided into four geographic segments, the American, Europe, Asia-Pacific, and Rest of the World. Major players in the IR imaging market are FLIR Systems Inc. (U.S.), DRS Technologies Inc. (U.S.), Fluke Corporation (U.S.), Axis Communications AB (Sweden), Samsung Techwin (South Korea), Seek Thermal Inc. (U.S.) and Sofradir Group (France). Undoubtedly, Americas is considered to be the leader in the overall IR imaging/sensors market.

There are at least five trends from application insight.

- Thermography: the ultra-low-end cameras with attractive pricing, lower than $1,000 drive the market. They will widen the customers of thermal technology. There are many new models released from leading companies, FLIR [18] and Fluke [19] in 2014 and 2015, and they are going to lead the price war due to their vertical business model.

- Automotive: the market leader Autoliv will continue to introduce night vision 3rd generation on new car models. New Euro NCAP could boost the market by promoting, in 2018, night time pedestrian collision mitigation solutions potentially using a thermal solution, but only if the cost is sufficiently low. FLIR Systems also provides camera cores for Rolls Royce, BMW, Audi and Mercedes Benz models through a partnership with Autoliv Electronics.

- Surveillance: infrared camera is a good surveillance robust equipment that can handle any environmental conditions for naval, air and ground security. SWIR works for counter-aerial, very long range land and sea, large area naval surveillance. Price erosion will continue (-12%/year) and will enlarge the scope of commercial applications like traffic, parking, and power stations.

- Consumer applications: this is the fast growth in 2013-2014. Personal vision systems (goggles, sight for security, and hunting, outdoor observation) dominate civil applications. Since there are many new entrants arriving from the outdoor visible business, it will continue to grow in future years.

- Smartphones: There are two smartphone modules (FLIR One, Opgal Android) and Seek Thermal been introduced at the ground breaking price of $349 by FLIR and $249 by Seek Thermal in 2014. A high number of pre-release reservations for the FLIR One already proves the commercial success of this smartphone platform. The smartphone business is an almost billion market. Those high volumes will only be possible if a huge cost reduction is obtained by the IR imaging industry.

At sensor level, major manufacturers (such as DRS, FLIR, Raytheon, ULIS, GWIC) have now moved to 8 inch production lines instead of 6 inch to reduce wafer manufacturing cost. Several have even outsourced their production to foundries to further reduce production cost. These two elements are preliminary signs of a strong microbolometer cost reduction that will open up cost driven applications such as smartphones.

At camera level, vertically integrated players, with internal sensor manufacturing, can benefit from their efficient cost structures to enter any commercial market with an aggressive camera price. For instance, DRS and FLIR lead the price war in surveillance while FLIR has introduced a low-cost infrared camera for firefighting in 2012. This represents a major advantage for vertically integrated players because they can leverage high volume manufacturing that the single market cameras specialists cannot.

1.2 Carbon Nanotube Based Infrared Sensor

1.2.1 Sensors Characterization

The major IR detector performance criteria indicating infrared detector performance are operating temperature, photo sensitivity, noise equivalent power (NEP) and detectivity.

- Photon sensitivity (Responsivity)

When noise is not a main consideration, the photon sensitivity can be calculated by the output (voltage or current) per watt of incident energy, show in Eq. (1.3).

$$R = \frac{S}{PA} \tag{1.3}$$

R: Responsivity, [V/W]

S: Signal Output, [V] or [A]

P: Incident energy, [W/cm^2]

A: Detector active area, [cm^2]

In photovoltaic infrared detectors, the output signals are extracted as photocurrent. It is expressed as Eq. (1.4), when the light is at a given wavelength irradiated on detector.

$$I_{SC} = \eta q \frac{PA}{h\frac{c}{\lambda}} = \frac{\eta q P A \lambda}{hc} \tag{1.4}$$

q: Electron charge, [C]

The responsivity of photovoltaic photodetector at specific wavelength (λ) will be simplified as Eq. (1.5).

$$R_\lambda = \frac{I_{SC}}{PA} = \frac{\eta q \lambda}{hc} = \frac{\eta \lambda}{1.24} \tag{1.5}$$

η : Quantum efficiency

However, the output of photoconductive detector is voltage. The output voltage V$_O$ will vary (ΔV$_O$) due to changes (ΔR$_i$) of internal resistance R$_i$ when exposing to infrared light.

$$V_O = \frac{R_L}{R_i + R_L} \cdot V_B \tag{1.6}$$

$$\Delta V_O = -\frac{R_L V_B}{R_i + R_L}^{2} \cdot \Delta R_i \tag{1.7}$$

$$\Delta R_i = -R_i \frac{q(\mu_e + \mu_h)}{\sigma} \frac{\eta \tau \lambda PA}{lwdhc} \tag{1.8}$$

τ: Carrier lifetime [S]

μ_e: Electron mobility [cm^2/(V · s)]

μ_h: Hole mobility [cm^2/(V·s)]

σ: Electrical conductance [S/m]

$$\Delta R_\lambda = \frac{\Delta V_O}{PA} = \frac{q(\mu_e + \mu_h)\lambda \tau \eta}{\sigma lwdhc} \cdot \frac{R_L R_i V_B}{(R_i + R_L)^2} \tag{1.9}$$

Although the photon sensitivity is so complicated, it can be expressed by Eq. (1.9). It is only applicable to few cases because R_i, μ_e, μ_h, τ and σ are dependent with each other. Moveover, it is also related to bias voltage (V_B) applied on the detector.

- Noise equivalent power: NEP

Noise equivalent power (NEP) is another critical value to measure the sensitivity of a photodetector. It is defined as the signal power that gives a signal-to-noise ratio of one in a one hertz output bandwidth [20]. The NEP (W/Hz$^{1/2}$) measures the quantity of incident light when the signal to nosie ratio (S/N) is one.

$$NEP = \frac{PA}{S/N \cdot \sqrt{\Delta f}} \tag{1.10}$$

N: Noise ouput, [V]

Δf: Noise bandwidth, [Hz]

The smaller NEP is, the photodetector will be more sensitive. For example, a detector with an NEP of $10^{-12}\mathrm{W}/\sqrt{\mathrm{Hz}}$ can detect a signal power of one picowatt with a Signal-to-Noise Ratio (SNR) of one after one half second of averaging. The SNR improves as the square root of the averaging time, and hence the SNR in this example can be improved to 10 by averaging for 50 seconds. Eq. (1.10) only refers to the electrical NEP. There is another NEP related to the detector system, called optical DEP. The optical NEP is equal to the electrical NEP divided by the optical coupling efficiency of the detector system.

- Detectivity: D^* (D-star)

Detectivity is the photon sensitivity per unit active area of a photodetector. This is widely used to compare the characteristics of different detectors [21]. The detectivity is given in Eq. (1.11), where A is the area of the photosensitive region of the detector, f is the frequency bandwidth [22]. When measuring D^* in experiments, it is related to temperature (T:[K]), wavelength of a radiant source (λ : [μm]) and the chopping frequency (f: [Hz]). Based on experimental results reported, the detector always has a peak sensitivity wavelength.

$$D^* = \frac{\sqrt{Af}}{NEP} \tag{1.11}$$

- Operating temperature

There are two sensing mechanisms for semiconductor bulk photon detector: the majority carrier and the minority carrier [23]. The sensing will be photoconductive in nature if it is majority carrier dominant, while it is both photoconductive and photovoltaic modes if the minority carrier dominates device output. Both carrier mobility and thermal noise are temperature dependent so that operating temperature is critical in characterizing of IR sensor.

In practically, it is necessary to consider wavelength, response time, temperature, cooling method, sensing area, number of sensing elements for infrared application.

1.2.2 Nano Material IR Photodetector

Carbon nanotubes have brought extensively attention both theoretically and experimentally since its discovery so that there are many nano electronic devices and NEMS using single or multiwall carbon nanotubes [24]. The ultra high surface-area-to-volume ratio and quasi 1D near ballistic electronic transport make it attractive in super capacitors [25], solar cells [26], mechanical oscillators [27], gas sensors [28]. Carbon nanotube based MEMS also drew increased scientific attention in single CNT, CNT film [24].

The observation of photoelectric effect in CNT has opened a number of avenues of research in both characterization and photonic applications of carbon nanotubes [29]. Infrared detection using CNTs was first realized and reported in [30] and [31]. The CNT based IR sensors, including individual single wall carbon nanotube (SWNT) based Schottky diode structure [32] [33] [34]and CNT field effect transistor (CNTFET) modulated structure [35] were also reported.

CNT is one dimensional nano materials with hexagon hollow cylinder structures, which shows outstanding mechanical, electrical and optoelectronic properties [36]. With development of more than 20 years, the theoretical analysis and potential applications are found. Depending on its chirality, carbon nanotubes have armchair, zigzag and chiral structure [37], which are classified as semiconductor and metal material . The bandgap energy of semiconductor CNT can be modulated by controlling the diameter of CNT in order to detect different wavelength of IR light [38]. The electron transport in CNT is confined in one dimension so that the thermal noise of CNT IR detector is extremely low due to phonon scattering suppression [39]. This creates the CNT device with stable response in room temperature.

The distinct properties of nano material distinguish it from traditional bulk materials. There are many researches turning to novel 1D, 2D materials and its derivatives. The carbon nanotubes are very good thermal conductors with ballistic conduction, so that it is used as a ultra small scale temperature sensor [40], which makes the sensor performance better than other devices in similar size. The resulting device exploits the superior thermal and electrical properties by deriving the temperature based on a change in electrical resistance [41]. Carbon nanotube not only works for thermal type but also quantum type infrared detector. There are many contributions on single wall carbon nanotube based infrared detector, in which the carbon nanotube is characterized as semiconductor materials with bandgap within near infrared wavelength. The second infrared detectors are demostrated using MWCNT. In [42] , different morphologies of MWCNT are synthesized to detect infrared radiation in room temperature via measuring photoconductance. The third is CNT films based IR photodetector. Most experimental data of these detectors suggest that the IR photoresponse arises mainly from the thermal effect, as in [43] [30]. It was also reported that photoexcitation effect predominated the IR photoresponse in CNT film at [44]. All of these infrared detectors can work in room temperature with fast responsivity in small bandgap.

1.3 Computational Imaging

Computational imaging (Photography) refers to digital image capturing and processing techniques which use digital computational method instead of optical processes. The goal is to overcome the limitations of traditional photography and enhance the way of capturing, manipulating, and interacting with visual media. As shown in Figure 1.2, the traditional film/array camera is mimics human eye for a single snapshot, single view, single instant, fixed dynamic range and depth of field for given illumination in a static world. The camera comes from basic geometry optics to form image. However, the computational image use one more process to recovery image, as shown in Figure 1.3. Computational photography which enhances or extends the capabilities of digital imaging, is one of most rapidly developing research field in computer vision, image processing and applied optics [45]. The output of these techniques can reconstruct information of scene which is not obtained by today's digital camera [46]. The current research has evolved many fields. First of such techniques is computational imaging, including high dynamic range imaging (HDR), light field imaging, color management, etc. [45]. The high dynamic range imaging in [47], is achieved by placing an optical mask adjacent to image sensor array, followed by an efficient image reconstruction algorithms. More reconstructions algorithm on HDR image are in [48] [49]. Light field imaging [50] [51] is used to analyze image parts that are not in focus and extract depth information by ray tracing. The novel computational photography involves optics coded exposure imaging, which inserts a patterned coded aperture to recover both depth information and an all focus image from single photographs. Compressive coded aperture [52] is combined in supper resolution image reconstruction from low resolution [53]. Another novel computational photography, which is based on new mathematical theory and algorithms of compressive sensing, combines sampling with compression into a single non-adaptive linear measurement process, named single pixel imaging [54]. The single pixel camera utilizes Spatial Light Modulator (SLM), comprised of millions of electrostatically actuated micromirrors, to project target image into a low dimension. The 2-D mirrors work as optical switch with two states ON and OFF, so that the single pixel camera measures the inner product between an $M \times N$-pixel image and two dimensional functions in matrix. It has been proved by Emmanuel Candes, Terence Tao and David Donoho [55] [56] [57] that the sparsity signal may be reconstructed with fewer samples than Nyquist Shannon theorem requires. Extra constraints are imposed so as to get a unique solution in underdetermined system. Since compressive sensing is Non-deterministic Polynomial time hard problem (NP hard), one approach is directly to use greedy selection algorithms. Because of its non-convex, there is no guarantee to find global minimizer and the solution is unreliable, though it is fast in this approach.

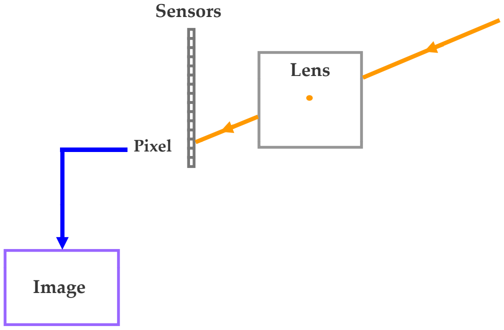


Figure 1.2 Conventional digital photography.

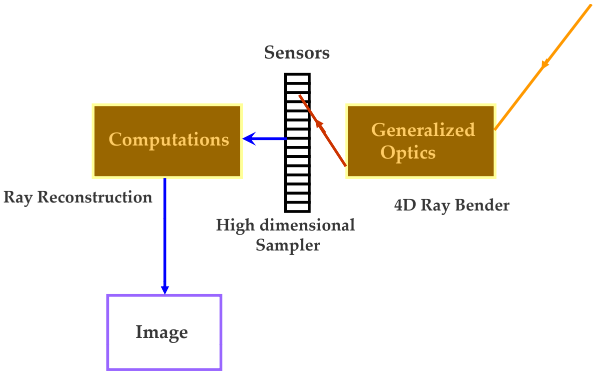


Figure 1.3 Computational photography.

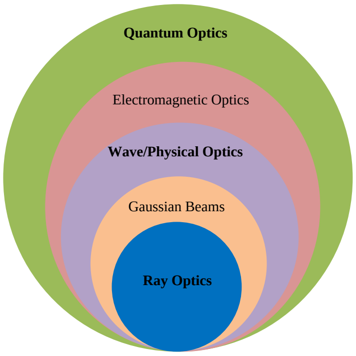


Figure 1.4 Optics classification.

1.4 High Dimensional Plenoptic Function of Light

The light behaviour has been investigated when Michael Faraday proposed that the light should be interpreted as a magnetic fields. The radiometry describes how energy is transferred from light sources to surface patches. In pervious research, the study of light can be divided as quantum optics, physical optics and geometry optics from micro to macro scale. The inclusion relation is shown in Figure 1.4. From the research area point of view, the quantum optics is the study of the interaction of object with light. If the object is in sub-wavelength or nanometer scale, it is referred to nanophotonics [58]. The physical optics's topics include interference, diffraction, polarization etc., which discuss physical theory. The abstract geometrical optics are also named ray optics which use approximation method to describe the light in terms of rays. The research will concentrate on how to mathematically express rays and get images by ray optics.

In ray optics, the plenoptic function was proposed for light field since 1991 by Adelson and Bergen [59]. In order to describe the nature luminous enviornment, there are seven parameters high dimensional model (7-Dimensional function) considered, shown in Eq. (1.12) (polar coordinates), Eq. (1.13) (Cartesian coordinates), where $V = (V_x, V_y, V_z)$ is the view point, $S = (\theta, \phi)/S = (x, y)$ is the direction of the ray light passing through the view point.

$$P = P(\theta, \phi, \lambda, t, V_x, V_y, V_z) \tag{1.12}$$

$$P = P(x, y, \lambda, t, V_x, V_y, V_z) \tag{1.13}$$

For a gray scale photograph taken by a pinhole camera, it only shows the intensity of light from a single view point in static. The wavelength is averaged over the spectrum, shown in Figure 1.5 left. The grayscale image can be parametreized by Eq. (1.14) or Eq. (1.15) in two distinct coordinates. A color image will add wavelength information to make it as a function of wavelength ($P(\theta, \phi, \lambda)$), as Figure 1.5 right. The color video stream will extend the information to cover time variable($P(\theta, \phi, \lambda, t)$). Moreover, the holographic movie would reconstruct of scene from every viewpoint such that seven parameters model is required. In order to make the model applicable, the assumption includes of rays passing through free space, the regions free of occluders such as opaque objects and the light traveling via a ray in constant along its length [60]. The 6D light field will be discussed in this research, as Eq. (1.16).

$$P = P(\theta, \phi) \tag{1.14}$$

$$P = P(x, y, \lambda) \tag{1.15}$$

$$P = P(x, y, \lambda, t, u, v) \tag{1.16}$$

1.5 Dissertation Overview

The central contribution of this dissertation is to design a reliable low dimensional material based IR sensor (CNT) and imaging system. A general overview of the research is discussed in section

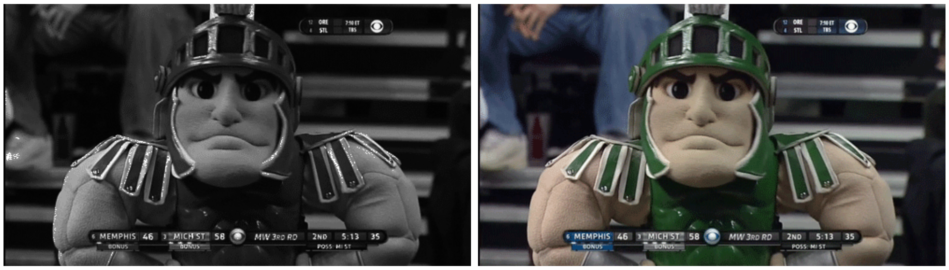


Figure 1.5 $P(x,y)$ and $P(x,y,\lambda)$.

of introduction. The following are six themes including nano material IR sensor, computational camera (single pixel IR camera), and four beyond IR camera systems, as shown in Figure 1.6.

1.6 Organization of the Study

The reliable CNT IR photodiode design, fabrication and testing method are presented in Chapter 2, including novel sensor structure design, an efficient sensor fabrication method and characterization of nanoscale IR sensor. It is followed by developing low dimensional material based IR camera, which comprised of compressive sensing and single IR sensor in Chapter 3. The camera readout system was also discussed. In Chapter 4, the compressive light field sensing method for IR spectrum is introduced, which makes IR light field sensing available. In Chapter 5, 3-D imaging method is described to explore more of scene in IR wavelength. In Chapter 6, the super resolution from multiple images to single image and its importance to single pixel imaging system are discussed, which open the door of nanosensor IR imaging system to commercialization. In Chapter 7, the nonconvex compressive video sensing and its implementation method was proposed to discover more IR messages in temporal domain. Finally, Chapter 8 presents the summary of research and future works.

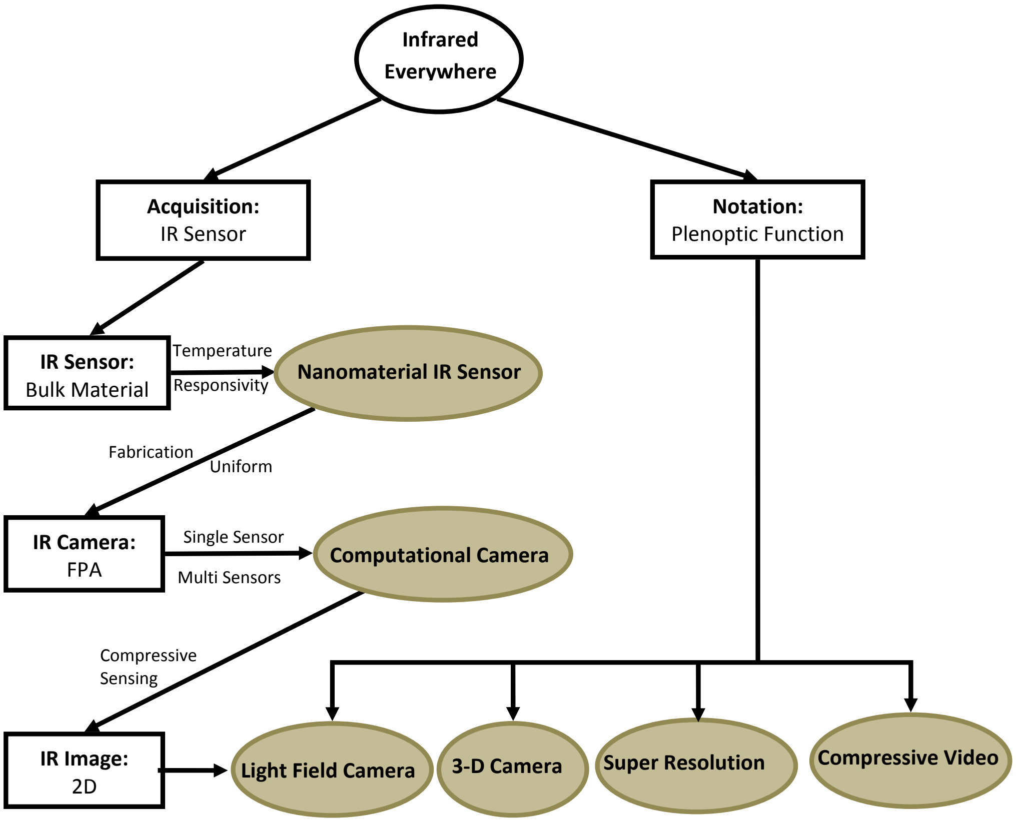


Figure 1.6 Dissertation overview.

Chapter 2

Low Dimensional Material Based Infrared Photodetector

2.1 Previous Work

Although nanotechnology had attracted a huge number of attentions since carbon nanotube (CNT) was firstly synthesized and reported by Sumio Iijima in 1991 at NEC [61], and the optical absorption spectra of the SWNTs were observed from visible to near infrared region [62], a high performance CNT based infrared detector was still unavailable. Over the past several years, numerous studies have been performed in the field of carbon nanotube based photodetectors. In all, it can be classified as single wall CNT, multi-wall CNTs and CNT films. For single wall CNT, it has armchair, zigzag and chiral structure depending on its chirality [63], which are classified as semiconductor and metal material. In order to differentiate wavelength of IR light, the bandgap energy of semiconductor CNT can be modulated by controlling the diameter of CNT [38]. The electron transport in CNT is confined in one dimension so that the thermal noise of CNT IR detector will be extremely low due to phonon scattering suppression [39]. This creates the CNT device with stable response in room temperature. So far, nearly all CNT based IR photodetectors sensing scheme are based on CNT metal junctions for photon electron transition [32]. The Schottky barriers, shown in Figure 2.1, will be formed at the contact regions between metal and semiconductor CNT. When IR light irradiates onto contact region, photon induced electrons will be injected from CNT to metal surface, to generate photocurrent in close circuitry.

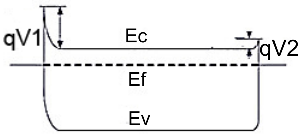


Figure 2.1 Band diagram of CNT metal contact.

However, CNT and metal contact has negative effect for photoconductance based IR photodetector [42] [44] so that the small contact resistance is preferred. The experimental results suggested that Pd and Ti contacts were superior to Au and Pt contacts, but the results for Ti were erratic, possibly due to the high chemical reactivity of Ti compared to other metals [64]. The calculations also demonstrated that there was no Schottky barrier to electron transfer between Pd and nanotube at the interface, because interface states, due to the charge transfer at the Pd/CNT contact, fills the band gap of the semiconducting CNT, resulting in a contact of metallic nature [65]. In order to reduce contact resistance in CNT-FETs, the use of a graphitic carbon (G-C) interfacial layer to semiconducting CNT can improve the electrical contact to the semiconducting CNT and reduce the subthreshold swing of transistors with these improved contacts [66]. Not only on the relationship between contact and photoresponse, it is known that the one-dimensional geometry of CNTs makes them highly sensitive to their electrostatic and electrochemical environment [67]. In [67], it also demonstrates that an electrochemical charge transfer reaction is the underlying phenomenon governing the suppression of electron conduction in CNTs devices. Besides, the device is also sensitive to electrical field. As shown in Figure 2.2, the “off”-state and “on”-state represent two different gate voltage conditions, which generate distinct band diagram in CNT-FET Schottky barrier. This kind of band bending affects the flow of electrons from source to drain and alters device properties.

Unfortunately, CNT metal contact sensing methodology suffers from limited sensing area and

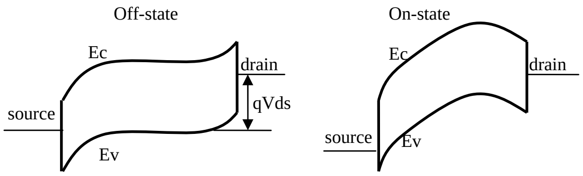


Figure 2.2 Band bending of a Schottky barrier in CNT-FET, for two gate voltages.

weak optical absorption. The fully explanation of how nanoscale material respond to infrared spectrum is critical in low dimension materials research. In this work, it is not only for the fundamental understanding of CNT photoinduced processes, but also for development of a new infrared photosensitive material with unique optical and electrical features.

2.2 CNT IR Sensor Design

As discussed in previous section, there are two kinds photoresponse mechanisms of CNT based infrared photodetector. They are quantum photovoltaic and photoconductance effect respectively, in which the single CNT photodetector happens on quantum photovoltaic while the CNT film or arrays based IR sensors depend on the photoconductance. In different CNT devices, the CNTs work as distinct functions. However, the CNTs-metal contact is most widely used structure for nano electrical circuits [68] and nano sensors [69]. In this research, a single CNT is preferred due to its unique performance so that CNT-metal contact is investigated more deeply. There are two types of CNT-metal contact area in most researches, the radial direction side contact and sidewall contact.

As shown in Figure 2.3, there are three locations of CNT in the side contact. The CNT-metal contact can be formed at bottom side (Figure 2.3 a), top side (Figure 2.3 b), and middle contact (Figure 2.3 c). The CNT on top of metal (bottom contact) can be manipulated by growth directly and Dielectrophoretic (DEP) assembly. However, the DEP method is not suitable for CNT under

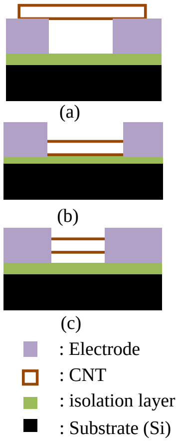


Figure 2.3 CNT metal contact with three distinct positions. a) CNT on the top metal; b) CNT under metal; c) CNT between metal.

metal. In the proposed application, CNTs are integrated into infrared optical sensor so that the Figure 2.3 (a) structure is more photoelectric conversion efficiency and the contact area will be transparent to IR irradiations. In this design, there is no top metal to block infrared light and the IR penetration depth on CNT can reach 450 *nm* [70] to increase efficiency.

2.3 Realtime DEP Fabrication

The techniques for manufacturing CNT based nano devices can be generally classified into bottom up and top-down two distinct methods [36]. In the first, CNTs are grown directly onto device substrate using chemical vapor deposition method [71]. With the directly growth method, single CNT devices are fabricated by growing a single CNT between a pair of prefabricated microelectrodes to make connections. The directly growth method is able to fabricate multiple single CNT devices at once. Thus it is good for making CNT based nano device arrays. However, the limitation of this method is that the properties of the CNTs can not be effectively controlled. Different CNTs may have distinct properties even they are produced at one batch based on its chirality [72]. Moreover, the production process may generate impurities around the microelectrodes and CNTs, which will affect the electronic properties of the CNT device. Thirdly, it is difficult to grow only a single CNT between the microelectrodes while they are bundles or films, losing the unique properties brought by the 1-D structure of CNTs. The most disadvantage of directly growth method comes from CVD process in very high temperature [73]. This limits the substrates such as silicon sapphire so that it could not widely work for flexible materials and some biocompatible MEMS applications.

The second category is to grow, purify and sort CNTs firstly, and align CNTs using assembly manipulation. There are two ways to manipulate CNTs including DEP manipulation [74] and nano probe based mechanical manipulation [75]. With the DEP deposition method, microelectrodes are first fabricated using conventional microfabrication. During localizing CNT process, a droplet of CNT suspension (appropriate concentration in ethanol) is dropped between the microelectrodes and an AC voltage is applied across the microelectrodes. CNTs will be attracted by the dielectrophoretic force and bridged on the electrodes to form an electrical connection. Although the number of CNTs attracted to the electrodes can be roughly controlled by varying the AC voltage and the concentration of the CNT solution, it is very difficult to deposit a single CNT using this method. Hence the DEP deposition method is normally used to fabricate devices with CNT films or CNT bundles. The nano probe based mechanical manipulation was also called nanorobot manipulation [75]. In [75], a single CNT attached at the tip end was manipulated using focused ion beam. In this method, the nanotube has to be metal-coated for manipulation and not good for building nanoelectronic devices.

DEP based assembly is potentially one of most important bottom-up technologies for fabricating nano materials based MEMS and NEMS blocks. It is liquid medium based method to transport micro/nano particles at room temperature. However, CNTs have very low dispersibility in many solvents due to its high Van der waals forces between each nanotube, inducing strong tendency to aggregate with each other. The most difficulty in controlling nanotube deposition using DEP is that a large number of tubes or bundles of tubes will be accumulated between gap so that a real-time controlled system is required to monitor the quantity. In [76], a conductance measuring was used to estimate a large number of multiwall carbon nanotubes with small impedance but not applicable to single wall CNTs. The approach in [77], used an in situ detection system and lock-in amplifier to read DC and AC current through DEP loop in order to align single wall CNT between gap. Due to the large impedance of SWCNT, the current is such weak between electrodes that the lock in amplifier needs time to integrate. Therefore, the quantity is out of control in this point.

The whole DEP based CNT assembly process has millions applications from electrical device to bio-sensors. It includes growth of CNT, deposition of as-grown CNTs on electrodes by dielectrophoresis (DEP). In bio-application, the extra process is to make self-assembly by functionalizing CNTs with different chemicals or even DNA molecules [78]. To a certain extent, these methods have their shortcomings in terms of repeatability, mass production. In generally, the both methods usually come together in fabrication process of CNT based device, where top-down methods are used to fabricate supporting structures such as contacting electrodes, and bottom-up methods are used to assemble CNTs and localize it onto desired position.

2.3.1 Introduction of DEP

The dielectrophoresis force originates from Pohl's theory and DEP forces F_{DEP} and torques T_{DEP} can be calculated [79] as Eq. (2.1) and Eq. (2.2). As shown in Figure 2.4, the nano particle is dispersed in medium and an electrical force will be applied on particle when an AC voltage is ON.

$$\mathbf{F}_{DEP} = (p \cdot \nabla)\mathbf{E} \tag{2.1}$$

$$\mathbf{T}_{DEP} = p \times \mathbf{E} \tag{2.2}$$

In Eq. (2.1) and Eq. (2.2), p is the induced dipole moment of the nano particle and $\mathbf{E}$ is the non uniform electric field applied on electrodes. The DEP forces can be simplified as Eq. (2.3) and Eq. (2.4) in general, where v_p is the volume of the particle and $\nabla|E|^2$ is the root-mean-square of the applied electric field.

$$\mathbf{F}_{DEP} = \frac{1}{2} v_p Re(f_{CM}) \nabla |\mathbf{E}|^2 \tag{2.3}$$

$$f_{CM} = \frac{\varepsilon_p^* - \varepsilon_m^*}{\varepsilon_p^* + 2\varepsilon_p^*} \tag{2.4}$$

Based on this equation, the direction of the DEP force is determined by the real part of f_{CM}. In Eq. (2.4), ε_m^* denotes the complex electrical permittivity of the liquid medium and ε_p^* is the complex electrical permittivity of nano particle. f_{CM} is the Clausius-Mossotti factor, which indicates whether the medium or the particle is more polarizable. When f_{CM} is larger 0, it is called positive DEP force, resulting the particle moving towards microelectrodes (high electric field region). When f_{CM} is less than 0, it is called negative DEP force, resulting the particle is moving away from high electric field region. The electrohydrodynamics of CNTs is not only related to DEP force, but also influenced by the effect of the fluid exerted on the particle. The viscous drag force is proportional to the relative velocity for a prolate ellipsoidal CNT.

The velocity dynamics in a fluid medium environment can be expressed by Newton's second law for a particle with mass m, as shown in Eq. (2.5). It only considers DEP force, viscous force and relative velocity $(u-v)$. The constant f shown in Eq. (2.6), is the translation friction factor which depends on size, shape, and fluid viscosity η. Recalling Perrin friction factors and further developed hydrodynamic approaches by Hardings and Small [80] [81], the friction factor moving at random is generalized shown in Eq. (2.7).

$$m\frac{dv}{dt} = F_{DEP} + f * (u - v) \tag{2.5}$$

$$\langle f \rangle = \frac{3\pi\eta l}{ln(l/r)} \tag{2.6}$$

$$v = (\frac{\mathbf{F}_{DEP}}{f} + u)(1 - e^{-(f/m)t}) \tag{2.7}$$

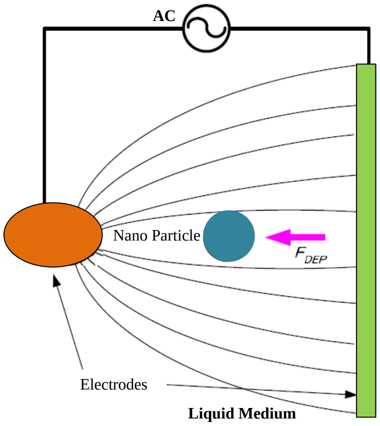


Figure 2.4 Illustration of the dielectrophoretic manipulation.

2.3.2 Assembly Method and System

DEP assembly refers to the manipulation of micro/nano particles using dielectrophoretic force, which exerts on polarized dielectric particles [82], when the particles present in a non-uniform electric field with distinct dielectric constant and polarizability to surrounding medium. As shown in Figure 2.5, the non-uniform electrical field is generated between symmetric electrodes with round shape, (SEM image shown in Figure 2.7) and carbon nanotubes have different dielectric constant to surrounding medium (Ethanol). There is a DEP force applied to CNTs when turning on external AC voltage, which will bridge CNT between gap [83]. In this novel system, an isolation method was proposed to measure impedance changes within DEP system loop. As shown in Figure 2.6, a customized transformer structure was used to read di/dt value to detect current changes. When a new nanotube is bridged, a sudden small change of impedance introduces a tremendous large di/dt, which will be reflected on both sides of transformer. The integrated weak signal detection module will easily read current change and send a feedback to turn AC off. Based on this method, the number of CNTs trapped by DEP is quantitatively identified. Compared to the method of reading electrical current directly, the proposed indirect approach responses much faster since it does not read absolute weak current value in DEP system.

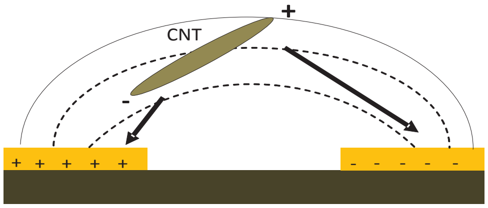


Figure 2.5 DEP force on CNT in a non-uniform electrical field (side view).

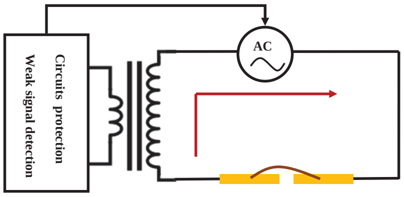


Figure 2.6 Real-time monitoring DEP system. Red row shows current loop when CNT is bridged between gap (yellow). The system will shut down AC source through feedback when impedance changes.

2.3.3 Quantitatively CNT Deposition and Device Fabrication

In DEP process, the Au electrodes were fabricated using Electron Beam Lithography (EBL) and lift-off technique to get 1 μm width and 1 μm gap on substrate. The CNT powders (Buckyusa) were dispersed in liquid medium (UN1170 Ethyl Alcohol) for 1.5 hour ultrasound. In the deposition process, a droplet of the CNT suspension will be dropped onto finished substrate and an AC voltage of 1.5 V peak-to-peak with frequency of 1 kHz is applied. When AC voltage is ON, the weak signal detection will also start to read current change. If the peak is larger than threshold, the AC voltage will be turned off automatically.

There is a critical treatment for CNT after AC voltage off. Since the bridged CNT is still in alcohol medium, the volatile alcohol will remove CNT out from desired position due to surface tension when the medium evaporates. The extra treatment for CNT is to merge the sample into DI water before alcohol dries. Although CNT is dissolved into alcohol, the surface of CNT is hydrophobic. When the sample is immersed in DI water, the residual medium (Alcohol) will be diluted. After one minute, the DI water is isotropic to the hydrophobic CNT so that the aligned CNT will stay on the original position when it is taken out of DI water.

There are two groups of experiments discussed. Figure 2.7 and Figure 2.8 show the SEM image of MWCNT and SWCNT aligned between gaps using proposed real-time monitoring. The weak signal detection system will count pulse generated by number of CNTs aligned. The results show one, two, three multiple wall CNTs deposition in Figure 2.7 and nanotubes can be precisely located. In this experimental setup, the CNT deposition yield is very high which reach 100% (20/20) for single CNT deposition. The two CNTs alignment experiment has 90% (18/20) device yield ratio. Figure 2.8 shows single wall carbon nanotube deposition from quantitatively control, which includes one nanotube and a CNT film deposition in localized area. Although the single

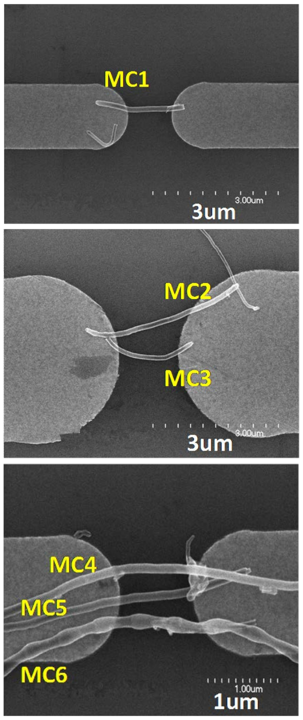


Figure 2.7 SEM image of multiwall carbon nanotubes. There is only one CNT (MC1) on top device, two CNTs (MC2 and MC3) on middle device and three CNTs on bottom device (MC4, MC5 and MC6).

wall CNT deposition is a little lower yield (15/20), it is possible to improve by reducing solution concentration. The thickness of CNT film is controllable, as shown in Figure 2.8 bottom.

2.4 Sensor Reliability and Response Enhancement

The nanomaterials are either diameter or thickness between 1 and 1000 *nm* [71]. Due to such small dimension, not only the material properties change, but also the sensing structure and mechanisms are distinct from conventional buck materials. On the other hand, nano structures such as nano tube, nanorods, nanobelts and nano fibres possess high surface to volume ratio, large penetration depth and fast charge diffusion rate, which are sensitive to gas such as H_2, CO, NO_2 and volatile organic compounds. Meanwhile, electrical charges trapped under CNT will also change

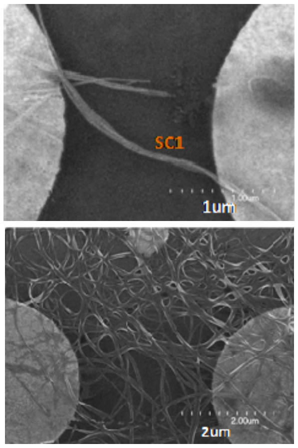


Figure 2.8 SEM image of single wall carbon nanotube. Top is single CNT (SC1) bridged. The bottom device shows single wall CNT film using real time DEP deposition.

device performance in nano electronics [84]. It is experimentally proved that the one dimensional geometry of CNTs makes them highly sensitive to their electrostatic and electrochemical environment [67]. In [67], it also demonstrates that an electrochemical charge transfer reaction is the underlying phenomenon governing the suppression of electron conduction in CNTs devices. From these points, the design of CNT based infrared detector will be more complicated than general bulk semiconductor materials.

2.4.1 Substrate Effect and Packaging on Nano Sensor

Since the interfacial/isolation layer on substrate (shown in Figure 2.3) has significant effect on IR photoresponse by their electrostatic and electrochemical environment, there are four distinct isolation layers deposited in experiments respectively, including SiO_2 (300 nm), Si_3N_4(180 nm), quartz (500 μm), parylene-C(4.5 μm), polyimide (10 μm) polymer. SiO_2 and Si_3N_4 are grown on silicon from University Wafer ($http://www.universitywafer.com/$). The substrate is p type (100), 0.01-0.02 ohm-cm.

The stability and reliability are two critical parameters to characterize nanoelectronics device.

It was found that the stability of the CNT device was affected by oxygen contamination. The exposure of single-walled CNT (SWCNT) samples and devices to oxygen appeared to have a strong influence on their electronic transport properties as reported previously [85]. In order to get a stable IR response of CNT sensor, the package layer using parylene-C thin film was deposited on top of CNT-metal contact so as for oxygen barrier.

The parylene-C is deposited by PDS 2010 parylene coater. The thickness of parylene was a critical parameter in the device because too thin layer will not isolate CNT from surrounding environment and too thick layer will have much absorption of IR irradiance. In experiments, 1 μm parylene was coated on CNT IR sensor using PDS2000 parylene coating.

In order to compare device performance, all measurements were performed using the same infrared source (100 mW 830 nm), and all date were collected by Agilent 4156C precise semiconductor parameter analyser in the room temperature (25°C). Figure 2.9 shows the darkcurrent in five IR sensors with different interfacial layer. It was larger than 1.75 nA on SiO_2 (as SiO2) interfacial layer while it was around 0.6 nA on quartz (as Quartz). However, the dark current was reduced to less than 0.3 nA on parylene-C (as Parylene), Si_3N_4 (as SiN) and polyimide (as Polyimide). The polymer can isolate oxygen-CNT contact under the Schottky barrier, especially the thicker polyimide will reduce the darkcurrent to 0.1 nA level. Meanwhile, the linearity measurements were shown in Figure 2.10. Five different sensors have very good linearity when IR irradiance power linearly increases. The results also show that the sensor with polyimide interfacial layer has largest response and the photocurrent is around 170 nA at 3000 mW/cm^2 IR irradiation. The SiO_2 layer has the worst response at around 30 nA photocurrent.

In CNT based nano electronic device, there is no exact function for electron transport although green functions are widely used to describe and simulate current. In the proposed CNT based IR sensor structure, the interfacial layer under CNT-metal contact has significant effects on electron

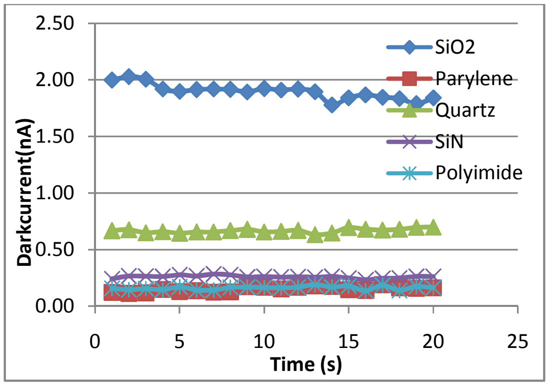


Figure 2.9 Dark current measurement results on CNT IR detector.

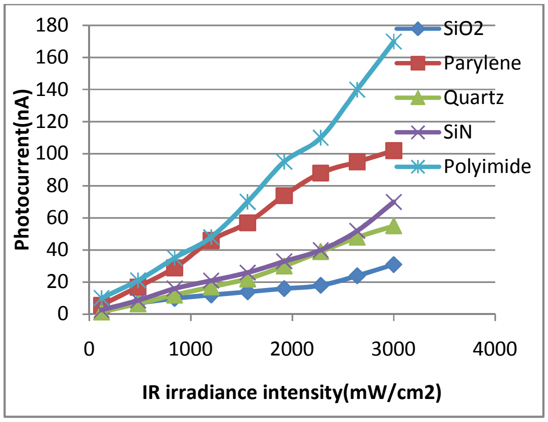


Figure 2.10 Linearity measurement results on CNT IR detector.

transport. The different interfacial layer under CNT-metal can change oxygen doping in CNT [67]. It is reported that the CNTs works as high p type semiconductor when fully exposing to oxygen. When the interfacial layer is changed, the chemical nature of CNT will be different so as to make p doping changes0.

In the experimental results, Si_3N_4 interfacial layer induced higher photocurrent than $Si0_2$. When $Si0_2$ was changed to Si_3N_4, there was no oxygen under CNT so that the contact barrier was high. Parylene and polyimide is also oxygen free. When CNT was deposited on parylene or polyimide, the electrochemical charge was less than on $Si0_2$, so that the CNT would be in low p type. Meanwhile, it increased the barrier height resulting a higher photocurrent.

2.4.2 Extrinsic Surface State Effect

In nanoscale material, surface state could also introduce extra electron energy state in CNT bandgap. The surface charges between CNT and metal, CNT and substrate are critical in the device, which are closely related to parasitic capacitance. The capacitance between two electrodes increases dramatically when gap decreases. This could cause the sensor behavior to depart from expected sensing performance. As shown in Figure 2.11, there are at least six parasitic capacitors in this device. C_s is the internal capacitance within CNT-metal Schottky barrier, which measures the build-in potential in CNT-metal contact. In Schottky diode, the depletion region is an insulator that separates the metal layer and doped semiconductor layer, forming a parallel plate capacitor C_s. The thickness of depletion layer can be modulated by the magnitude of externally applied voltage. C_1 and C_2 are substrate capacitors ($\ll C_s$). C_{sr} and C_{sl} are parasitic capacitors between electrode and substrate, which is source to deteriorate sensor response. From capacitance equation Eq. (2.8), the parasitic capacitance depends on the gap size between two electrodes and isolation materials underneath. Although C_1 and C_2 is larger than than C_i, it is coupling to ground and has no effect

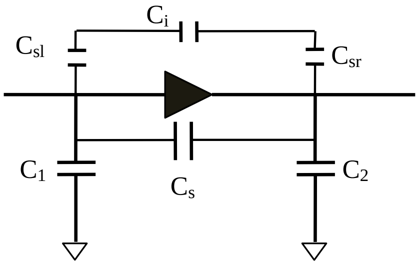


Figure 2.11 Parasitic capacitance model of CNT metal Schottky barrier.

on sensor response. ε is the dielectric constant of silicon dioxide and d represents the gap size. When gap size decreases, the C$_s$ increases dramatically for nanoscale sensor. This large capacitor stores more charges between two electrodes and changes electrical potential distribution along carbon nanotube. Moreover, it may reduce Schottky barrier height to make IR detector performance worse. In the proposed CNT IR device, a low-k material was deposited as interfacial layer to reduce charge distribution under carbon nanotube. On top of SiO$_2$ layer (150 *nm*) there was 10 μm polyimide spined on top as interfacial layer. The polyimide thickness and surface flatness is the key process. In this research, the polyimide (HD MicroSystems, Inc. PI-2555) was spined twice on silicon based wafer with the speed of 2000 *rpm*. After spin process, it was put in oven and cured the polyimide at 300°C for 2 hour with 5°C per minute starting from room temperature. The resulting polyimide thickness is around 10 μm.

$$C_i = \frac{\varepsilon * S}{d} \tag{2.8}$$

The capacitance measures charge storage ability between two isolation plates. It is obvious that the charge storage under sensor will change sensing performance. As shown in Figure 2.12, when metal atoms are deposited on isolation layer, there is positive charge layer formed on substrate. Based on this model and Figure 2.11, parasitic capacitors from C$_i$ and C$_{sr}$ or C$_{sl}$ will affect

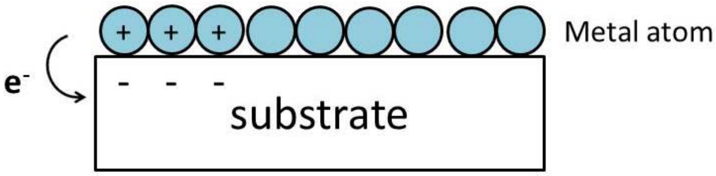


Figure 2.12 Surface charge storage on substrate.

charge redistribution in substrate. Meanwhile, electron work function difference between metal and substrate material will also contribute on charge distribution. The electron work function of Au is around 5.3 *eV* [86], while the energy bandgap of silicon dioxide is 9 eV, the polyimide is 4.32 eV for two layers [87]. This difference will generate charge accumulation on the surface, which contributes to surface voltage. The huger difference forms larger dynamic surface capacitance. In order to make infrared sensor stable and high responsivity, the polyimide contributes less charge stored on C_{sl} since the work function of Au and polyimide layer is less than other materials discussed.

2.4.3 Sensor Response Enhancement

Although the IR sensor response can be optimized by device structure, the fill factor of CNT-based IR sensors is still limited by low incoming electric field at their nanometer scale sensing area. The photoresponses of CNT-based sensors are relatively low. It is mainly because the detection methodology based on the CNT/metal schottky junction suffers from limited sensing area and weak optical absorption. There are many ways proposed to enhance the performance of the CNT-based IR sensors, of which the promising approach is use optical antennas to enhance the local electric field [88]. Helical antennas have a high gain over a broad band of frequency characteristics. The radiation along the helix axis is found to be the strongest when the circumference of the helix is of the order of one wavelength. To improve the fill factor of CNT-based IR sensor, the three-

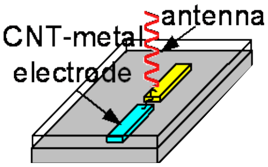


Figure 2.13 CNT-based IR sensor response enhancement by helical antenna.

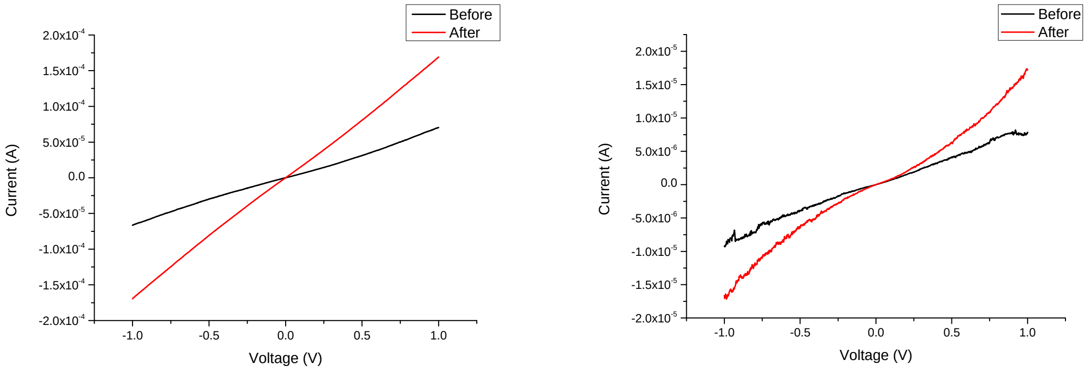


Figure 2.14 I-V curve of CNT IR Sensor. a) device A; b) device B.

dimensional helical structures with micro and nano-features are needed. In this research, the helical nanobelt (HNB) structure was fabricated from the InGaAs/GaAs by a top-down fabrication process in which a strained nanometer-thick heteroepitaxial bilayer curled up to form three-dimensional (3D) helical structure with nanoscale features [89]. It served as the optical antennas to improve the fill factor of the CNT-based IR sensors, which is schematically illustrated in Figure 2.13.

The electrical and photoresponse properties are both tested and discussed to fully characterize the performance of the integrated detector. The results of electrical property tests for both detectors are shown in Figure 2.13 and Figure 2.14, with the red line for the detector after integration and the black line for the bare one. It can be readily found that after the integration, the detector is more sensitive to the change of the electrical signal, and therefore the HNB antenna is of a better performance for the IR sensing from the perspective of electrical properties. Several factors could contribute to this effect. Firstly, the HNB works as a charged particle, which will generate a small electrical field around the CNT-metal contact and therefore changes the I-V relationship. On the other hand, due to the mechanical assembly of the HNB antenna, this can also be attributed to the change of the position of the CNT and/or the contact between the CNT and electrode.

2.5 Nanoscale IR Sensor Characterization

2.5.1 Sensors and Measurement Method

In CNT based infrared photodetectors, the most basic structure is CNT-metal schottky diode, in which the CNT is aligned between two electrodes by quantitatively controlled dielectrophoretic (DEP) assembly [90]. As shown in Figure 2.15 (top), the SEM image shows Au-CNT-Au structure and its dimension on SiO_2, where the CNT length is around 6 μm. There are three areas from left to right (L, M and R): left Au-CNT (1 μm), CNT connection (4 μm) and right CNT-Au (1 μm). In this Au-CNT-Au symmetry structure, both photoconductance and photovoltaic effect have possibility to lead photocurrent in device and the dominate effect depends on the CNT area and CNT-metal contact type, shown in Table 2.1. At left and right side, the Au-CNT/CNT-Au contact can introduce both photoconductance and photovoltaic effect for IR response while the center CNT (M area) only generate photoconductance. In order to certify the foremost photoresponse source, a reliable testing system is required to measure CNT IR sensor response. The characterization of nanoscale photodetector is a huge challenge due to optics diffraction limit [91] and the commercialized infrared laser beam spot size can only narrow to 10 μm scale which is thousand times of single wall CNT diameter [61]. In Figure 2.15 (bottom), it shows the relative size between laser beam (World Star Tech UH5-100G-808, 100 mW 808 nm single mode laser module) and CNT after fine focus, in which the bright white oval shape shows the laser beam and the center black dot represents the CNT under laser. The CNT length is around tenth of laser beam diameter. It is hardly to know the localization of CNT and laser beam in this configuration. Therefore,

Table 2.1 Au-CNT-Au structure and its photoresponse.

Side Name	Possible Photoresponse
Area L	Photoconductance, Photovoltaic
Area M	Photoconductance
Area R	Photoconductance, Photovoltaic

in [31] [92] [34], it just showed the maximum photocurrent with estimated input irradiance energy.

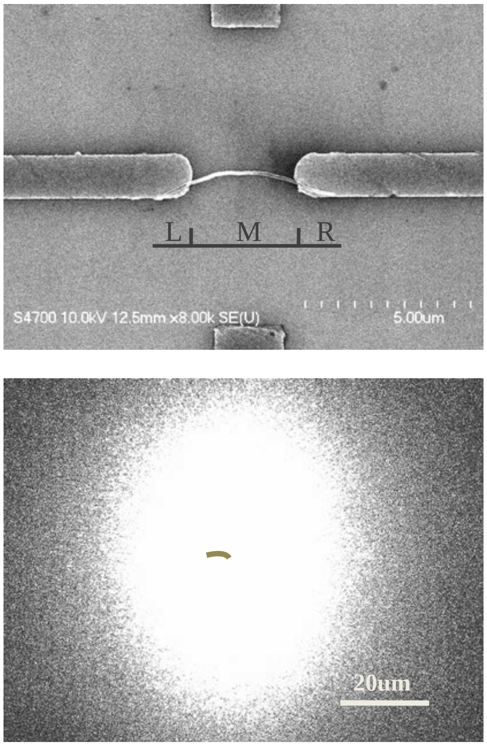


Figure 2.15 Top: SEM image of Au-CNT-Au structure. Bottom: The relative size between CNT detector and IR laser beam spot.

In the proposed testing bench (Figure 2.16(a) and Figure 2.16(b)), a digital microscope (Keyence VHX-600) and precise five axis substage (Kleindiek Eucentric Five Axis Table) are used to localize CNT photodetector. The long working distance of digital microscope leaves the space for IR irradiation on detector. Both microscope and laser are locked on non-movable position and the microscope monitors the height (*z* direction) and position (*x* and *y* direction) of CNT detector. The nanoscale CNT device is moved on precise substage and a charge integration readout circuit measures the current flowing in it. The inset at left corner of Figure 2.16 (a) shows the relative

Table 2.2 Eucentric five axis table specificiations.

Items	Specificiations
Dimensions	L:72 mm - W: 50 mm - H: 44 mm
Travel (x and y)	10 mm
Travel (z)	3 mm
Travel (R)	360 deg
Travel (T)	$\pm$ 90 deg
Absolute accuracy	T $<$ 0.2 deg
Repeatablity	T $<$ 0.03 deg
Linear resolution	$<$ 0.5 nm
Rotational resolution	$< 6\text{x}10^{-6}$ deg (10^{-7} rad)
Speed	up to 1 mm/s
Resolution	$<$ 0.5 nm

position of detector and laser spot on stage. Figure 2.15 (b) is more precise about the dimension. Figure 2.16 (b) is the experimental setup using laser, substage and digital microscope. Table 2.2 lists the specifications of eucentric five axis (x, y, z, R and T), in which the resolution, speed and travel range is accurate enough for repositioning device.

In order to leverage the detector in horizontal, there are four markers (perpendicular 'L' shape and its mirror image) designed on substrate, as shown in Figure 2.16 (c), M1-M4. During calibration process, the detector is moved by stage controller to focus marker respectively. This process is finished till four markers are in focus when it is moved to the center of microscope without changing z direction. In the measurement, the center point of digital microscope (via microscope display screen) should keep focus when the detector is moving with the substage on x-y plane. The measurement trajectory is around 1 μm step size, as shown in Figure 2.16 (d) and the center of photodetector will be moved on this pathway. The measurement process includes x-axis and y-axis scanning. In order to set the center focused corresponding point under the same IR irradiation, the focus point/line is always in the center of digital microscope no matter where the detector is.

As shown in Figure 2.17, there are three points selected from device (A, B and C) and three imaging point (A', B' and C') formed on image sensor after objective lens and tube lens. The clear focused points from microscope are in the same height. However, when the detector plane is not perpendicular with symmetry axis of microscope (point A_1 point C_1 light path), there is only one narrow line (point B_1) focused as shown in Figure 2.17 right. As in Figure 2.16 proposed measurement system, the digital microscope is not vertical on detector plane, there must be one line focused when moving the photodetector on *x*-*y* plane. By controlling this focus line in the middle of microscope image on *y* direction, the detector will be moved on the same horizontal level.

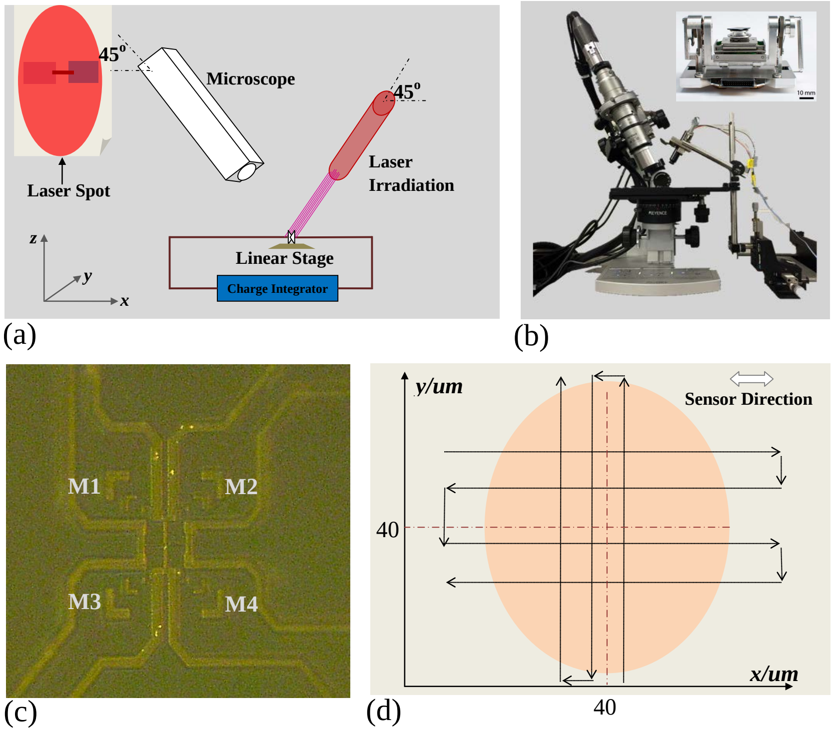


Figure 2.16 a) Proposed testing bench using digital microscope, laser and five axis substage. b) Hardware setup, inset is substage. c) Four points calibration marker for detector. d) Raster scanning: experimental measurement pathway for centroid of photodetector.

2.5.2 Experimental Results

In this section, CNT IR sensor response on *x*-axis, *y*-axis, bias voltage and contact length are discussed. The measurement is firstly conducted after aligning CNT photodetector along *x*-axis,

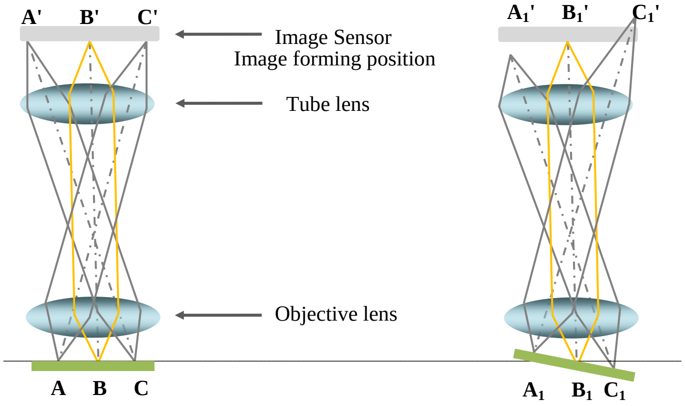


Figure 2.17 Focused and unfocused light rays on digital microscope.

in which the maximum response can reach to 28 *nA*, shown in Figure 2.18. The photocurrent decreases when moving the detector up and down (y >40 μm and y <40 μm). The photocurrent (I_p) relation on line y (= 20, 30, 40, 50, 60 μm) is ordered as Eq. (2.9) and Eq. (2.10) on its corresponding point. The line with maximum response is close on the center in y direction (y = 40 μm).

$$I_{p,y=40\mu m} > I_{p,y=30\mu m} > I_{p,y=20\mu m} \tag{2.9}$$

$$I_{p,y=40\mu m} > I_{p,y=50\mu m} > I_{p,y=60\mu m} \tag{2.10}$$

In Figure 2.18, there are three areas, including positive response area, negative response area and the unknown area between these two. The photocurrent is on positive direction, negative direction and unstable respectively. In overall, there are two opposite direction current sources in the device when IR irradiates on detector, because it generates positive and negative photocurrent with zero bias on it. This could be explained by two Schottky barrier formed by Au-CNT (left) and CNT-Au (right). When the detector is on left side of laser spot, the CNT-Au Schottky diode dominates the response, while the Au-CNT Schottky diode generates more photocurrent on right side. The maximum response does not happen in the center line but with a little offset on both side.

As shown in Figure 2.18, it locates around 32 μm at left and 55 μm at right half. Therefore, the photoconductance effect can not dominate the photoresponse because the maximum IR irradiation is on the center due to its single mode gaussian beam. The maximum photocurrent should be on center if photoconductance contributes the most.

In the measurement, the maximum photocurrent is generated on the large slope area of gaussian laser beam, where the two diodes will stay in distinct areas with large IR energy difference. The photodetector on higher power IR irradiation will output more current after neutralizing small part of charges with another side. In the unknown area between positive and negative response, the photocurrent is noisy and unstable. It jumps from positive side to negative side or reversely. The reason is that gaussian beam will produce nearly uniform output in center area. The two current in facing photodetectors will be canceled by each other so that it is very hard to generate a stable photocurrent. As shown in Figure 2.18, the more unknown area happens the sensor is closer to center, which is consistent with gaussian distribution.

Figure 2.19 shows the photoresponse along y direction. The photocurrent is positive when the detector is located on left side, while it generates negative response on right side. Meanwhile, the photocurrent decreases when the sensor goes further left (from $x = 33$ μm to $x = 30$ μm) or further right (from $x = 48$ μm to $x = 51 \mu$m). All the photoresonse are symmetry along $y = 40$ μm due to that the detector direction is perpendicular to moving trajectory and the laser beam output is symmetry along $y = 40$ μm. The photocurrent is proportional to the difference of IR irradiation between Au-CNT contact (left) and CNT-Au (right) contact. However, the photocurrent curve is not fully symmetry at $x = 32$ μm and $x = 33$ μm because the precise substage changes the position at 1 μm step size but with measurement position errors. If photoconductance dominates the photoresponse, each line should be gaussian curve as the result that the photocurrent directly reflects the laser output distribution on y direction.

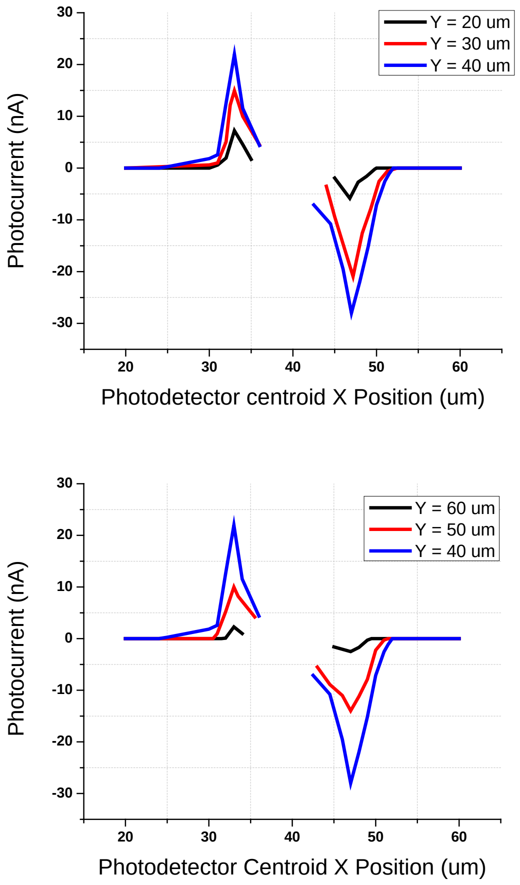


Figure 2.18 Photocurrent measurement along *x* direction with distinct *y*.

In infrared photodetector, the figure of merit includes NETD, responsivity etc [93]. Responsivity is referred to photosensitivity which is related to quantum efficiency (the number of electrons released per incident photon). When noise is not a main consideration, the photo sensitivity can be calculated by the output (voltage or current) per watt of incident energy, shown in Eq. (2.11), where R: Responsivity, [V/W], S: Signal Output, [V] or [A], P: Incident energy, [W/cm^2] and A: Detector active area, [cm^2]. The responsivity can be calculated by localizing CNT IR photodetector due to that the IR irradiance is gaussian distribution dependent on *x*-*y* position. In this research, the proposed detector maximum responsivity can reach to 16.8 μA/mW.

$$R = \frac{S}{PA} \tag{2.11}$$

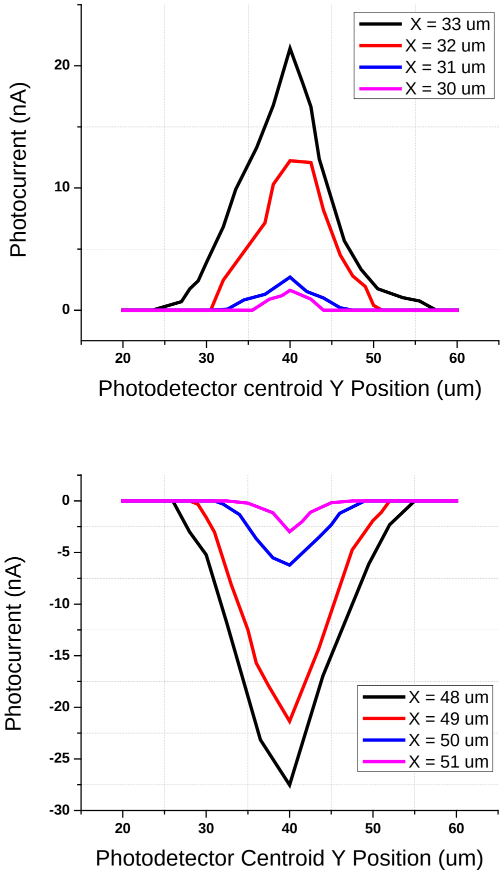


Figure 2.19 Photocurrent measurement along *y* direction with distinct *x*.

The nano device is very sensitive to external environmental changes, e.g. chemical [94], electrical signals [95]. In the CNT photodetector characterization, the bias voltage is sweep signal from -10 mV to 10 mV as shown in Figure 2.20 and the darkcurrent is from -160 nA to 160 nA, which is almost linear curve due to small voltage range. However, the photocurrent (maximum value) is always within 16.5±1 nA and the photoconductance has no effect on total current in sensor.

Since the Schottky barrier is critical in CNT based photodetector, the metal materials and con-

Table 2.3 CNT metal contact length and the direction of output photocurrent.

Left Length (μm)	Right Length (μm)	Photoresponse
8	1.5	Single direction
6	1.8	bi-direction
3.5	1.2	Single direction

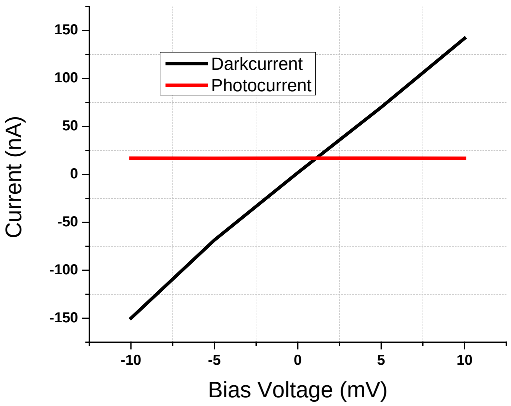


Figure 2.20 Photoresponse and darkcurrent on different bias voltage.

tact area will play a key role in output. The Cu and Ag metal CNT contact are also characterized. A group of ten Au-CNT-Au, Ag-CNT-Ag and Cu-CNT-Cu Schottky diode based IR photodetectors are measured with IR irradiation (100 mW 808 nm). As shown in Figure 2.21, the Au-CNT contact has more response than Ag and Cu in average due to its high work function (Au: 5.1 eV, Ag: 4.26 eV, Cu: 4.7 eV). The contact length is also characterized by different size, shown in Table 2.3. The first and third detector only respond to one side and the 6/1.8 μm detector has positive and negative current. The 1.8 μm side only generates around 1 nA scale photocurrent but 6 μm side produce about 40 nA response (All the measurements are from maximum point). The experimental results indicate that photovoltaic dominates photoresponse on CNT-metal Schottky detector although the photovoltaic voltage can not be sampled due to its tiny value merged in noise. The detector IR response are dependent on CNT-metal contact size and metal workfunction.

2.6 Chapter Summary

The stability and reliability of CNT based IR sensor were analyzed and presented fully in this study. In the first, by analyzing the parasitic capacitance in CNT metal Schottky barrier, a novel structure was proposed. The sensor on low-k polyimide interfacial layer has around 100 nA photocurrent at

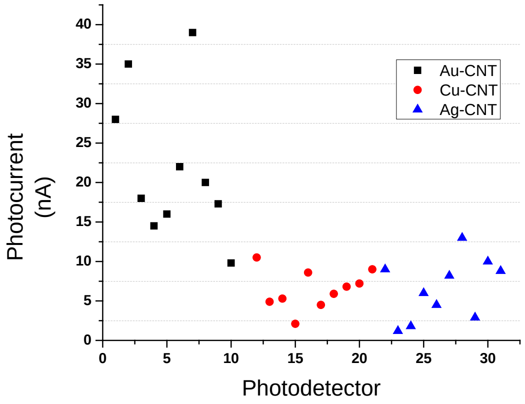


Figure 2.21 Photocurrent comparison on Au-CNT, Cu-CNT and Ag-CNT.

830 nm 3000 mW/cm^2 irradiation. The parasitic capacitance reduces the sensor response while increasing substrate surface voltage. It dominates nano electrical device performance, nano material sensors when the device feature size shrinks into sub-micro or nano scale. Secondly, a novel isolation electrical feedback system was introduced into DEP system. By measuring the impedance changes, the system can quantitatively count the number of carbon nanotubes bridged between two electrodes. The experimental results show that the system response speed is fast to single wall and multi-wall CNTs, although the impedance differs much. This system will also applicable to thin layer graphene deposition control and other nanomaterials deposition and localization. Thirdly, the robust test bench using digital microscope and precise five axis substage is used to measure detector photoresponse. The relative position between nanoscale sensor and IR beam is localized by mapping the photocurrent on laser spot. The distance between photodetector and infrared laser lens is leveraged by digital microscope. The experimental results show photovoltaic quantum effect dominates CNT-Metal Schottky based IR detector and the photoresponse is dependent on contact size and metal materials. Our proposed measurement method provides a robust and precise approach to characterize sub-micro and nano scale photodetector which is important for sub-wavelength scale photodetector characterization.

Chapter 3

Single Pixel Infrared Camera

3.1 Previous Work

Over the past few years, there have been considerable breakthroughs in the thermal imager market including the fact that prices have dropped considerably. With the advances in technology and materials, infrared cameras are finally being designed for the end user. Regarding to the cost, there was no new fully-featured imager less than $\$20,000$ USD. However, there are many fantastic choices out there for building applications with a wide range of features in prices ranging from $\$2,000$ to $\$9,000$ USD. As complex as some systems may seem, infrared cameras are comprised of some basic components: lens, detector, processing electronics, display, controls and power supply. Some features such as thermal sensitivity and detector size are useful in evaluating performance.

Thermal Sensitivity: This is the most important specification to evaluate IR camera. The thermal imager is able to resolve temperature differences at least 0.1° (100 mK) or lower. The smaller number indicates the better (i.e. more sensitive) of the system. In overall, a handful of 40-50 mK $(0.04 - 0.05^\circ)$ system is also available to provide fantastic image quality and clarity. The lower sensitivities are capable of discerning smaller temperature variations typically encountered in marginal inspection conditions (when the inside to outside wall surface temperature difference is low). In other words, the additional cost of improved sensitivity is an investment that can have real returns.

Detector Array Size: Most infrared sensor arrays have fewer pixels than visible-light cameras.

IR imagers available for the civilian market are a long way from the 5-8 megapixel visual arrays which are used to seeing on most smartphone cameras. However, more pixels generally means greater detail and higher resolution infrared cameras can measure smaller targets from further away and create sharper thermal images, adding up to more precise and reliable measurements. Excellent infrared systems for civilian application are now being made with 120×120 (14,400), 160×120 (19,200) and 320×240 (76,800) focal plane arrays (FPAs). The FPAs smaller than 120×120, though financially attractive, don't provide sufficient spatial resolution. On contrary, the FPAs larger than 320×240, such as 640×480 (307,200), produce an impressive image but cost more. In addition, it must be aware of the difference between detector spatial resolution and display resolution.

Image Display: A high-quality LCD display screen is essential to diagnosing an image. There must be clarified the display resolution and sensor array resolution. In product description, some manufacturers boast about a high resolution LCD and hide their low resolution detector. For instance, LCD resolution may spec at 640×480, but if the IR detector pixel resolution is only 160×120, or 19,200 pixels, the greater display resolution accomplishes absolutely nothing. The quality of the thermal image and its measurement data are always determined by the detector resolution.

Frame Rate: 9 Hz systems have become widely available and work just as well as 30 Hz and 60 Hz systems. However, the higher frame rate has the better ability to render and capture moving targets. Lower frame rates are less tolerant to movement and will blur the image if crossing a scene too quickly. Although there is an important consideration for industrial thermographers who are inspecting certain types of rotating equipment (motor shafts, bearings or couplings), it is far less of a concern in building applications where the targets are stationary.

The thermal infrared (IR) camera that attaches to a smart phone (FLIR One/Seek Thermal) is now available, which brings infrared technology into consumer electronics. The novel $250 Seek Thermal infrared camera evaluates its effectiveness in helping wildland firefighters find lingering smoldering areas during the mopup stage of fire suppression. FLIR ONE is another lightweight easily connect and use in smartphone. It explores the natural world with no additional cords, cases, devices or screens. The IR-Blue is an affordable thermal imaging accessory for iPhone and Android devices, which uses a 64 zone 16×4 pixel non-contact infrared sensor array to read the temperature in viewing and connects using bluetooth to iPhone or Android device to show the temperature reading as colors. The novel infrared cameras are mostly in research lab. A single pixel IR camera was proposed in [69], where the camera system used a single CNT photodetector to compressively sample the linear projection of the image onto binary random patterns. By employing compressive sensing algorithm, high resolution image can be achieved with fewer samples than original image dimension.

In 2006, compressive sensing based new digital image/video camera directly acquires random projections of a scene without first collecting the pixels/voxels [96]. The camera architecture employs a digital micromirror array (DMD) to optically sample linear projections of the scene onto pseudorandom binary patterns. The key hallmark is its ability to obtain an image or video with a single detection element (“single pixel”) while measuring the scene fewer times than the number of pixels/voxels. Since the camera relies on a single photodetector, it is also adapted to image at wavelengths where conventional CCD and CMOS imagers are blind.

3.2 Spatial Light Modulator based Imager

3.2.1 Compressive Sensing

The conventional Nyquist-Shannon sampling theorem requires the sufficient sampling rate at least twice of signal frequency in order for fully reconstruction guaranteed. The compressive sensing, a new computing paradigm directly samples the signal in compressed form so that the sampling rate can be significantly reduced, which has attracted extremely interest in imaging [54] [69], geophysical data analysis [97], control and robotics [98], communication [99], medical imaging processing including MRI, CT [100]. In compressive sensing, there is no need to design sensors with higher bandwidth than original signals to follow and capture [101]. It can be seen as a sum of the linear projection from original signal to measurement matrix. The compressive sensing is a combined sampling and compression process, which is the most efficient way to sample signals from single processing point of view. E. Candes, etc. gave the mathematic proof of using random measurement matrix to recover the original signal by solving minimization of the ℓ_0 and ℓ_1 optimization problem [57].

Given an unknown signal $\boldsymbol{x}$ ($\boldsymbol{x} \in \mathbb{R}_N$), compressive sensing takes M times linear measurements from measurement matrix to original signal $\boldsymbol{x}$, as shown in Eq. (3.1). If $M = N$ (N is the dimension of unknown signal $\boldsymbol{x}$), the signal $\boldsymbol{x}$ could be easily reconstructed by solving linear equation or else it would be an underdetermined question. However, the original signal can also be reconstructed with less measurements ($M << N$) using compressive sensing, where Φ is a $M \times N$ measurement matrix, which transforms the measurement as a linear projection from original signal to measurement matrix. $\boldsymbol{y}$ is the measurement result. The optimal solution for $\boldsymbol{x}$ could be reconstructed [102] by solving the minimization of the 1-norm optimization. More discussion and comprehensive introduction on compressive sensing can be seen on [103] [104].

$$\boldsymbol{y} = \Phi \boldsymbol{x} \tag{3.1}$$

The core of compressive sensing is the solver of problem of Eq. (3.1). It is NP-hard unfortunately. In $\boldsymbol{y} = \Phi \boldsymbol{x}$. Φ ($\Phi \in \mathbb{R}^{m \times n}$) is the measurement matrix ($m < n$).

Instead of solving the NP hard directly, Donoho [104] proves that if $\boldsymbol{x}$ is sparse and Φ is under some conditions, such as the null space property [105], the incoherence condition [106] and restricted isometry property [55], the Eq. (3.1) will be equivalent to Eq. (3.2). The nature image ($\boldsymbol{x}$) sparsity will be discussed in Chapter 5.

$$\begin{aligned} \tilde{\boldsymbol{x}} &= \text{argmin} \, \| \boldsymbol{x} \|_1 \\ \text{s.t.} \quad \Phi \boldsymbol{x} &= \boldsymbol{y} \end{aligned} \tag{3.2}$$

A lot of algorithms have been proposed to solve this ℓ_1 problem, including Orthogonal Matching Pursuit (OMP) [107], Iterative Shrinkage-Thresholding (IST) [108], CoSaMP [109], subspace pursuit [110], Accelerated Proximal Gradient (APG) [111], and Alternating Direction Method (ADM) [112] and its linearized version (LADM) [113].

3.2.2 Single Pixel Imager

The structure of the IR sensing system is built upon the compressive sensing theory introduced in the previous section. Figure 3.1 depicts the system setup using a single pixel CNT photodetector as image sensor, in which IR images are directed onto a DMD through a set of lenses. The IR images on the DMD represents the original signal x in Eq. (3.1) ($x \in \mathbb{R}^n$, n is the number of image pixels). The DMD generates patterns according to a measurement matrix Φ ($\Phi \in \mathbb{R}^{m \times n}$) so as to compress the IR images. Each pattern on the DMD is comprised by n pixels, thus there are m

different patterns form the measurement matrix Φ. The compressed IR images are reflected and focused to a CNT photodetector. The IR signal arriving at the CNT photodetector represents the linear projection of the image onto the measurement matrix, which could be considered as the inner product of IR image x and each row vector in the measurement matrix Φ. The photocurrent is recorded by a fast readout system integrating with a charge integrator, an Analog-to-Digital Converter (ADC), and a Data Acquisition (DAQ) card. The amplitude of photocurrent represents the value of y in Eq. (3.1) ($y \in \mathbb{R}^m$, m is the number of measurements). It should be noted that compressive sensing sets $m < n$, therefore, the IR image x is compressively sampled into the y. Based on the measurement results y and the designed measurement matrix Φ, the original signal x can be recovered using an image reconstruction algorithm.

3.3 Weak Signal Readout Method

The readout circuit is one of the key components in imaging system. With the limitation of nanomaterial electron transport, the photocurrent is from *pA* to *nA* scale in CNT infrared sensor [32] [33]. More importantly, it is bias dependent. In order to read photocurrent in CNT infrared sensor, a low noise, high gain readout circuit is needed, in which a fully current readout system includes a current to voltage module, and a Digital Signal Processing module (DSP).

3.3.1 Current to Voltage Conversion Method

There are two classes of CNT photodetector signal monitoring, voltage and current based. The former requires great voltage difference (at least *mV* scale) on sensor when light irradiates on it, especially for photovoltaic device. However, the IR irradiation is nonlinear to voltage generated in quantum effect infrared sensor but linear proportional to photocurrent. Meanwhile, in CNT

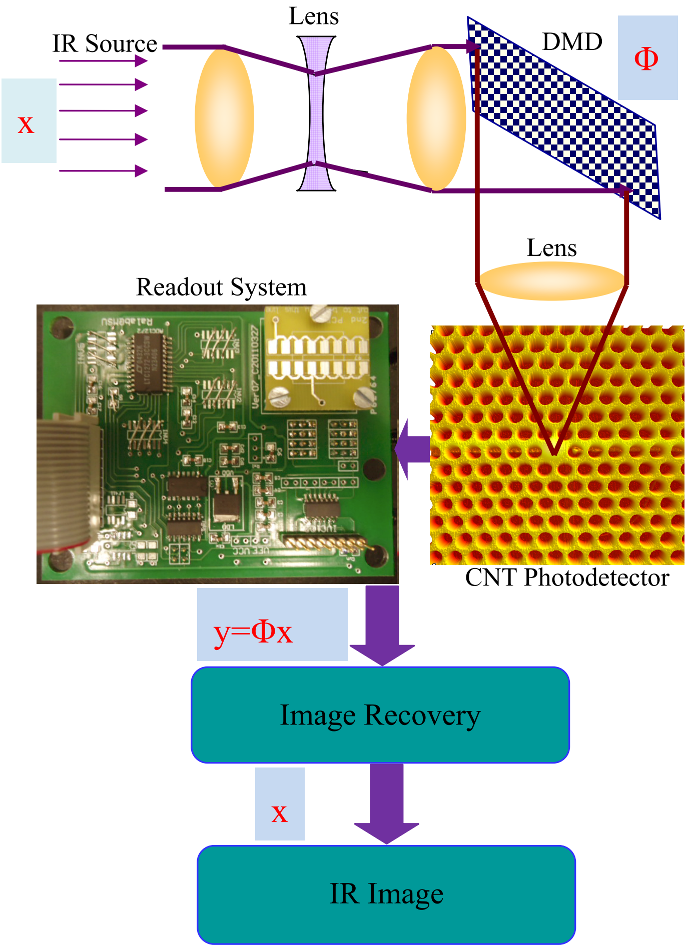


Figure 3.1 System setup of compressive sensing based imaging system using a CNT photodetector, response enhanced by photonic cavity.

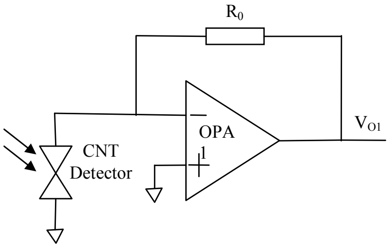


Figure 3.2 Schematic of R type IV converter.

detector experiments [114], CNT photovoltaic effect is not as silicon photodiode. It generates *nA* scale photocurrent while the voltage doesn't change much. Therefore, a current to voltage conversion will be designed for CNT photocurrent measurement. The current monitoring requires the readout circuit present zero input impedance to the detector, and absorb the detector's current without producing a bias voltage. Since the photocurrent highly depends on the applied bias, it is desired to use zero bias for current monitoring on CNT detector.

3.3.2 Resistor based Current Readout Method

In order to convert pA/nA photocurrent to voltage, a current to voltage amplifier (transimpedance amplifier)becomes the most important module in readout system. The basic current to voltage converter (IV converter) is to use photocurrent (Ip) multiplied by a resistor. These circuits, shown in Figure 3.2, need very large (giga ohms) and precise resistor to detect *pA* current [115]. The thermal noise of the large resistor will also limit the resolution of the circuit, although the deep negative feedback (OPA_1) set CNT detector at zero bias in this condition.

3.3.3 Capacitor Based Current Readout Method

Capacitive Trans-Impedance Amplifier (CTIA) also works as a current to voltage converter and it can be designed to have high charge to voltage conversion ratio with low noise. Figure 3.3 is

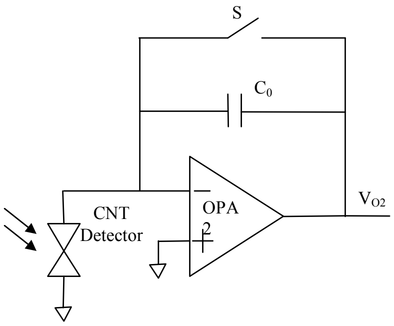


Figure 3.3 Schematic of C type IV converter.

the basic circuit of CTIA using switch capacitor as adjust resistor, in which there are two working phases. In phase1, switch S is ON, OPA_2 is reset the output as zero. In phase2, switch S is OFF in T_{off} and the photocurrent (Ip) will charge negative node (-) of OPA_2 to redistribute it, where $Q = I_p * T_{off}$, $V_{o2} = Q/C_0$, (T_{off} represents switch S off time). In this design, the circuit can detect ultra low current based on the noise level in input. It also achieves a large Power Supply Rejection Ratio (PSRR), high open loop gain and large dynamic range. The CTIA gain is given by $G = Q/C_0$.

3.4 ROIC Structure for CNT IR Sensor

3.4.1 Zero bias Readout Circuits

As discussed in previous section, there are two requirements on readout circuits. One is the highly resolutions to picoampere scale and another is no bias or bias modulated on the detector. In order to reduce noise and read such low current from MWCNT detector, a CTIA, shown in Figure 3.3 was designed in this IR readout system. Since the offset of input generates a serious bias on CNT IR sensor due to input asymmetry, in the proposed design, a dummy CNT IR sensor will be connected on positive input of OPA, as shown in Figure 3.4, such that the input impedance match will be in

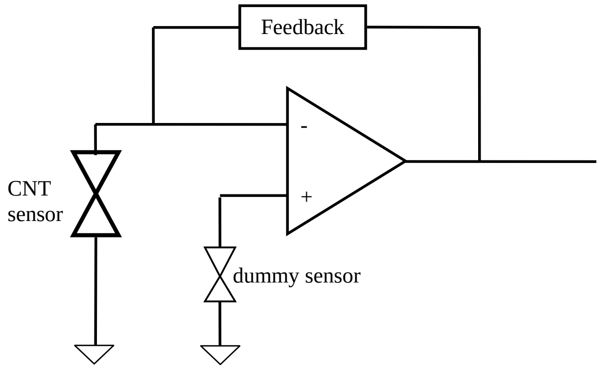


Figure 3.4 Zero bias readout circuit.

same order.

3.4.2 High Gain Current to Current Converter

In order to get high speed CNT based IR sensor readout system, the current needs amplified before converting to voltage.

In this part, a proposed current amplified readout method for carbon nanotube infrared sensor was described herein, shown in Figure 3.5. There are four electrodes, in which left part is CNT sensor, while right part is conventional Bipolar Junction Transistor (BJT) structure. In this configuration, the CNT sensor is connected to the base of BJT, so that the photocurrent will flow to the base node under condition of infrared irradiation. The base current could be amplified by BJT working principle [116].

3.4.3 High Speed Readout

In high resolution ADC, delta-sigma (ΔΣ, or sigma-delta, ΣΔ) modulation is a method for encoding high resolution signals into lower resolution using pulse-density modulation. A high performance, 24-bit ΣΔ analog-to-digital converter was used in experiments. It combined wide input bandwidth

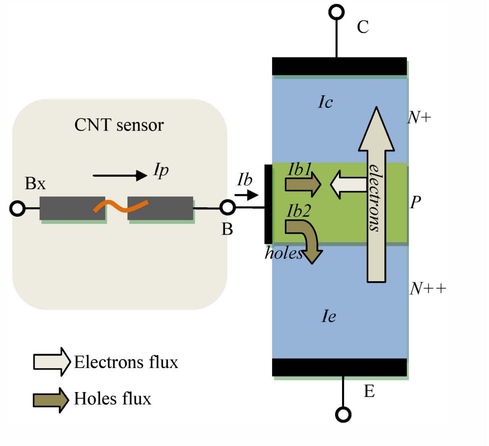


Figure 3.5 Current to current converter for CNT IR sensor.

and high speed $\Sigma-\Delta$ conversion with a performance of 106 dB SNR at 625 *kSPS*, making it ideal for high speed data acquisition. Wide dynamic range, combining with significantly reduced anti-aliasing requirements, simplifies the design process. In addition, the device offers programmable decimation rates and digital FIR filter. It is ideal for applications demanding high SNR without a complex front end signal processing design.

3.5 Hardware Experimental Performance and Applications

3.5.1 Readout System Testing

In this section, the hardware setup and some experimental results are presented. For the sake of reducing the noise after CTIA, a Low Pass Filter (LPF) was introduced to limit the bandwidth and optimize the thermal noise contribution on CTIA. Meanwhile, it weakens all high frequency noise, the power frequency noise included. Figure 3.6 is the experimental diagram of readout system.

Based on this readout system, a group of experiments were designed to verify the readout performance. As shown in Figure 3.7, it measures the linearity and stability of ROIC, where *x*-axis is corresponding to laser intensity. The digital output at T = 1*h* keeps close as value at T =

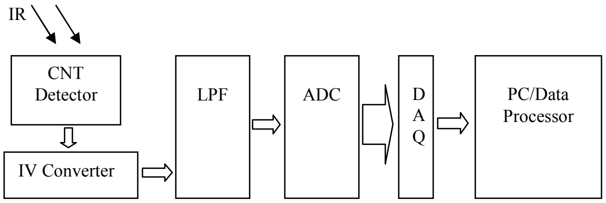


Figure 3.6 Diagram of readout system.

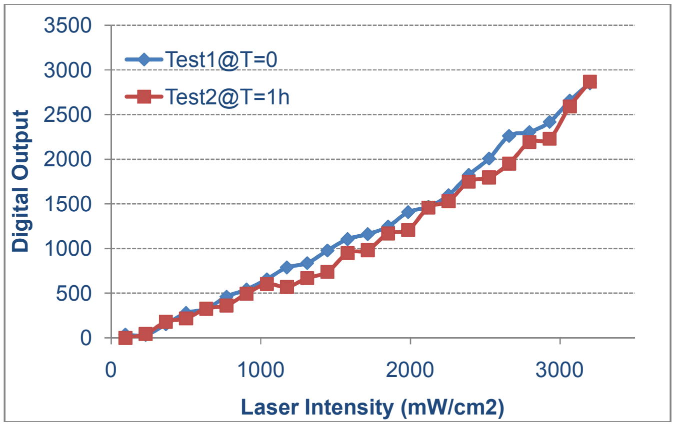


Figure 3.7 Readout linearity test on CNT based IR detector.

0 with linear response. Meanwhile, the readout output was also compared with Agilent 4155C Semiconductor Analyzer (refer to Curve Tracer) measurement in Figure 3.8. The ROIC output follows Curve Tracer very well. The results show that the readout system can reach *pA* resolution with high stability performance.

3.5.2 Readout Applications

In the CNT based single pixel IR camera, it integrated the precisely weak signal readout system to fulfil a camera. The hardware set up was shown in Figure 3.9, where IR irradiation was controlled by Digital Micromirror Device (DMD). Based on the hardware setup, three experiments were realized, shown in Figure 3.10. A rectangle bar was moved from top to bottom and the recovery image could follow target image very well. In visually, there are two features in the restored image, noisy and partial recovery image. The noise originates from sensor and sampling process.

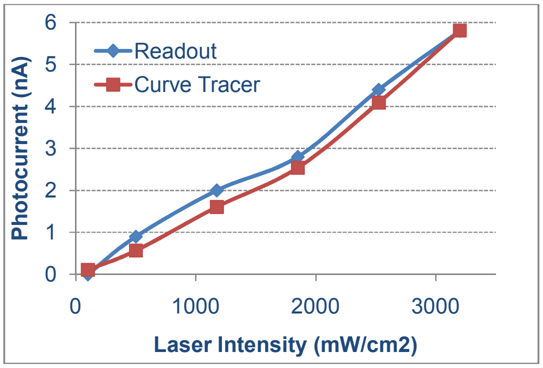
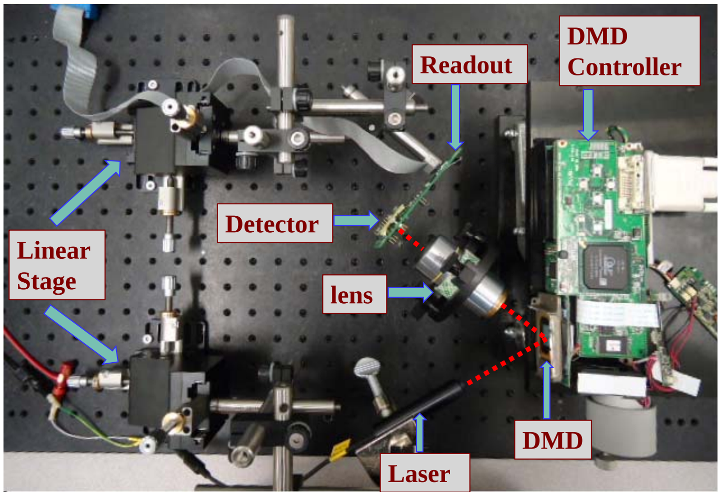


Figure 3.8 Readout comparison between proposed system and semiconductor parameter analyzer (Agilent 4155c).

Meanwhile, the results also show that the lower left corner is destructive (partial image) due to such tiny sensing area. From these results, it is found that the readout system work perfectly in CNT based IR single imaging system.

3.6 Chapter Summary

Nano-photodetectors have demonstrated promising performance to detect IR signals, especially CNT photodetectors. However, a small absorption area and the difficulties to fabricate a large scale photodetector array impede its application in imaging systems. In order to overcome these problems, a compressive sensing based IR camera system was developed. Firstly, the ultra high speed, high resolution readout system was presented in this study. It could test the current to sub *nA* in CNT based IR detector with low noise so as to integrate it in CNT based single pixel IR imaging system. The experimental results show that the number of measurements required to recover the images can be much fewer than the pixel number of original images. This camera system was capable of observing the dynamic movement of a laser spot in near infrared wavelength. In general, the compressive sensing method might make the nanosensor based infrared camera achievable.



(Side View)

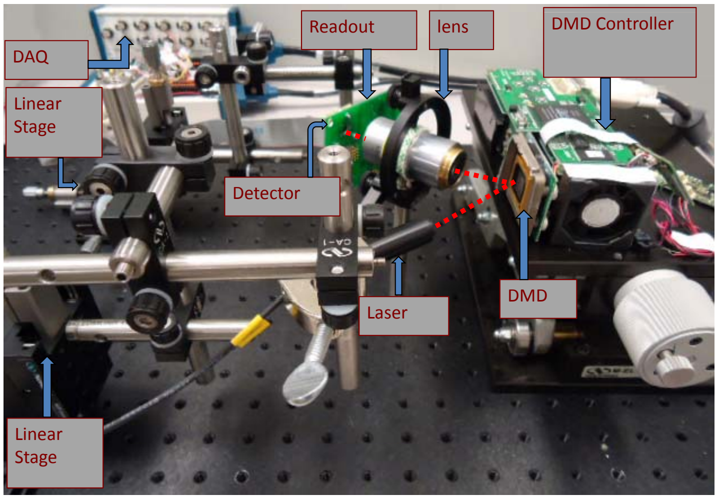


Figure 3.9 Hardware setup of single pixel IR imaging system.

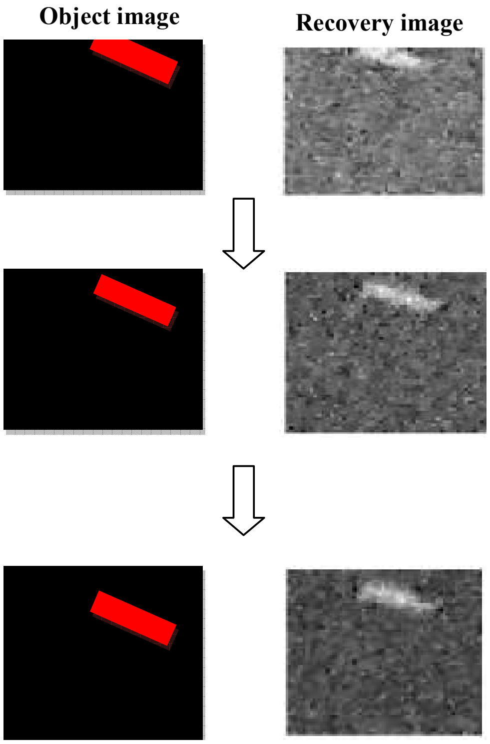


Figure 3.10 Images recovery based on single CNT detector.

Chapter 4

Light Field Imaging

4.1 Previous Work

The conventional photographs only record the sum total of light rays for each point on image plane so that they tell little about the amount of light traveling along individual rays. The focus and lens aberration problems have challenged photographers since the very beginning, therefore, light field photography was proposed to solve these problems. In geometric optics, the fundamental carrier of light is a ray and it is focused on incoherent light with objects far larger than the wavelength. The radiance is represented to measure the amount of light which is interpreted as a field since 1846 [117]. In 1936, Gershun introduced this light field terminology, the first systematical physical theory of light [118], which defined the light field as a function of position and direction in regions of space free of occluders. The plenoptic function firstly modeled the light field using seven parameters, shown in Eq. (4.1) in general [59].

$$P = P(x, y, \lambda, t, V_x, V_y, V_z) \tag{4.1}$$

$$P = P(x, y, u, v) \tag{4.2}$$

The rays in space can be parameterized by three coordinates x, y, and z and two angles θ and ϕ for any stationary, quasi monochromatic light field,as shown in Figure 4.2. The 5D radiance function describes the appearance of scene through all light rays (2D) emitted from 3D point. It is

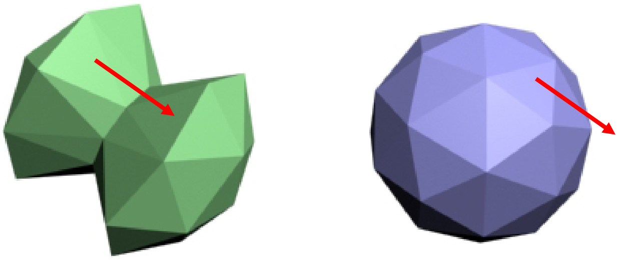


Figure 4.1 Concave object radiance (left) and convex object radiance (right).

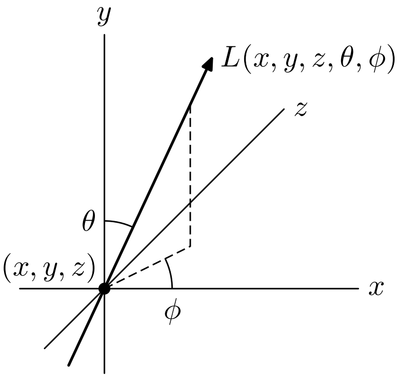


Figure 4.2 Parameterizing light ray in 3D space by position (x,y,z) and direction (θ,ϕ).

a five-dimensional function, that is a function over a five-dimensional manifold equivalent to the product of 3D euclidean space and the 2-sphere. The 5D plenoptic function is not practical since it is impossible to capture views by point to point. However, it is can be reduced to four dimensions by two assumptions. Firstly, the region of interest is restricted to locations outside of convex hull of the object because the light leaving from one point may end by another point on a concave object, as shown in Figure 4.1. Secondly, it is reduced as four parameters with the assumption of rays passing through free space regions and free of occluders, such as opaque objects, and scattering media, such as fog. Under these assumption, the light traveling is constant along its pathway, eliminating one dimension of variation [60].

Based on 4D light field model, there are many light field photography implementations proposed in literatures. Recently, lenslet-based light field system has been integrated into digital cameras [119]. Light modulating in mask-based systems have evolved to be more light efficient than pinhole arrays [120]. Nevertheless, both approaches sacrifice image resolution because the number of sensor pixels is the upper limit of the number of light rays captured. To avoid this limitation, alternative design has been proposed which favors spatial resolution over angular resolution [121]. In order to fully preserve image resolution, current options include camera arrays or taking multiple photographs by a single camera [122]. In overall, the time sequential approaches are limited to static scenes since high speed camera arrays are costly and bulky. The idea of compressive light field acquisition itself is not new either. In [123], it simulates a compressive camera array for light field sensing. Recently, researchers have started to explore compressive light field acquisition using a single camera, such as optical coding strategies including coded apertures and coded lenslets. It is a combination of coded mask, aperture and random mirror reflections. Unfortunately, they require multiple images to be recorded such that they can't capture dynamic scenes, though they succeed in reducing the number of shots compared to noncompressive counterparts.

Light field camera is not only active in research lab, it is also very popular in business prospects. Lytro was founded by Stanford university computer graphics laboratory alumnus Ren Ng to commercialize the light field camera. They developed a consumer light field digital camera capable of capturing images using a plenoptic technique. Raytrix is another light field camera company which has sold several models of plenoptic camera for industrial and scientific applications since 2010, with resolutions starting from one megapixel. Pelican Imaging has thin multi-camera array systems intending for consumer electronics, which use from 4 to 16 closely spaced micro-cameras instead of a micro-lens array image sensor. Meanwhile, Nokia has invested in Pelican Imaging to produce a plenoptic camera system with 16-lens array camera to be implemented in Nokia smartphones. The Adobe light field camera is a prototype 100-megapixel which takes a three-dimensional photo of the scene in focus using 19 uniquely configured lenses. Each lens will take a 5.2-megapixel photo of the entire scene. The CAFADIS camera is another plenoptic camera developed by the

University of La Laguna (Spain). CAFADIS stands (in Spanish) for phase-distance camera. Since it works for distance and optical wavefront estimation, the product produces several images refocused at different distances, depth mapping, all-in-focus images and stereo pairs from a single shot. A similar optical design can be used in adaptive optics in astrophysics so as to correct the aberrations caused by atmospheric turbulence in telescope images. In order to accelerate these tasks, different algorithms running on GPU and FPGA, operate on the raw images captured by the camera. Mitsubishi Electric Research Laboratories's (MERL) light field camera is based on the principle of optical heterodyning and uses a printed film (mask) placed close to the sensor. Any hand-held camera can be converted into a light field camera using this technology by simply inserting a low-cost film on top of the sensor. This mask-based design avoids the problem of loss of resolution, since a high-resolution photo can be generated for the focused parts of the scene.

However, the core of all light field cameras in lab or commercialized system are large visible sensor array. These approaches are not applicable to IR spectrum in overall. A novel compressive sampling approach will be proposed in this research as it is necessary for IR camera to capture more information than simple 2D camera. In this thesis, the light field imaging is a narrowly name of light field omitted time and wavelength (or stationary, quasi monochromatic light), as shown in Eq. (4.2), also referred to plenoptic camera [124] or synthetic aperture camera [125].

4.2 4D Light Field Model

4.2.1 Light Field Model in Lens

The 4D light field, defined as radiance along rays in empty space, was firstly proposed in form of the light field [126] and the lumigraph [127]. In this model, the observer and the scene can be separated by a surface so that the radiance is represented by a function of light rays passing through it. The model is named two plane parameterization, as shown in Figure 4.3, where the plane (s,t) and (u,v) set the relative coordinates for light ray passing through. The (u,v)-plane is the viewpoint plane in which all camera focal points are placed on regular grid points while the (s,t)-plane is the focal plane. In order to create such a 4D plenoptic model for real scenes, a large number of views are taken. By this way, they are discrete sampling of the plenoptic function. This could be enhanced as continuous light field by assumption of the real objects to be lambertian. Any point of the object has the isotropic radiance in all possible directions so that the light rays can be fully represented by interpolating algorithms.

As shown in Figure 4.3, the radiance can be represented by $f(s,t,u,v)$ for any rays passing between (s,t) and (u,v). The pair of points between *st* plane and *uv* plane correspond only to one ray. The disadvantage of two plane parameterization is that the method can not represent rays parallel to the two parallel planes. However, as shown in Figure 4.4 in the real camera sampling system, the lens plane and sensor film are assigned as *uv* and *st* plane distinctly, because every ray that contributes to a photography must pass through the lens and terminates somewhere on the film. It also explains the camera limits light rays within Angle Of View (AOV). In Figure 4.5, $f(s_0,u_0)$, $f(s_0,u_1)$, $f(s_0,u_2)$ show the light rays passing through u_0, u_1 and u_2. The u $((u,v))$ is the directional axis because the u intercepting on the lens determines the direction at which the rays will be collected by sensors. The s $((s,t))$ refers to the spatial axis. In general, it can be mapped to a point on the ray-space diagram on Figure 4.5 for any rays in diagram Figure 4.4.

In conventional camera, the focus will be adjusted by changing the distance between the film and the lens. The ray space will also be changed as the camera focus on the different depth. As shown in Figure 4.6 (top), when moving the sensor array (film) plane closer to the lens, the rays from selected point will correspond to a line with positive slope, shown in Figure 4.7 (left) and all rays from same point must be on this positive slope line in ray-space diagram. On the contrary, it

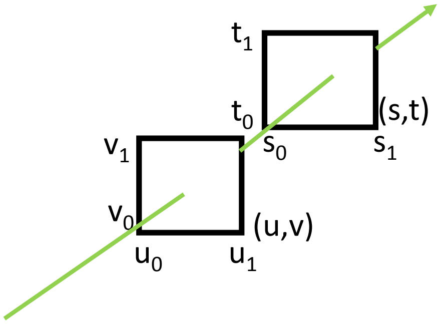


Figure 4.3 Two plane parameterization for light field.

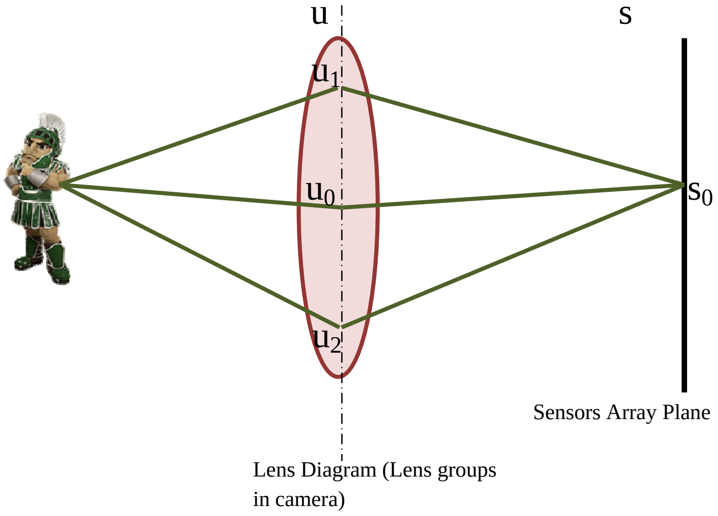


Figure 4.4 Two plane parameterization in SLR camera.

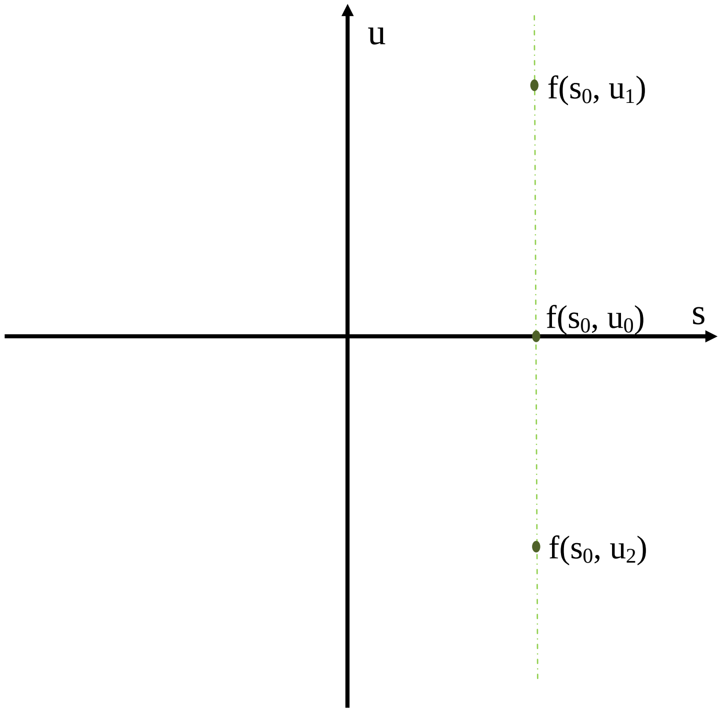


Figure 4.5 Two plane parameterization in Cartesian coordinates.

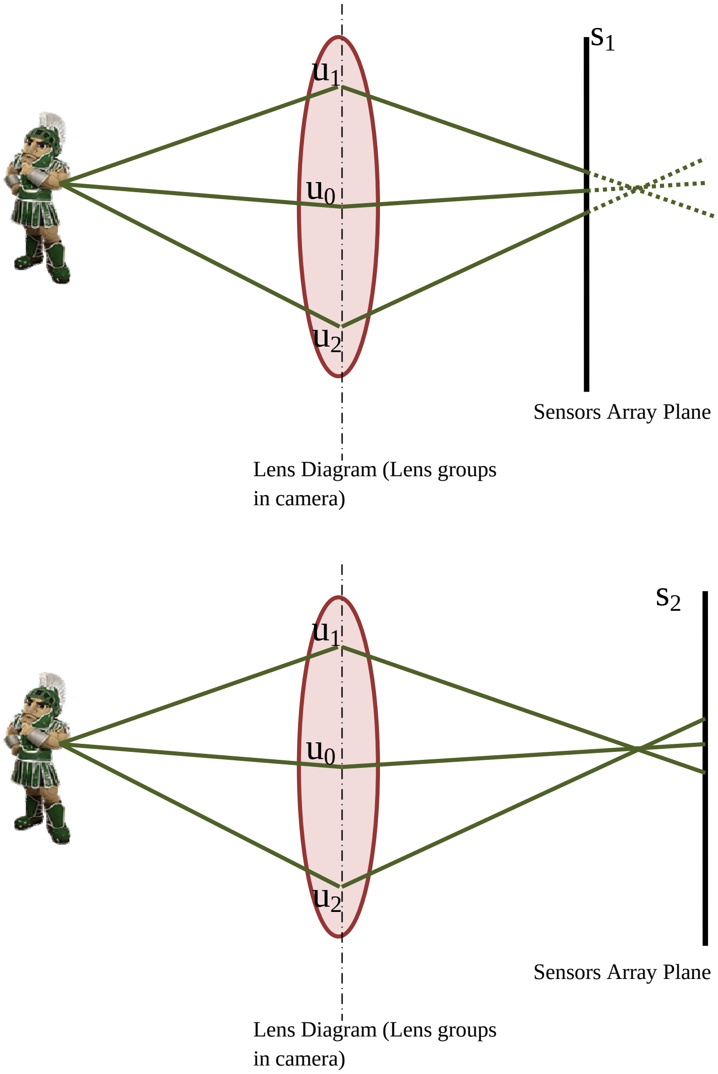


Figure 4.6 Light ray diagram of camera (unfocused).

will be a negative slope line when moving the film further from the lens, as shown in Figure 4.7 (right). The slope on the ray-space depends on the separation between the real image plane and film plane [60].

4.2.2 Light Field Model in Mirror

In order to get fully light field information of scene, there is at least 4D space to support all values as $f(s,t,u,v)$. In the DMD based light field sampling, it also requires two plane parameters to

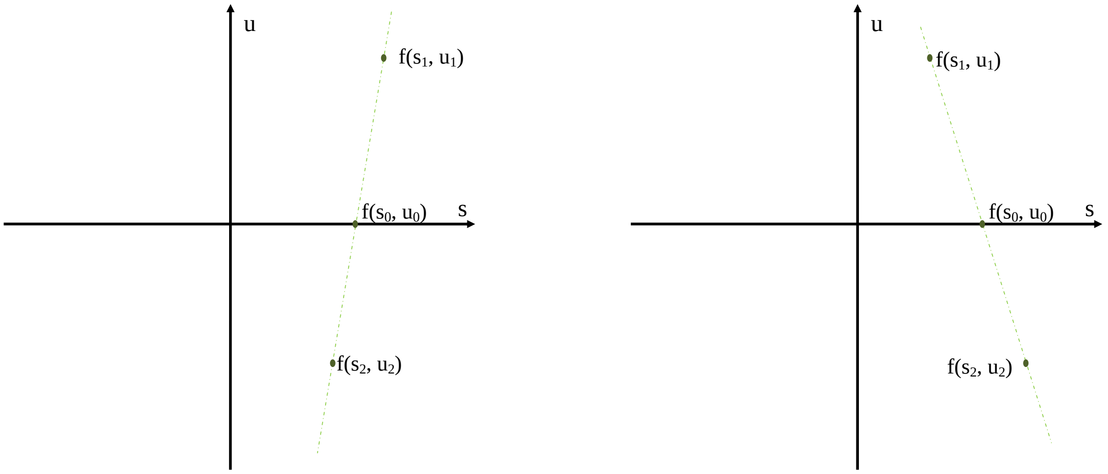


Figure 4.7 Rays in Cartesian coordinates (unfocused).

characterize light rays. In the mirror reflection, it is a reflected duplication of an object that appears identical but mirrored. In geometry, the mirror image of an object or two-dimensional figure is a virtual image formed by specular reflection. It is the same size as the original object except the object or figure has reflection symmetry (also known as a P-symmetry). Reflection in a mirror also results in a change of chirality, more specifically from a right-handed to a left-handed coordinate system (or vice versa). As a consequence, if one looks in a mirror and lets two axes (up-down and front-back) coincide with those in the mirror, this will give a reversal of the third axis (left-right). The mirror image appears to be three-dimensional if the observer moves or uses binocular vision. Although the mirror changes the light distribution in the halfspace in front and behind it, the mirror image does not violate the conservation of energy as the mirror simply redirects the light. In terms of the light distribution, the DMD only changes the direction of lights stopped on its surface so that the mirror plane could be one internal plane of light field modeling, as shown in Figure 4.8. The second plane would be selected in front of DMD and it is the mask (aperture) plane. Through the mask and DMD plane, all light rays passing within these two objects can be modeled.

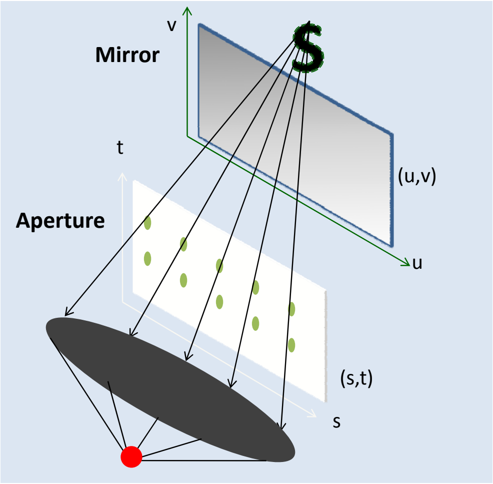


Figure 4.8 Two plane parameterization in DMD based imaging system.

4.3 Mask based Single Pixel Light Field Sensing

4.3.1 Optics and System Design

As shown in Figure 4.9, the basic single pixel camera is composed of main lens, digital micro mirror device (DMD, working as spatial light modulator), second lens for converging light to single sensor, and IR sensor. After main lens, the scene or desired object is projected onto DMD array to form a virtual image behind, shown in Figure 4.9 inset. The DMD can be controlled individually and DMD patterns are decided by measurement matrix. Through multiple patterns, it will generate a serial of measurements, in which each mirror corresponds to part of the desired object and contributes to one pixel in recovery image. The key optical component in this light field imaging system is aperture on mask, which controls the light ray directions. As shown in Figure 4.10, multiple angular images are captured in single pixel camera in order to get all directional rays. There are two aperture holes of $A1$ and $A2$, which generate two distinct recovery image $di1$ and $di2$ in image space. They represent two distinct angular information (group of light rays) from scene. The whole light rays can be recorded through multiple aperture positions.

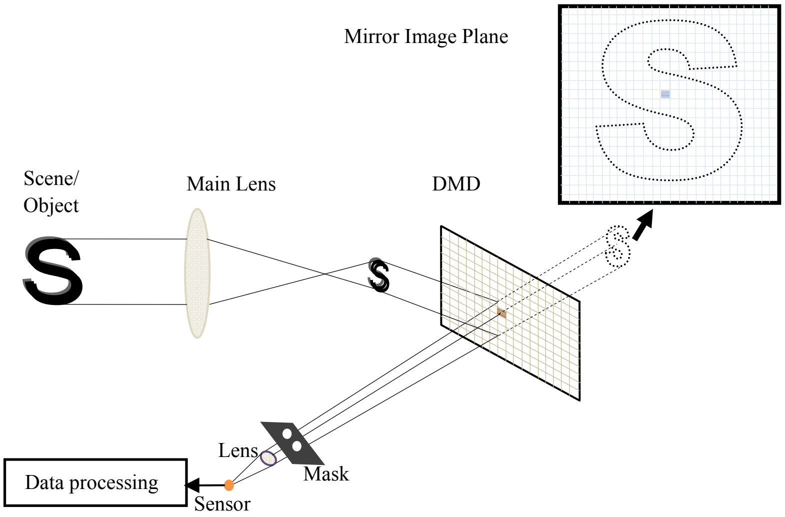


Figure 4.9 Schematic diagram of single pixel light field sensing.

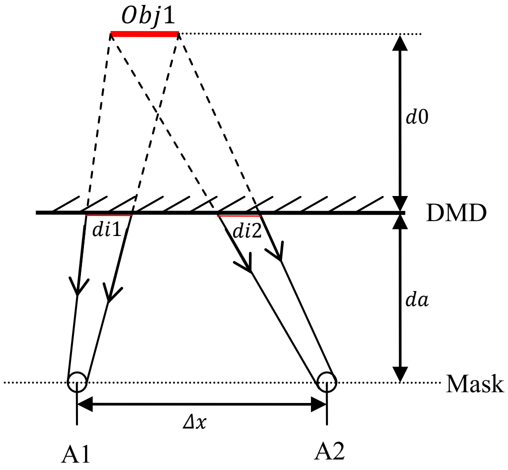


Figure 4.10 Distinct angular image from two aperture.

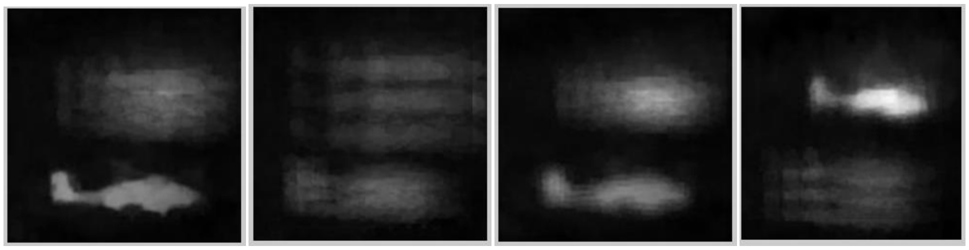


Figure 4.11 Synthetic aperture imaging. The focus plane is becoming far away to main lens from left to right.

4.3.2 Experimental Performance

The most popular application of light field camera is to study the synthetic aperture photography, on which the focus is adjustable using a single shot. In the experimental setup, there are two aircrafts located in two distinct planes along the main lens, one close to the lens and another is further away from lens. During sampling, 25 angular (directional) images are reconstructed. For each angular, it will sense distinct rays from same object. The plane and parallax method was used to show refocus of single pixel IR imaging, shown in Figure 4.11. The left shows that camera works in near focus on the bottom aircraft while the far aircraft will become clear when the focus plane moves away. Meanwhile, both objects will have kind of blur when focusing between them.

4.4 Double Compressive Light Field Sensing

The light field sensing are characterized by spacial resolution and angular resolution though it is trade-off in conventional design. This problem could be solved by multiple cameras design. However, the more cameras introduce more difficult optics design and larger space. In compressive light field sampling, it is obvious that spatial sampling is sparsity as single pixel camera [128]. The angular information could also be sparsity since the difference is relatively small between adjacent angular images. In this research, double compressive light field sensing includes the basic spacial and angular sensing. The spatial compressive sensing is discussed in Chapter 3. The second compressive regime is applied for angular image difference. By taking the first angular image as full recovery, the system only samples the difference between second and first angular image. In the following recovery algorithm, it also only reconstructs the difference and the whole second reconstructed image is combined by this difference with first angular image.

4.4.1 Modeling of Double Compressive Light Field Sensing

The simplified light field is modeled as 4D information, including 2D spacial images and 2D angular values. The light field sampling is described as 2D images along 2D angular values so that the angular images could be modeled as series, shown in Figure 4.12. Each individual angular image is a single sampling and recovered by single pixel imager algorithm. Along the angular axis, there is also some redundant in some way. The difference between two adjacent angular images are quite sparse in spatial domain, as shown in Figure 4.13, in which the intensity is in grayscale and the nonzero values (significant changes) are shown in Figure 4.13 (b). Based on the difference signal, another compressive sensing, referred to double compressive sensing, can be applied on next angular image reconstruction.

4.4.2 Recovery Algorithm

As discussed in Chapter 3, ℓ_1 minimization is most popular in compressive sensing due to it is ℓ_0 equivalence under some conditions. However, Total Variation (TV) regularization makes the recovered image quality sharper and preserves the edges or boundaries more accurately. It not only reconstructs sparse signals or images but also succeeds when the gradient of the underlying signal or image is sparse. There is only a limited number of TV solvers available, including SOCP [129],

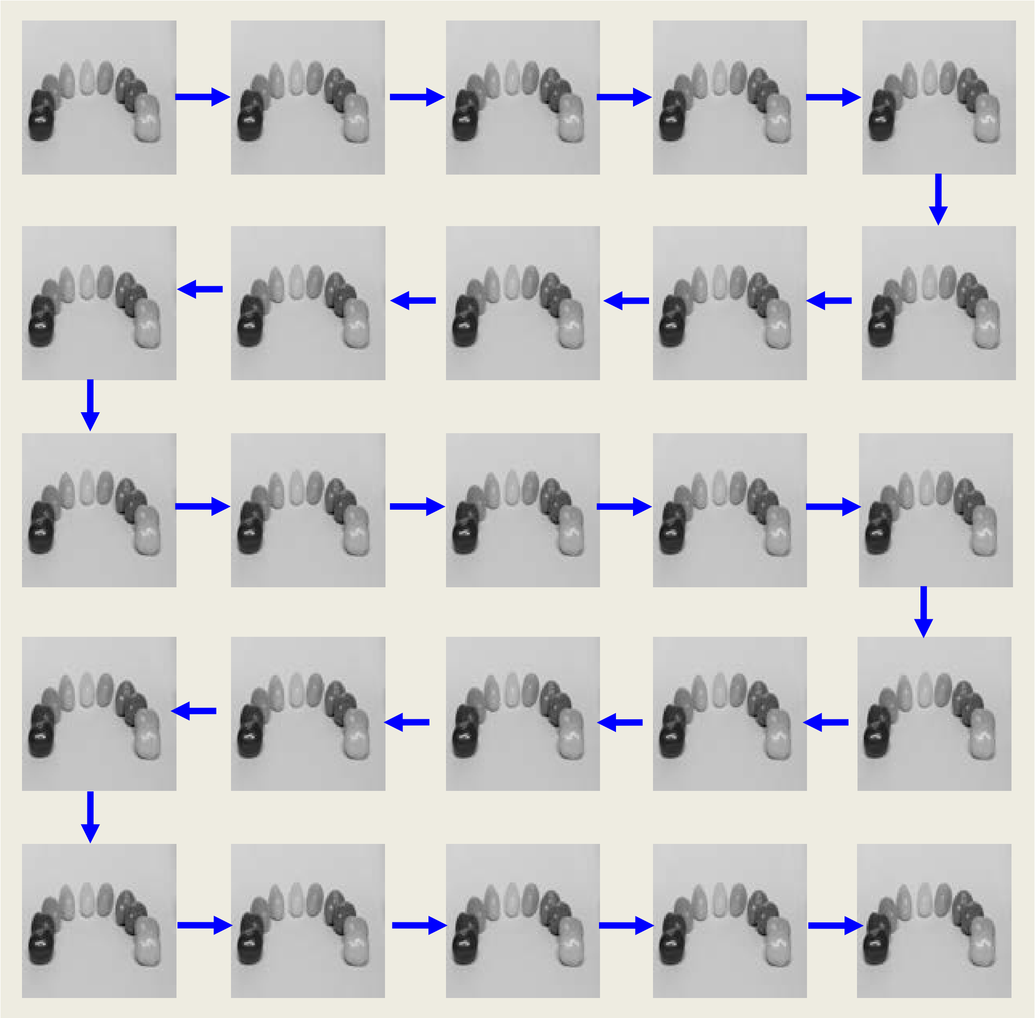
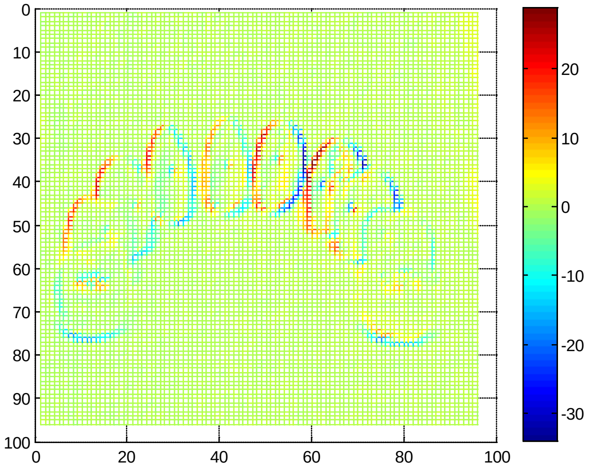


Figure 4.12 5×5 angular images of Stanford jelly beans.

ℓ_1-Magic [57], TwIST [130] and RecPF [131]. In this research, TVAL3 [128] is used for double compressive recovery. The TV regularization solves the problem of Eq. (4.3), where $\tilde{\boldsymbol{x}}$ is estimated value, $\boldsymbol{x} \in \mathbb{R}^{s \times t}$, $D\boldsymbol{x}$ is the discrete gradient of $\boldsymbol{x}$, Φ is the measurement matrix (random Bernoulli). $\boldsymbol{y}$ is the observation of $\boldsymbol{x}$ via some linear measurements. It can be either 1-norm (corresponding to the anisotropic TV) or 2-norm (corresponding to the isotropic TV).

$$\begin{aligned} \tilde{\boldsymbol{x}} &= \operatorname{argmin} \| D\boldsymbol{x} \| \\ \text{s.t.} \quad \Phi\boldsymbol{x} &= \boldsymbol{y} \end{aligned} \tag{4.3}$$



(a)

(b)

Figure 4.13 Adjacent angular image difference. a) intensity difference, b) significant changes (nonzero changes) of angular image.

In TVAL3 algorithm, instead of employing the augmented Lagrangian method to minimize the TV model Eq. (4.3) directly, an equivalent variant of Eq. (4.4) is considered and its unconstrained corresponding augmented Lagrangian function is Eq. (4.5). Suppose that $\boldsymbol{x}^{(l)}$ and $\omega_i^{(l)}$ respectively denote the approximate minimizers of Eq. (4.5) at the (l) iteration which refers to the inner iteration while solving the subproblem. The alternating direction method solves two subproblem respectively.

$$\begin{aligned}&\min_{\omega_i,\,\boldsymbol{x}} \sum_i \parallel \omega_i \parallel,\\ &\text{s.t.} \quad \Phi\boldsymbol{x} = \boldsymbol{y} \quad \text{and} \quad D_i\boldsymbol{x} = \omega_i \quad \text{for all} \quad i.\end{aligned} \tag{4.4}$$

$$\begin{aligned}\min_{\omega_i,\,\boldsymbol{x}} \quad \varphi_A(\omega_i,\boldsymbol{x}) = \sum_i (\parallel \omega_i\| - v_i^T(D_i\boldsymbol{x} - \omega_i) + \frac{\beta_i}{2} \parallel D_i\boldsymbol{x} - \omega_i\|_2^2)&\\ -\lambda^T(\Phi\boldsymbol{x} - y) + \frac{\mu}{2} \parallel \Phi\boldsymbol{x} - y\|_2^2&\end{aligned} \tag{4.5}$$

4.4.3 Experiments with Double Compressive Sensing

In this section, the double compressive sensing numerical results are analyzed. Firstly, the basic compressive sensing is applied on angular image recovery. The sampling ratio must be greater than 30% so as to get high image quality, as shown in Figure 4.14 (top), while the bottom line shows recovery images using difference respectively. Though it is hardly to distinguish the difference in visually, it is much clear in Figure 4.15 where two methods are compared using same measurements. The double compressive sensing can reduce sampling ratio down to 9% with high image quality while basic compressive sensing requires as high as 25%.

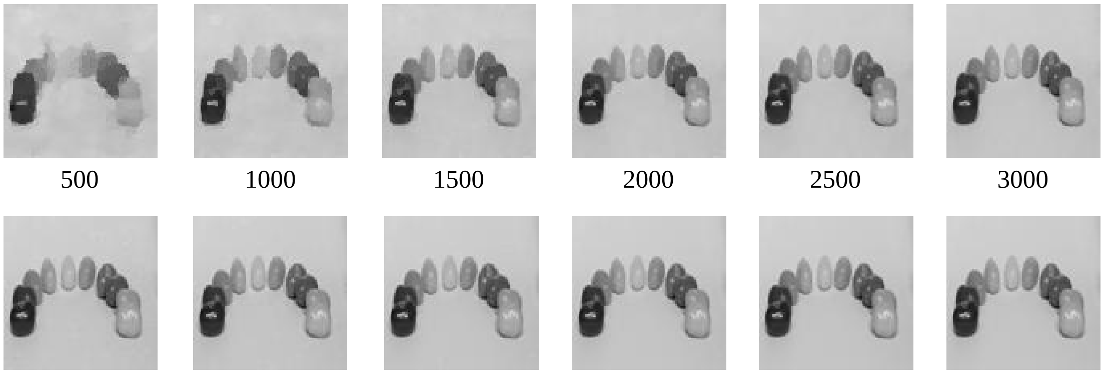


Figure 4.14 Angular image recovery comparison between basic compressive sensing and double compressive sensing.

The second simulation is based on Stanford Jelly Beans ($http$://lightfield.stanford.edu/lfs.html). There are 100 (10 × 10) images selected and the first column is chosen as reference image in each

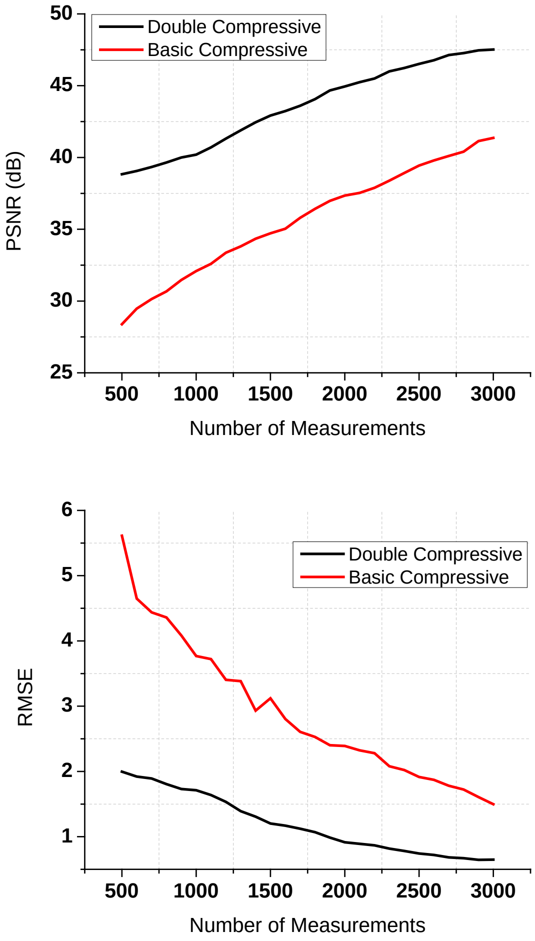


Figure 4.15 Angular image recovery residual error by RMSE and PSNR.

row. As shown in Figure 4.16, there are 10 rows, 5 columns (50 angular images) recovered. The estimation of next angular image comes from current image and the reconstructed difference. Since the residual error is such small and image spatial resolution is only 96×96, it is hard to discover difference in visual. Table 4.1 lists the RMSE and PSNR for each image compared to ground truth, where angular ID represents the image row/column (0102: 01 row 02 column). In overall, the residual error is still within 30 dB in five images according to PSNR.

Table 4.1 Characterizing angular images accumulation residual error by PSNR and RMSE.

Angular ID	0102	0103	0104	0105
PSNR (dB)	40.2367	34.7074	32.5634	31.2847
RMSE	1.7042	2.8735	3.2524	3.5392
Angular ID	0202	0203	0204	0205
PSNR (dB)	42.4017	35.2519	32.9163	31.5164
RMSE	1.3074	2.7402	3.1839	3.5185
Angular ID	0302	0303	0304	0305
PSNR (dB)	40.535	34.3587	32.738	31.1188
RMSE	1.5826	2.9546	3.207	3.5282
Angular ID	0402	0403	0404	0405
PSNR (dB)	40.5569	34.6422	32.665	31.3414
RMSE	1.696	2.9551	3.2956	3.5862
Angular ID	0502	0503	0504	0505
PSNR (dB)	41.9746	34.988	33.1187	31.3126
RMSE	1.3578	2.8385	3.2109	3.6197
Angular ID	0602	0603	0604	0605
PSNR (dB)	41.2399	34.1652	32.28	31.073
RMSE	1.5256	3.1313	3.5052	3.7688
Angular ID	0702	0703	0704	0705
PSNR (dB)	41.5864	34.4899	32.3982	30.3772
RMSE	1.4209	2.9379	3.3659	3.7864
Angular ID	0802	0803	0804	0805
PSNR (dB)	40.4599	34.5469	32.1183	30.7085
RMSE	1.5341	2.8866	3.3293	3.6208
Angular ID	0902	0903	0904	0905
PSNR (dB)	41.9962	34.4988	32.5221	30.8233
RMSE	1.3524	2.9075	3.3043	3.6575
Angular ID	1002	1003	1004	1005
PSNR (dB)	39.9315	34.1226	32.3102	30.9432
RMSE	1.7786	3.0697	3.4578	3.7069

4.5 Chapter summary

In this chapter, the light field sensing is discussed and compared with conventional imaging. A single pixel infrared camera system based light field sensing is proposed to capture infrared light rays. By taking multiple angular images through coded aperture, we can use single CNT IR sensor to reconstruct full 4D light field from large object. The synthetic aperture photography will remove the IR focus problem away. This brings broadly application in virtual reality field. The double compressive sensing reduces sampling rate using the redundant of angular image difference. The experimental results show that the more angular images will achieve higher angular resolution, better refocus in synthetic aperture photography for single pixel IR cameras.

.

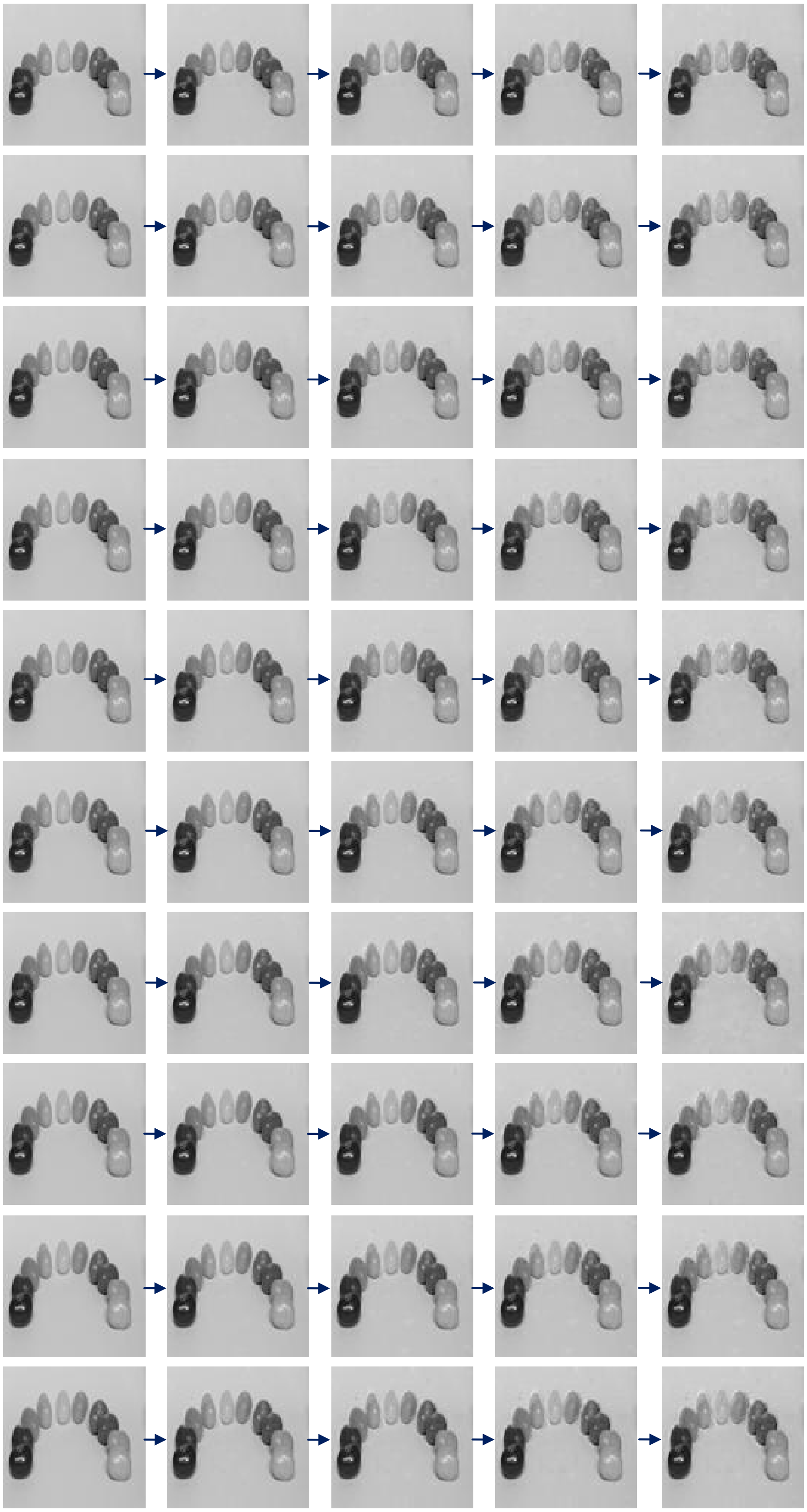


Figure 4.16 Angular image recovery from double compressive sensing, the first column is reference image and the other four are restored depends on its left.

Chapter 5

3-D Imaging

5.1 Previous Work

Three dimensional infrared cameras are potentially used in autonomous vehicles, robot manufacturers, security firms, industrial and video game manufacturers. All of these industries use either some sort of infrared imaging system or depth calculation. Especially with the development of autonomous vehicle and advanced robot industry, the three dimensional infrared camera becomes more and more important. Autonomous vehicles, including self driving cars, use several different methods of understanding their surroundings [132]. In general, there are two current techniques to achieve a three-dimensional image of an environment in market. The first is range distance cameras which use brute force computation to calculate the image based on single point scanning system [133]. It will produce a 2D image showing the distance to points in a scene from a specific point. Therefore, the problem with this system is that it is very difficult to obtain video rate imaging in a dynamic environment due to the computation intensity required to collect all the data for a single image. Besides, they are unable to distinguish what something they detect is. Anything of a close size (a deer or a parked car) looks similar in those imaging systems. Stereo photography is another 3D camera, which uses stereoscopic vision system, attempting to overlay the images collected by the two cameras to create a three-dimensional image. It makes the camera simulate human binocular vision [134] to capture three-dimensional scene. However, this technique is less than ideal due to problems with camera placement or inadvertent movement, calibration and other difficulties [135]. In addition, these systems are large and expensive. Due to the manufacturing limitation of large infrared sensor array, there is very little 3D IR camera based on binocular vision.

In the 3D camera market, it did have some 3D visible light cameras, which come into three different configurations: time of flight cameras, projected texture stereo vision cameras and phase shift cameras. Time of Flight (ToF) is one of effective ways to map the distance for 3D imaging. In [136] [137], the system combined ToF and standard RGB camera to do real-time 3D scene augmentation with virtual objects so that both 3D geometry and color information are matched. From on shelf market, time of flight cameras are the most common, which are typified by the SR4000 from Mesa Imaging. It's largely used in industrial inspection processes. However, it should be noted that the near-IR (or visible) light ToF camera is with the price approximately $\$10,000$. Projected texture stereo vision cameras are examples of 3D visible cameras and represented by the Ensenso N10. This camera has higher resolution than ToF configuration, and is capable of providing video images. Phase shift camera requires two CMOS sensor arrays, which increase calibration difficulty. They are represented by the Fujifilm Finepix Real3D camera, a commercially available system capable of taking 10 megapixel image. In overall, all three cameras' configurations are designed to work with visible light rather than infrared and are therefore incapable of operating in the mid-IR spectrum or longer wavelength.

There are at least three following factors considered when selecting a 3D camera, including operating in highly dynamic environments, camera resolution and price. It is undoubtedly that the infrared is better than visible camera for highly dynamic environments, because the infrared sensing can filter visible noise in daytime and works at nights. However, the infrared 3D camera needs high performance IR sensor, which introduces additional difficulties stemming from a cooling requirement. It requires cryogenic cooling to reduce the background noise especially for middle wave spectrum. Therefore, the 3D IR camera is not portable.

In this thesis, the low dimensional material based IR sensor and imaging method were proposed to solve these problems. In the proposal, a combination of distance measurement and stereo camera will allow a system to measure 3D coordinates and photorealistic texture at a video frame rate in the IR spectrum. The compressive sensing combines the advantages of the single-point technique with the advantages of the stereoscopic, e.g. it has a single-point detector so it does not require the difficult calibration as the stereoscopic technique. Besides, it is much less intensive computing than current single-point method as it uses a compressive sensing. As such, it is possible to capture video-rate 3D images that are unobtainable with the current generation of single-point systems. A critical requirement in compressive sensing imaging is that the detector has to switch multiple times instead of a single shot. Especially for a high resolution image, a detector is required to switch tens of thousands times to obtain an image. This requirement has made it extremely difficult to apply compressive sensing onto IR imaging. However, the nanomaterial based infrared photodetector provides the solution to these IR-specific problems.

5.2 Time of Flight 3D Imaging

5.2.1 Working Principle

The ToF method comes from depth measurement techniques. It refers to the passive process of measuring the depth of a scene by quantifying the changes that an emitted light signal encounters when it bounces back from objects in a scene. As shown in Figure 5.1, a 3D time-of-flight camera works by illuminating the scene with a modulated infrared light source and observes the distance to a 3D object by measuring the absolute time of which a light pulse needs to travel from a source into the 3D scene and back after reflection. Since the speed of light is constant and known, ($c = 3 * 10^8 m/s$), this is also called phase shift method which measures the phase shift between the

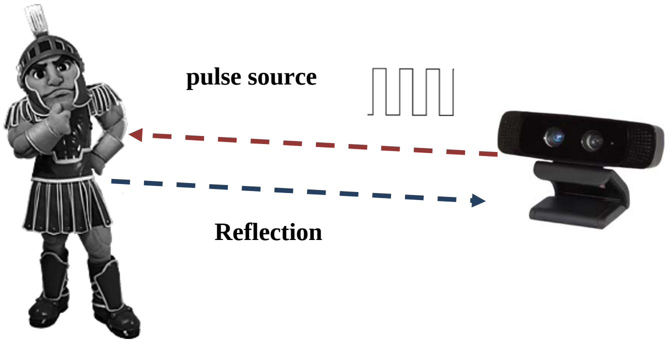


Figure 5.1 3D time-of-flight camera operation principle.

illumination and the reflection. In order to improve measurement accuracy, the illumination is typically from a solid state laser or a LED operating in the near-infrared range (850 *nm*) invisible to the human eyes and works in short bandwidth [138]. The receiver sensor is designed to respond to the same spectrum as irradiation which converts the photonic energy to electrical current. However, the ambient light also contributes to object reflection so that the depth accuracy will go worse (lower signal to noise ratio) with higher ambient component.

5.2.2 Time of Flight Modeling and Application

The phase shift determines the distance from desired object. In order to detect phase shift between the illumination source and the reflection, the light source is pulsed or modulated by a Continuous-Wave (CW) [138]. The modulation would be a sinusoid or square wave. The latter is more common because it can be easily realized using digital circuits. There are two ways to record time. The first is to integrate photoelectrons from the reflected light and another is to start a fast counter at the first detection of the reflection which requires a ultra fast photodetector. As shown in Figure 5.2, the phase difference is calculated by Eq. (5.1) from the relation between two different electric charge (left) or four different electric charge (right). The phase control signals have 90 degree delay from each other. Therefore, the distance can be calculated by Eq. (5.2).

$$\mathbf{D}_{Depth} = \frac{c}{2}\frac{\Delta\varphi}{2\pi f} \tag{5.1}$$

$$\mathbf{d} = \frac{1}{2}c\Delta t(\frac{Q_2}{Q_1 + Q_2}) \tag{5.2}$$

On contrary, the continuous wave modulation method retrieves phase shift by demodulation of received signal, which crosses correlation of received signal with emitted signal. For an emitted sinusoidal signal $g(t)$ with modulation frequency ω, the received signal could be $s(t)$ after reflection from 3D surface, shown in Eq. (5.5) and Eq. (5.6), (b: constant bias, a: amplitude, ϕ: phase shift). The cross correlation of $g(t)$*$s(t)$ is $c(\tau)$ in Eq. (5.7). By sampling $c(\tau)$ at four sequential instants with 90 degree phase offset to get $Q_i = c((i-1)\cdot\pi/2), i = 1,...,4$, the four resulting electric charge can be used to estimate the phase difference $\Delta\varphi$ as Eq. (5.3), where the distance in Eq. (5.4), c is the speed of light in constant. It also directly obtains the parameters a and b in Eq. (5.8).

$$\Delta\varphi = arctan(\frac{Q_3 - Q_4}{Q_1 - Q_2}) \tag{5.3}$$

$$\mathbf{d} = \frac{c}{2f}\frac{\Delta\varphi}{2\pi} \tag{5.4}$$

$$g(t) = cos(\omega t) \tag{5.5}$$

$$s(t) = \mathbf{b} + \mathbf{a}cos(\omega t + \phi) \tag{5.6}$$

$$c(\tau) = s * g = \int_{-\infty}^{+\infty} s(t)\cdot g(t+\tau)dt = \frac{\mathbf{a}}{2}cos(\omega\tau + \phi) + \mathbf{b} \tag{5.7}$$

$$\begin{aligned} a &= \frac{\sqrt{(Q_4 - Q_2)^2 + (Q_3 - Q_1)^2}}{2} \\ b &= \frac{Q_1 + Q_2 + Q_3 + Q_4}{4} \end{aligned} \tag{5.8}$$

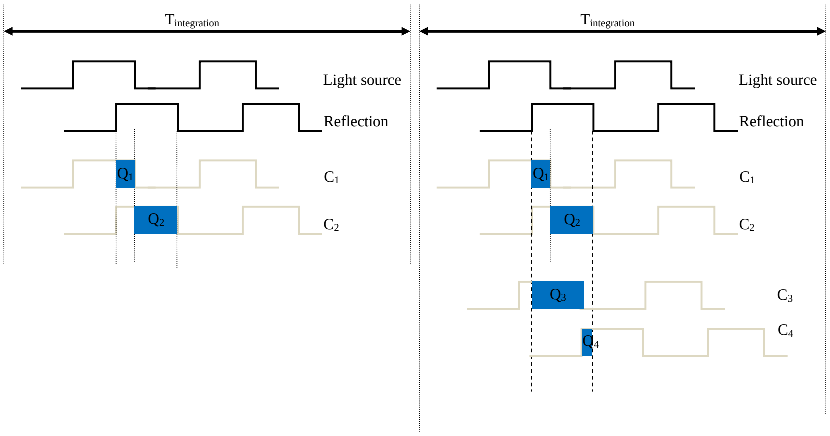


Figure 5.2 Two time-of-flight methods: pulsed (left) and continuous wave (right).

In CW 3D depth measurement, the reflected amplitude (**a**) is a function of the optical power. The offset (**b**) is a function of the ambient light and residual system offset. The high amplitude, high modulation frequency and high modulation contrast will increase accuracy while high offset can lead to saturation and reduce accuracy. At high frequency, the modulation contrast will attenuate due to the physical property of the silicon, which puts a practical upper limit on the modulation speed. The advantages of pulsed modulation ToF method is direct measurement, where the irradiation light pulse energy is tens order higher than background illumination to reduce background noise. Most importantly, the direction of illumination and observation are collinear.

The most obvious drawback is long time measurement due to scanning point by point. Meanwhile, the light pulse inaccuracy comes from light scattering on distinct surface. From a hardware perspective, it is difficult to generate short light pulses with fast rise/fall time and low repetition rate. The continuous wave modulation is applicable to variety of light sources due to no short and strong pulse required. Besides, the long integration time introduces motion blur and limits the frame rates. The measurement accuracy is dependent on the power of the emitted IR source and the material of target surface.

5.3 Stereo Vision 3D Imaging

5.3.1 Working Principle

The stereo vision imaging comes from human binocular vision which requires two or more lenses with separate sensor. The binocular vision has distinct advantages compared to single eye. It gives precise depth perception by two eyes and allows a creature to see more of, or all of, an object behind an obstacle [139]. With two eyes, there would be overlapping of vision from same scene, which introduces a slightly different viewpoint between it. As a result, binocular vision provides depth information by calculation. In human eye imaging, the differences is provided as binocular disparity to brain so as to calculate the depth of visual scene, including the nature of the stimulus and brain process.

In binocular vision, the retinal disparity, or the separation between objects as seen by left and right eye, are key values to evaluate depth. As shown in Figure 5.3, the distance between two eyes is almost always 6.5 *cm* for adult. The object image in left eye and right eye is slightly different. This retinal disparity will provide the relative depth of object due to no referral distance. It might be argued that people could estimate the distance by eyes. This is because there is a calibration in brain when seeing the object. Since the 3D image is highly dependent on the image difference, it is very easy to understand that the closer objects generate the smaller retinal disparity [140].

5.3.2 Stereo Vision Modeling and Application

Stereo vision generally uses two cameras separated by a distance, in a physical arrangement similar to the human eyes. As shown in Figure 5.4, the left shows the simple pin hole camera model and this is also the basis of laser triangulation and radar distance measurement. The α and β are computed by two distinct images. The depth z can be calculated, as Eq. (5.9).

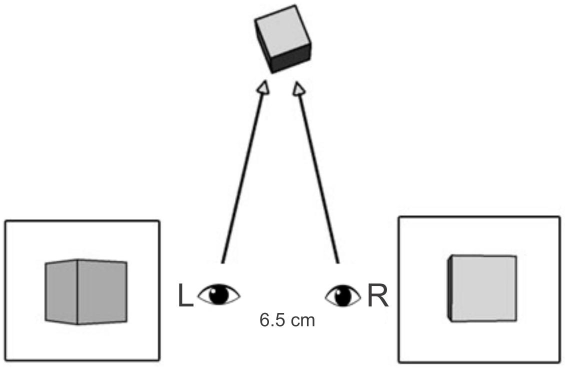


Figure 5.3 Retinal disparity from eyes.

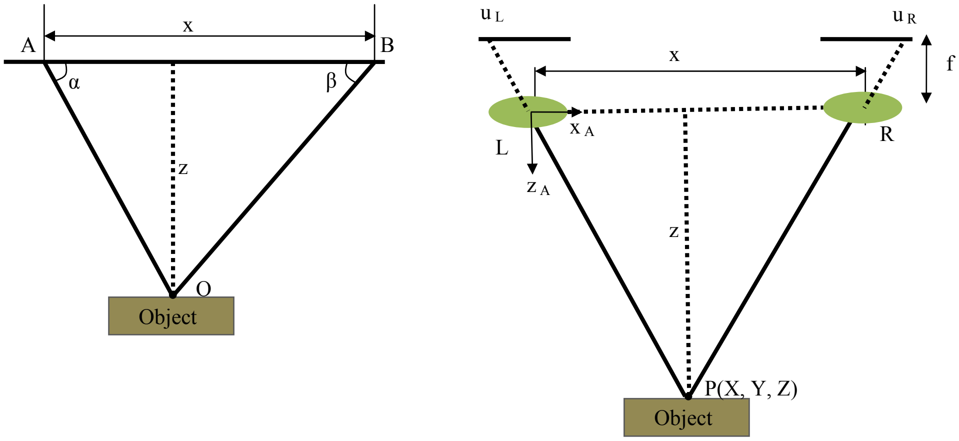


Figure 5.4 Stereopsis depth through disparity measurement (left) and simplified stereo vision system (right).

$$z = \frac{x}{\frac{1}{tan\alpha} + \frac{1}{tan\beta}} \tag{5.9}$$

Figure 5.4 right shows the diagram of a simplified stereo vision setup. The both cameras are mounted exactly parallel to each other, and with the exact same focal length, where

x: The baseline, or the distance between two cameras.

f: The focal length of camera.

P: The real world point defined by the coordinates X, Y and Z.

u$_L$: The projection of P in image plane by left camera.

u$_R$: The projection of P in image plane by right camera.

The model locates in 2D cartesian coordinates (x$_A$, z$_A$) where the distance between two cameras are x, the point P(X, Y, Z) are at u_L and u_R in image plane respectively. By acquiring two distinct images, the X-coordinates of the point u_L and u_R can be given in Eq. (5.10). The disparity refers to the distance between u_L and u_R. It is obviously to calculate depth by Eq. (5.11), though an actual stereo vision set-up is more complex. The same fundamental principles still apply anyhow.

$$\begin{aligned} u_L &= f * X/Z \\ u_R &= f * (X - b)/Z \end{aligned} \tag{5.10}$$

$$D_{Depth} = f * \frac{b}{u_L - u_R} = f * b/z \tag{5.11}$$

The major challenge of stereo vision is the correspondence calibration which involves complex, computationally intensive algorithms for feature extraction and matching. Compared with ToF method, the stereo vision implementation cost is very low, as most common off-the-shelf cameras can be used. This technique is intuitive presentation to humans so that both human and machine are looking at the same images. In stereo vision, the error of depth resolution is a quadratic function of the distance. However, the stereo vision infrared 3D camera is very uncommon due to price, resolution, temperature limitation. In this thesis, a compressive 3D imaging sampling was proposed for infrared stereo camera design.

5.4 Compressive 3D Imaging

5.4.1 Sparsity in 3-D Image

The single pixel camera system has been introduced based on compressive sensing to overcome the current limitation and challenge in manufacturing large scale photosensor arrays [54]. It is more attractive by using a low dimensional materials based infrared sensors. Compressive 3D imaging will both have the advantage of ToF sensor and stereo vision method.

The most important issue is to find redundant and duplicate information or sparsity in 3D image so as for compressive sensing. Generally, it is proved that most natural digital images are sparse [55]. For example, the background of a image may have many pixels with the same color and texture information. Much of this redundant information ends up being discarded during the compression process, making these high resolution cameras very inefficient. In case of a nonsparse signal, a sparsity process is required to transform the non-sparse signal into a spare signal by using some special basis, such as wavelet, curvelet and Fourier [141], as shown in Eq. (5.12), where s is the sparse representation of the non-sparse signal in basis, y is the observation.

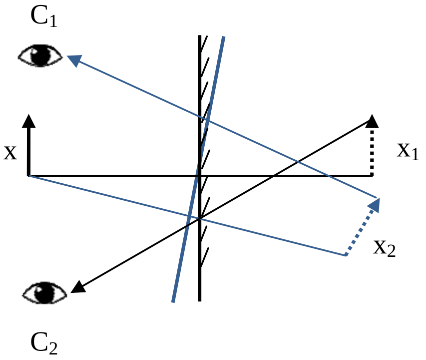


Figure 5.5 Sparsity in DMD based 3D sampling.

$$\begin{aligned} x &= \Psi \times s \\ y &= \phi \times \Psi \times s \end{aligned} \tag{5.12}$$

In digital micromirror device based single pixel camera sampling system, there is another redundant information during sampling except for the image itself. In Chapter 3, it already discussed how DMD worked. As shown in Figure 5.5, when the object 3D scene (x) locates in front of DMD, there will be two distinct reflected images from DMD “ON” state and “OFF” state, x_1 and x_2. It can be seen as the two images of stereo vision. During each sampling process, when the DMD pattern is D, the C_1 and C_2 can be derived as Eq. (5.13). C_2 is dependent on C_1 for a desired scene so that the sampling would be redundant. By introducing anther IR sensor, it will generate second image for stereo vision system.

$$\begin{aligned} C_1 &= DX_1 \\ C_2 &= (I-D)X_2 \end{aligned} \tag{5.13}$$

5.4.2 Compressive 3D Sampling

The prototype 3D camera is silicon infrared sensor based. In a 2D single pixel camera, it will consist of a group of lenses, DMD and photodetector, along with some associated optics and hardware. There are two approaches for developing a 3D camera using DMD, as shown in Figure 5.6 and Figure 5.7. In the first design, the camera will detect IR emission from two distinct angular mirrors via the DMD status "ON" and "OFF". When the DMD is "ON", the reflected light will be focused on sensor "I" from the upper lens. When the DMD is "OFF", the infrared light passes through the lower lens into sensor "II". The difference between DMD "ON" and "OFF" status is an angular change of 12 degree. This way, light incident to sensors "I" and "II" will provide angular information about the object being imaged. After the light is captured by the photodetector, the information flows to a readout circuit and ADC, which will converter analog photocurrent into digital values. The image recovery is based on sampled values and red-cyan color model is applied to build uncalibrated stereo image.

There is another 3D camera design, shown in Figure 5.7, named as the mask configuration. In this optics design, there is only one IR sensor, while an electrically controlled mask is placed before the detector to get two angular images. The object will be projected on to the DMD plane through the main lens. When the mask opens the upper window, only the light that is incident to upper window will be let through to the detector. This allows the light to serve as one of the two images required to obtain a 3D scene. Likewise, the same process takes place when the lower window opens, allowing the system to have both images necessary to reconstruct a 3D image falling on a single detector. The following readout circuits and 3D modeling will be same as Figure 5.6.

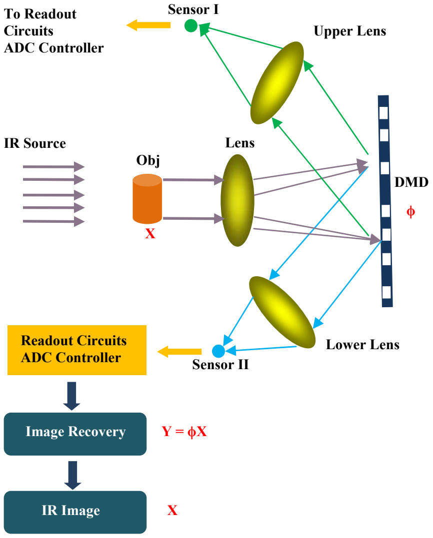


Figure 5.6 Dual detectors 3D imaging system.

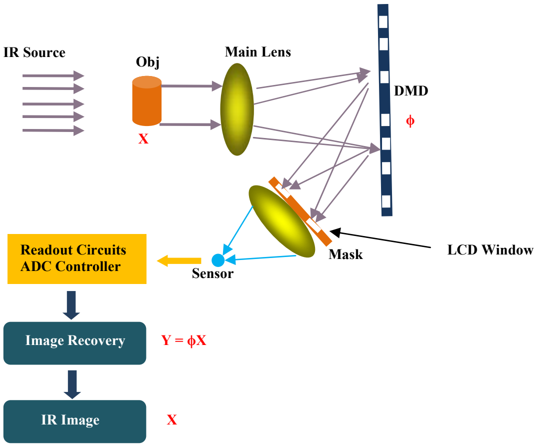


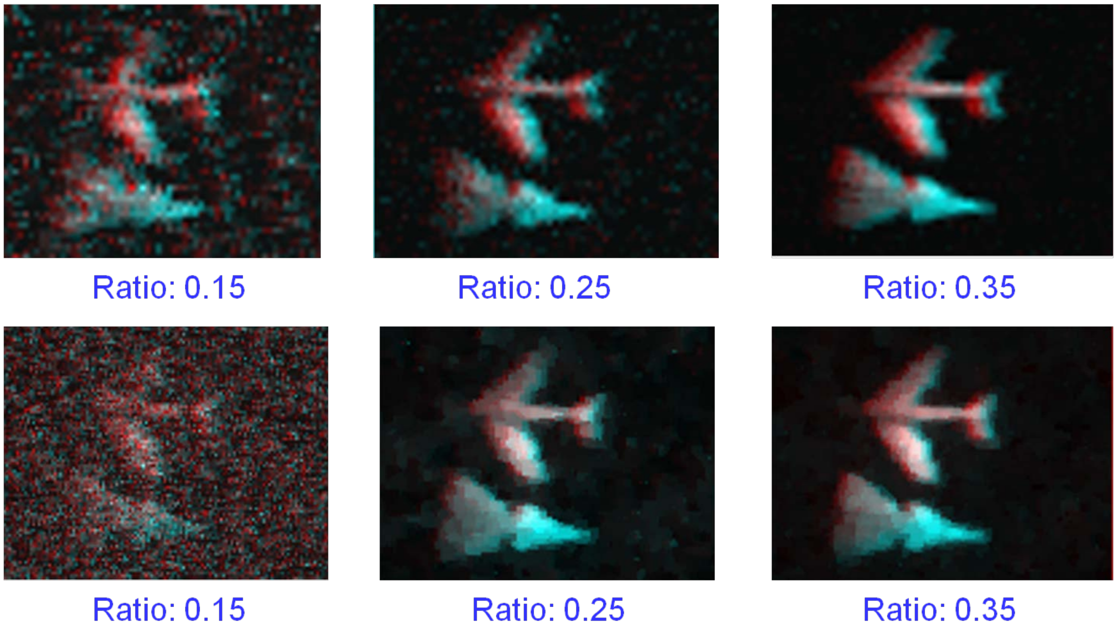
Figure 5.7 Mask based 3D imaging system.

5.4.3 Experiments with Prototype Camera

The experimental results demonstrate the generation of stereoscopic images using the desired imaging system. Figure 5.6 depicts the illustration of the 3D imaging system which captures 3D scenes. Two photodetectors are employed and their positions are aligned to the reflected signal from DMD, so that the reconstructed image on each detector represents the directional information. In the experiment, an infrared light source (830 *nm*) produces illumination on the target scene. When the signals are reflected from the digital micromirror devices, a series of photocurrents are generated on each detectors, which are acquired by a high-speed readout circuitry. Finally, images are reconstructed based on the compressive sensing algorithm. As a result, two images can be obtained from two detectors simultaneously and they are considered as sub-aperture images for 3D computation. The resolution of the recovered images is 64×64 pixels, with 30% sampling ratio. It is noted that these two sub-aperture images also represent two different perspective views of the source signal. Therefore, a stereoscopic image can be computed and generated as shown in Figure 5.8. There are two noise sources, including sampling noise and recovery error of compressive sensing.

5.5 Chapter summary

The development of these two systems solves the problems confronting the current generation of 3D IR sensors. These problems include the computation time and calibration issues associated with current 3D image systems and the cryogenic cooling system required by most current IR sensors. The compressive sensing algorithm developed will be capable of decreasing the computational time required for imaging to a point where the image itself can be updated at a video frame rate of 30 Hz. The algorithm can be adjusted to ensure photorealistic quality of environments and enable



Ratio: 0.15 Ratio: 0.25 Ratio: 0.35

Ratio: 0.15 Ratio: 0.25 Ratio: 0.35

Figure 5.8 3D image reconstruction in red/cyan color.

precise depth calculations. The use of CNT photodetector and its associated signal enhancement that are developed in chapter 2 will eliminate the need for the cryogenic cooling system. To conclude, the proposed 3D imaging system is based on the integration of compressive sensing and binocular vision. By introducing two photodetectors into the system, two sub-aperture images can be obtained from two different perspective views. Current experimental results indicated that infrared images can be recovered based on this principle and the images can be used to compute stereoscopic photography. It opens the door to the 3D imaging through investigating the new camera architecture using compressive sensing and computational light-field.

Chapter 6

Super Resolution Imaging

6.1 Previous Work

The High Resolution (HR) imaging are required and necessary for many electronic imaging applications because HR image offers more details of object, especially for medical image diagnosis, pattern recognition in computer vision. The most directly way to enhance resolution is to increase the chip size (together with lens changes) or reduce the pixel size to increase the number of pixels per unit area. Both techniques are dependent on microfabrication manufacturing. They are limited by device fabrication, price and sensor sensitivity. The disadvantage of former method is to increase capacitance which makes the readout circuit in low charge transfer rate. It also requires a high precision optics [142]. For the latter solution, reducing the pixel size will also decrease the amount of light coming to each sensor so that the shot noise will degrade the image quality. The minimum active pixel area is around 40 μm^2 for a $0.35 \mu m$ CMOS process [143]. However, the resolution also can be enhanced using signal processing (referred to super resolution). The recovered high resolution image is obtained from multiple low resolution images observed. As shown in Figure 6.1, the red points represent pixel center of reference Low Resolution (LR) image. The blue and green point show two LR images with subpixel shift for the same scene. The three low resolution images are subsampled (aliased) as well as shifted with a subpixel. In order to increase resolution, the LR images must have subpixel shifts or else there is no new information for image reconstruction. This could be obtained from one camera with several captures or from

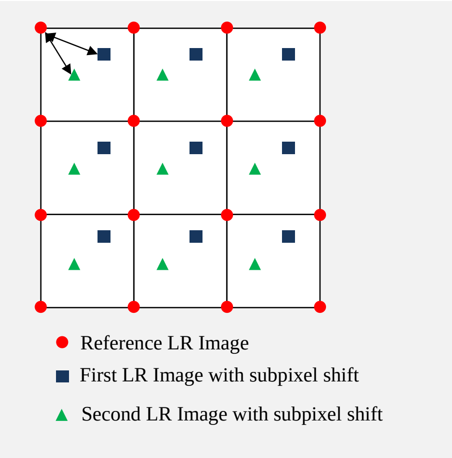


Figure 6.1 Multiple images super resolution.

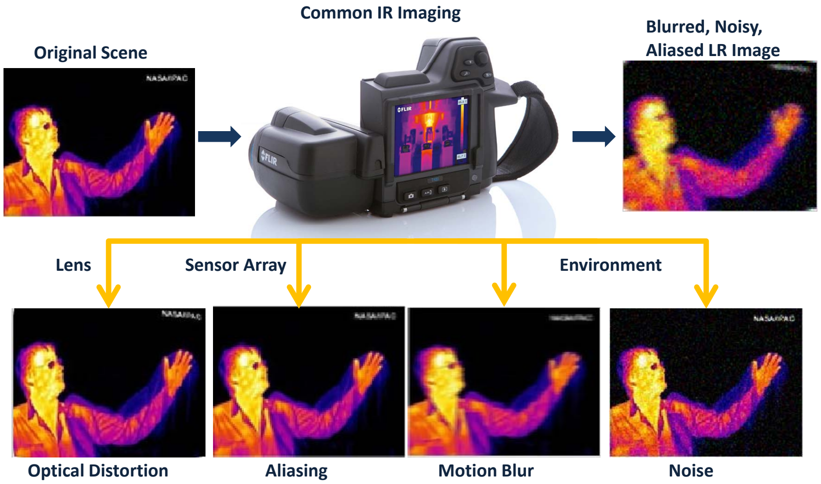


Figure 6.2 Four causes of LR image acquisition.

multiple cameras located at different positions. In conventional camera (visible, infrared SLR camera), there are at least four factors inducing LR image, including optical distortions (out of focus, diffraction limit, etc.), motion blur due to limited shutter speed, noise that occurs within the sensor or during sampling. As shown in Figure 6.2, the recorded image usually suffers from distortion, blur, noise, and aliasing effects.

There are two main categories of Super Resolution (SR) method, including multiple images (frames) and single image based. As shown in Figure 6.1, the multiframe SR reconstruction has longest history and it involves three steps, image registration, interpolation and restoration. The most difficulty is to estimate the motion of LR input frames corresponding to reference frame, referred to image registration. The typical method is to look up for interest points in the low-resolution image set, then use robust methods to estimate the point correspondences and compute homographies between images. In [144], the iterative method was used to estimate registration parameters, shifts and rotation. The block matching is also applied to register input images in [145]. The Bayesian method searched a continuous space of shift and rotation together with conventional MAP reconstruction algorithm to estimate the high resolution. The dense and scene illumination changes are also used to estimate motion of each pixel in [146], because the photometric will be changed as well when scene illumination changes. The zoom and defocus approaches are discussed in [147] and [148] to build up the constraints for recovering super resolution image, where the registration and estimation could be in a joint framework. In [149], the joint MAP estimation algorithms capture the dependence between LR image registration and HR image estimation although it may introduce overfitting problem. All above approaches are explicit motion estimation which is critical to SR reconstruction. Besides, the fuzzy motion estimation based on block matching is also used in denoising algorithm in [150].

Compared to multiple images based super-resolution, single image based is more applicable since there is only one low-resolution image required, especially for portable applications. The most straightforward way of single image super resolution is interpolation. The nearest neighbor interpolation model the unknown point by its nearest neighbor point [151]. For each point on the HR grid, the closest known LR pixel is selected as the value at the point in HR grid. The disadvantage is jaggy effect on the HR images. Bicubic interpolation utilized a cubic kernel to interpolate but created blur effect [152]. The edge directed interpolation method is presented in [153] for super resolution. It estimates the local covariance coefficients from a single low resolution image, and applies the same coefficients to reconstruct high-resolution. The contourlet transform and wavelet based linear interpolation is proposed in [154]. The directional filtering and data fusion were used to edge guided nonlinear interpolation to preserve sharp edges and reduce ring artifacts in [155]. The another single image based super resolution is statistics methods. It utilizes statistical edge dependency information as in low and high resolution images in [156]. The learning based method operates by building a model from example, learns from and makes predictions on data. It is a powerful tool for image super resolution, although it requires two large training data sets. In order to get super resolution, a profile of image gradient was described as the shape and sharpness in recovery algorithm [157]. In [158], it proposed a method via an example-based strategy which divided the high-frequency patches of a low-resolution image into different classes. The signal sparsity and learning methods were combined together in [159]. The low resolution images were represented in sparse domain by computing corresponding coefficients, and the high resolution images are generated via these coefficients.

In image reconstruction, it requires high computational effort, even for problem of moderate size, especially for real time processing. The super resolution becomes more attractive in single pixel imaging system due to time complexity of compressive sensing. Many efficient algorithms have been developed for solving this minimization, including linearized Bregman method, runtime 10 s for 4096 data set [160], fixed point continuation method, 6.2 s for 4096 data set [161], Bregman iterative method [162]. Although a fast coordinate descent method was proposed in [163], of 0.94 s for 4096 data set, it was specialized for sparsity data, however it might not applicable to image recovery. Fourier domain computation was introduced to replace iterative linear solvers in MRI reconstruction via $\ell 1$ minimization, which cost 11.5 s to reconstruct of 256×256 image in Matlab on 1.2 GHz laptop computer in [164]. The problem was that the runtime would increase to 346 s for 1024×1024 uterus image.

Another approach is to get fast recovery via hardware improvement on FPGA [165]. It integrates orthogonal matching pursuit (OMP) and approximate message passing (AMP) in FPGA to get 0.63 *ms* for 32×32 image block recovery. Due to hardware limited, it is still hard to do 1024×1024 image size. The Total Variation (TV) norm image restoration makes the recovered image quality sharper and more accurately by preserving the edges or boundaries. It works when the gradient of the image is sparse. Most images can be well approximated by TV norm so that there are huge TV minimization methods proposed. The weakness of these algorithms are still either too slow or less robust compared with $\ell 1$ minimization algorithm, especially for large image size [128]. In order to produce image with super resolution, high image quality in single pixel camera system, it will cost more than 10 *s* for 256×256 image in 1.4 GHz desktop using TVL3 algorithm. Moreover, the runtime will increase dramatically when image size goes high.

The combined super resolution and compressive sensing attracts more interesting in infrared spectrum than visible image. The higher resolution IR images means better ability to spot small targets at longer distances. It is hardly to directly reconstruct high resolution IR image from compressive sensing, such that the signal processing approach would be best option for IR image super resolution.

6.2 Observation Modeling

The digital imaging system suffers form hardware limitations, acquiring images with various kinds of degradations, as shown in Figure 6.2, including the optical distortion, the motion blur due to aperture time, the sensor noise. Finally the frames captured by the low resolution imaging system are blurred, aliasing, and noisy version of the underlying true scene.

$$\begin{aligned}\mathbf{Y}_k = D_k B_k M_k \mathbf{X} + n_k, k = 1,2,...,K \\ \mathbf{X} = [x_1, x_2, ..., x_N]\end{aligned} \tag{6.1}$$

As shown in Eq. (6.1), $\mathbf{X} = [x_1, x_2, ..., x_N]$ denotes the desired high resolution image, sampled above Nyquist sampling frequency. $\mathbf{Y}_k$ is the k^{th} LR image with subpixel shift from reference LR image ($k = 1$). All vectors are represented in lexicographical order, and if $M = 1$, it is single image super resolution observation model. M_k describes the motion information for k^{th} LR image. B_k is the blur models effect and D_k is the down sampling operator. n_k represents the noise model. The linear equation is in Eq. (6.2). The k^{th} LR image is denoted as $Y(s,t), (s = 1,2,...N_1, t = 1,2,...N_2)$, N_1, N_2 are horizontal and vertical direction resolution of LR images. The parameters L_1, L_2 are down sampling factors for each direction respectively ($N = L_1 N_1 L_2 N_2$). In mathematically, the motion warp matrix size is $L_1 N_1 L_2 N_2 \times L_1 N_1 L_2 N_2$, blur matrix B_k is $L_1 N_1 L_2 N_2 \times L_1 N_1 L_2 N_2$, while subsampling matrix D_k is $(N_1 N_2)^2 \times L_1 N_1 L_2 N_2$ size.

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_K \end{bmatrix} = \begin{bmatrix} D_1 B_1 M_1 \\ D_2 B_2 M_2 \\ \vdots \\ D_K B_K M_K \end{bmatrix} \mathbf{X} + \begin{bmatrix} n_1 \\ n_2 \\ \vdots \\ n_K \end{bmatrix} \tag{6.2}$$

As shown in Eq. (6.1) and Eq. (6.2), the motion warp matrix represents the global or local translation and rotation. This can be estimated by the difference between reference and particular LR images. It is an interpolation method when the fractional unit of motion is not equal to the HR sensor grid. The blur matrix are very hardware related, which depends on optical system, focus, diffraction limit, aberration. It is usually a known value. The downsampling matrix D_k models the aliasing in LR image. These three matrix are correlated to super resolution process closely.

6.3 Multiple Images based Super Resolution

6.3.1 Nonuniform Interpolation Approach

Currently most of the non-uniform interpolation based super resolution image reconstruction methods proposed in the literature consist of the three stages: registration, interpolation and restoration. As the crucial step to the success of the super resolution, the displacement between two pixels is calculated as given information. The sampled angular images are irregularly data, lack of high frequency components and in presence of noise from the optical system.

The simplest way to interpolate a HR image is nearest neighbor interpolation. All LR images are registered in its position, shown in Figure 6.1. It interpolates pixel by its nearest neighbor point. For each point on the HR grid, the closest known LR pixel is selected and used as the value at the point in HR grid. This has lower complexity but with jaggy effect. The most applicable non-iterative algorithm is gradient adaptive interpolation. The inserted pixel value is influenced by the local gradient of pixels, where all LR image motions will be estimated by distinct algorithms. According to the motions, the low resolution images are mapped to the uniform high resolution grid. The three nearest pixels around the interpolated HR pixel will be selected by Euclidean distance between the desired pixel and its neighboring pixels. The interpolated gray values can be calculated by Eq. (6.3), where $f(i,j)$ is the interpolated point, $f(i_k, j_k)$ k= 1,2,3 are its three nearest points. $W(i_k, j_k)$ is the gradient weight function, and m is a positive value, μ is a positive value close or smaller than 1. $f_x'(i_k, j_k), f_y'(i_k, j_k)$ are vertical derivative and horizontal derivative. The gradient-based adaptive interpolation takes into account of local gradient. The smaller the local gradient of a pixel is, the more contributions it should have on the interpolated pixel [166].

$$
\begin{aligned}
f(i,j) &= \frac{\sum_{k=1}^{k=3} W(i_k,j_k)S(i_k,j_k)f(i_k,j_k)}{\sum_{k=1}^{k=3} W(i_k,j_k)S(i_k,j_k)} \\
S(i_k,j_k) &= (1-\Delta x(i_k,j_k))(1-\Delta y(i_k,j_k)) \\
\Delta x(i_k,j_k) &= |i_k - i| \\
\Delta y(i_k,j_k) &= |j_k - j| \\
W(i_k,j_k) &= (-\mu G(i_k,j_k)+1)^m \\
G(i_k,j_k) &= \frac{|f'_x(i_k,j_k)|+|f'_y(i_k,j_k)|}{2\sqrt{(f'_x(i_k,j_k))^2+(f'_y(i_k,j_k))^2}}
\end{aligned}
\tag{6.3}
$$

6.3.2 Frequency Domain Approach

Since the multiple frames super resolution approach reconstructs the HR image from LR images, there must be a relationship between LR images and desired HR image. The frequency domain method utilizes this relationship, which is based on:

1) The shifting property of Fourier transform shift theorem. $\mathscr{F}\{f(t-t_0)\}(s) = e^{-j2\pi st_0}F(s)$.

2) The aliasing relationship between the Continuous Fourier Transform (CFT) of an original HR image and the Discrete Fourier Transform (DFT) of observed LR images.

3) There is an assumption that an original HR image is bandlimited.

The first two principles are critical to frequency domain super resolution because it builds the connection between the fourier transform of reference HR and its subpixel shift or the transformation of HR and its downsampling signals. Given a reference continuous HR image $x(t_1,t_2)$, the corresponding fourier transform is $X(w_1,w_2)$. When a subpixel LR image is taken, it is downsampling of an k^{th} shifted image from $x(t_1,t_2)$, defined as $x_k(t_1,t_2) = x(t_1+\delta_{k1}, t_2+\delta_{k2})$ in spacial domain (t_1, t_2 are representing high resolution). By the frequency transform shift theorem, the continuous fourier transform of shifted image $X_k(w_1, w_2)$ can be derived as Eq. (6.4).

$$X_k(w_1, w_2) = exp[j2\pi(\delta_{k1}w_1 + \delta_{k2}w_2))]X(w_1 + w_2) \tag{6.4}$$

All k^{th} shifted images can be expressed as Eq. (6.4) in frequency domain. Based on inverse Fourier transform, the k^{th} LR image will be Eq. (6.5). By comparing Eq. (6.4) and Eq. (6.5), the discrete fourier transform of LR images and CFT of HR image is connected and modeled as $Y = \Phi X$, such that the HR image super resolution becomes an linear inverse problem.

$$r_k(t_1, t_2) = \frac{1}{T_1 T_2} \sum_{n_1=0}^{L_1-1} \sum_{n_2=0}^{L_2-1} \times (\frac{2\pi}{T_1}(\frac{t_1}{N_1} + n_1), \frac{2\pi}{T_2}(\frac{t_2}{N_2} + n_2)) \tag{6.5}$$

6.3.3 Regularized SR Reconstruction Approach

The non-uniform interpolation and frequency domain approach are both non-iterative algorithm for super resolution. However, the iterative super resolution is also attractive due to the excellent performance. As shown in Eq. (6.1), the observation model could be simplified as Eq. (6.6). The deterministic regularized approach solve this problem by the prior information, which can be converted to optimization by choosing a x to minimize the lagrangian in Eq. (6.7). For least square prior knowledge, it suggests that most of images are naturally smooth with limited high frequency values. In Eq. (6.7), α represents the lagrange multiplier, commonly referred to the regularization parameter. This is critical in reconstructing HR image since the larger value of α will generally lead to a smoother HR image. The least square approach works for a small number of LR images available (or the problem is underdetermined) or the observed data has too much noise due to registration error and noised sampling. On the other hand, if a large number of LR images are available and the amount of noise is low, small α will lead to a good solution. The HR recovery can be converted to an iteration problem as Eq. (6.8). The convex problem is differentiable with quadratic regularization, where β represents the convergence parameter.

$$\mathbf{Y}_k = W_k\mathbf{X} + n_k, k = 1,2,...,K \tag{6.6}$$

$$min\sum_{k=1}^{p} ||\mathbf{y}_k - \mathbf{W}_k\mathbf{x}||^2 + \alpha||C\mathbf{x}||^2 \tag{6.7}$$

$$\begin{gathered}(\sum_{k=1}^{p} \mathbf{W}_k^T\mathbf{W}_k + \alpha\mathbf{C}^T\mathbf{C})\hat{\mathbf{x}} = \sum_{k=1}^{p} \mathbf{W}_k^T\mathbf{y}_k \\ \hat{\mathbf{x}}^{n+1} = \hat{\mathbf{x}}^n + \beta(\sum_{k=1}^{p} \mathbf{W}_k^T(y_k - W_k\hat{\mathbf{x}}^n) - \alpha C^T C\hat{\mathbf{x}}^n)\end{gathered} \tag{6.8}$$

6.3.4 Multiple Frames Sampling in Single Pixel Camera

In conventional multiple frames super resolution, the LR images are taken by multiple cameras or single camera in distinct position/time. However, the compressive multiple frames sampling can be implemented by simple optics design. As shown in Figure 6.3, the basic single pixel camera is composed of one main lens, digital micro mirror device (working as spatial light modulator), lens for converging light to single sensor, and IR sensor. The imaging system working process is discussed in Chapter 3 and Chapter 4.

In order to get high resolution image, multiple position angular images are captured in single pixel camera. As shown in Figure 6.4, when the object ‘S’ locates in front of mirror, there will be a virtual image with reflection symmetry. When there are two imaging systems (eyes) capturing it

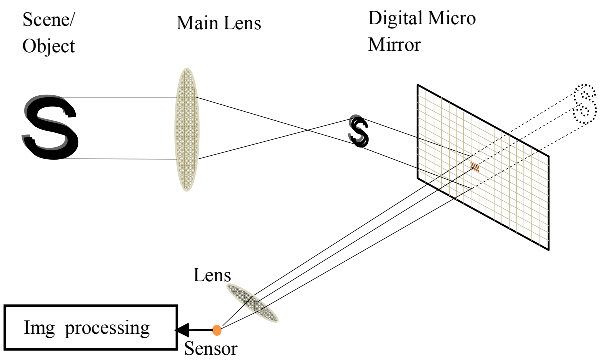


Figure 6.3 Schematic diagram of single pixel camera.

from distinct directions, each micro mirror will correspond to different area on two angular images. The micro mirror in single pixel camera is similar as one pixel in sensor array. If all micro mirrors are continuously as a complete plane, the two separate images are exactly same only with image translation (the reflections are same for different angular), as shown in Figure 6.4 right. There are two pin holes of A1 and A2 which generate two distinct recovery images di1 and di2 and the image displacement can be calculated by Eq. (6.9).

$$\Delta d = \frac{d0}{d0 + da} * \Delta x \tag{6.9}$$

As shown in Figure 6.5 right, four different mirrors (in different color) will reflect different part of light rays to single sensor through two pin holes. For example, the left end point will map to different mirror through aperture A1 and A2, which contributes to different image pixels. The more angular images are captured, the higher resolution will be achieved, because there are more subdivisions of object than one pin hole recovery. In this research, 16 different angular images are sampled, as shown in Figure 6.5 left and each aperture pin hole size is close to mirror size in 0.5 *mm*. There are 16 measurements corresponding to 16 angular images, of which the difference are within sub pixel. In experimental setup, an electrical controlled aperture mask is put between

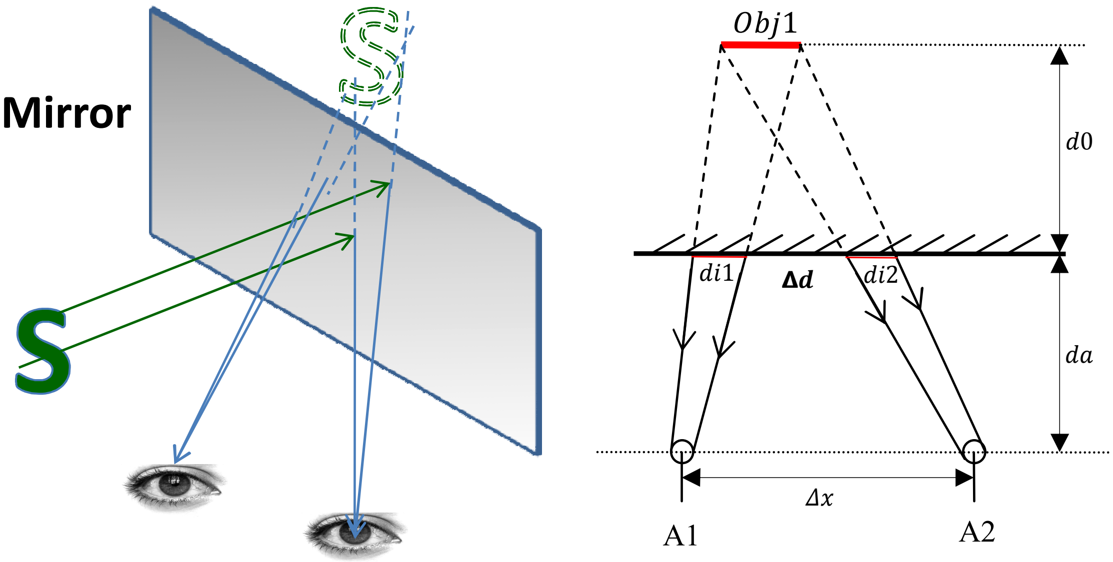


Figure 6.4 Two apertures design in DMD imaging system.

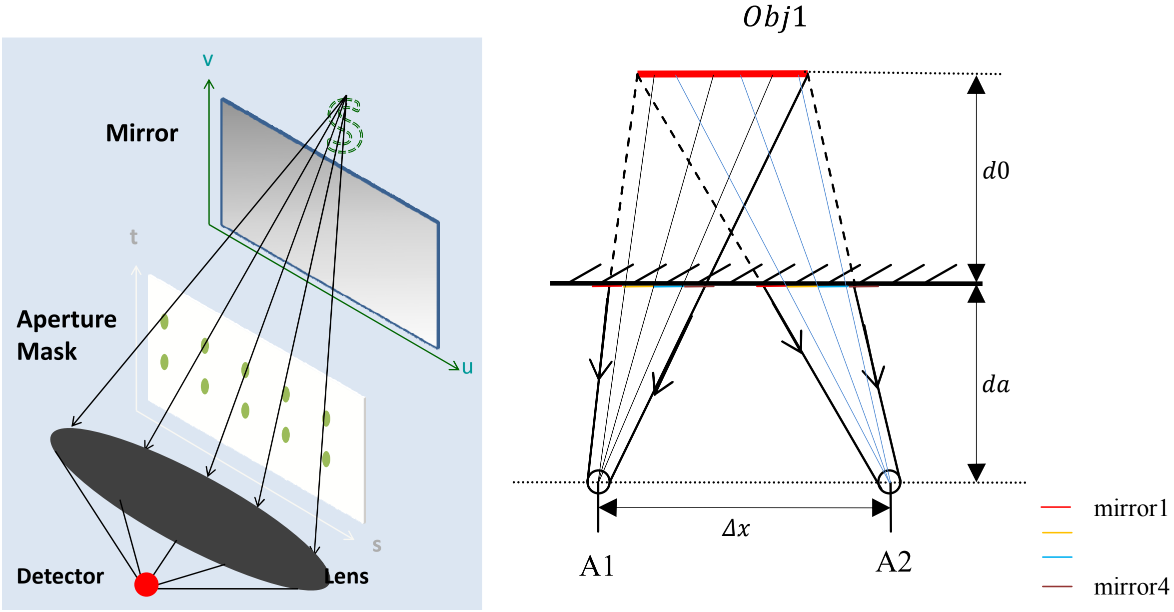


Figure 6.5 Multiple apertures design for high resolution imaging.

DMD and lens to control subpixel difference. By this way, the multiple frames sub-aperture image compressive sampling is implemented.

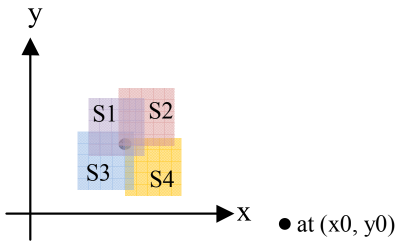


Figure 6.6 Measurement covered in neighborhood of X_0 point.

6.3.5 Experiments with Prototype Camera

Due to the limitation of computation and accuracy of compressive sensing system, the acquired images are in low resolution. In this experiment, the continuously gradient based interpolation model was implemented in single pixel IR camera system. The registration in compressive sensing camera is the crucial step to the success of the super resolution image construction. In DMD based single pixel camera system, the registration is easier than SLR camera and the displacement between two pixels within any different angle images is calculated as given information so that the registration is linear and easily obtained. In this research, the continuous gradient based interpolation also named surface interpolation is introduced as model. As shown in Figure 6.6, within a local neighborhood centered at $S_0(x_0,y_0)$, the intensity value at position $S(x+x_0,y+y_0)$, $f(s,s_0)$ is approximated by a polynomial expansion with basis, shown in Eq. (6.10). However, as shown in Figure 6.6, the measurement for each pixel covers a large area in low resolution. The measurement will be a integration on specific area, shown in Eq. (6.11), where the observation matrix is **I**. In order to solve for the projection coefficient P for point S_0, an extra constraint needed. In this research, least squares estimation or L_2 minimization was applied to get unique solution.

$$\begin{aligned} f(s,s_0) &= p_0(s_0)+p_1(s_0)x+p_2(s_0)y+p_3(s_0)x^2 \\ &+p_4(s_0)xy+p_5(s_0)y^2+p_6(s_0)x^2y+\ldots \end{aligned} \tag{6.10}$$

$$
\begin{aligned}
\mathbf{I} &= \iint f(s,s_0)dxdy \\
&= c_0 + c_1x + c_2y + p_0(s_0)xy + p_1(s_0)x^2y + \\
&p_2(s_0)xy^2 + p_3(s_0)x^3y + p_4(s_0)x^2y^2 + p_5(s_0)xy^3 + \ldots
\end{aligned} \tag{6.11}
$$

A group of experimental results were shown to verify the images from high resolution single pixel infrared camera. Figure 6.7 shows the hardware of system, including IR source, object, main lens, DMD and its controller, digital mask. In this configuration, the IR source irradiates on white spartan object, generating pseudo diffuse reflection on the surface. The object irradiation will pass through main lens and form a virtual object image behind DMD. A digital mask and its pattern are followed for angular image capturing.

As shown in Figure 6.8, there are three groups high resolution image. In left, it is 96×96 single angular image, and the right images are $192\times 192(4\times)$, $288\times 288(9\times)$, $384\times 384(16\times)$ pixels distinctly. It is obvious that all high resolution images get the edge of inside black hole, while the original 96×96 image has blur inside and on the edge because the high resolution images has noisy removal during calculation. At the same time, the sharpness becomes better than original LR image and the details inside football becomes clear in which the image goes from “white - gray - black” from left to right. Meanwhile the texture becomes clear.

6.4 Single Images based Super Resolution

Due to the hardware computational limitation, the single pixel camera has relatively low recovery resolution. Although previous interpolation method could enhance image resolution by multiple images, it is complicated with high time cost. The single image based super resolution methods, including interpolation-based, learning-based, statistics-based and others were reported [167]. In

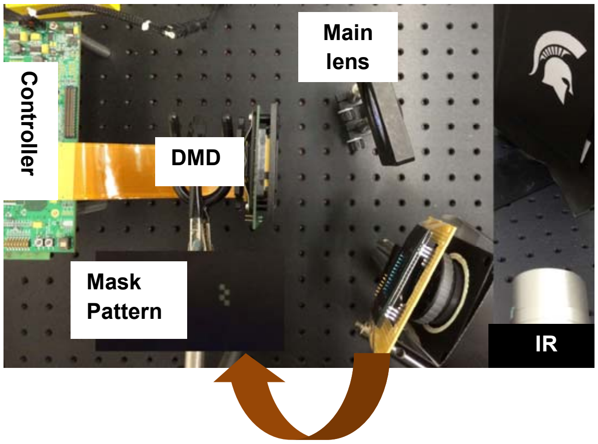


Figure 6.7 Prototype hardware of super resolution single pixel camera.

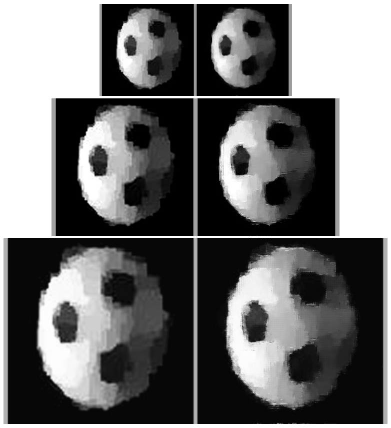


Figure 6.8 Experimental results, from top: 4×, 9×, 16× .

this research, a spline based reproducing kernel hilbert space and approximative heaviside function are deployed to model smooth part and edge component of image respectively. By adjusting the heaviside function parameters, the proposed model will enhance distinct part of images. Compared to multiple scenes or video sequences super resolution method, the proposed algorithm is applicable on real-time camera or video system. The overall system can get super resolution IR image based on single IR photodetector, making IR camera better measurement accuracy and observing more details at long distance.

6.4.1 Observation Modeling

The digital IR image acquiring system is hardware dependent. The output image comes with various kinds of degradations from motion blur, aliasing and noisy version of true scene [143]. As shown in Eq. (6.12), $X = [x_1, x_2, ..., x_N]$ denotes the desired high resolution image above Nyquist sampling frequency, where Y is the single LR image observed, B is the blur models effect and D is the down sampling operator, n represents the noise model. This is simplified model of Eq. (6.1).

$$\begin{aligned} \mathbf{Y} &= DB\mathbf{X} + n, \\ \mathbf{X} &= [x_1, x_2, ..., x_N] \end{aligned} \tag{6.12}$$

As shown in Figure 6.9, the multiple LR images superresolution will reconstruct each HR grid distinctly by extra sampling. However, the single LR image will combine the feature of HR image into one, so that an extra constraint will be necessary for HR reconstruction.

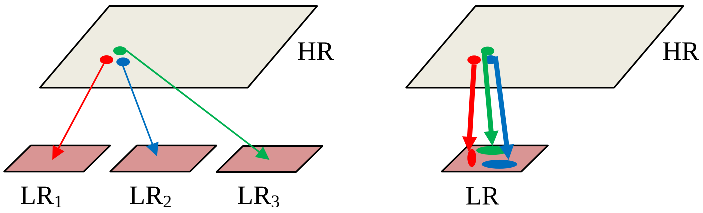


Figure 6.9 Classical multiple images and single image super resolution.

6.4.2 Spline based Reproducing Kernel Hilbert Space and Approximative Heaviside Function Model

The image intensity $f(x,y)$ is divided by smooth component and edges, as $f(x,y) = h(x,y) + g(x,y)$. In the model, the smooth part ($h(x,y)$) is modeled by functions in a reproducing kernel hilbert space (RKHS) and every evaluation function is bounded. Edge is described as step function in $g(x,y)$. For an arbitrary real value $h(x)$, $x \in [0,1]$, its taylor series can be expanded at $x = 0$ as Eq. (6.13). $h_0(x)$ and $h_1(x)$ is corresponding to first and second part of $h(x)$, also referred to lagrange remainder. The span of Eq. (6.14) will form a RKHS [167], $\mathbb{H}_0 = \text{span}\{\phi_1, \phi_2, \phi_3, ..., \phi_m\}$ and $h_0(x) = \sum_{n=1}^{m} d_n \phi_n(x)$, $d_n(n = 1,2,...,m)$ is coefficients.

$$\begin{aligned} h(x) = \sum_{n=0}^{m-1} \frac{x^n}{n!} h^{(n)}(0) + \int_0^1 \frac{(x-t)^{m-1}}{(m-1)!} h^{(m)}(t)dt \\ = h_0(x) + h_1(x) \end{aligned} \tag{6.13}$$

$$\begin{gathered} \phi_n(x) = \frac{x^{n-1}}{(n-1)!} \\ n = 1,2,3,...,m \end{gathered} \tag{6.14}$$

The remainder $h_1(x)$ can also be expanded on a RKHS [167]. Let $G_m(x,u) = \frac{(x-u)^{m-1}}{(m-1)!}$ and the boundary conditions $h_1^n(0) = 0, n = 0,1,...,m-1$. $\mathbb{H}_1$ is a Hilbert space on $[0,1]$ with reproducing kernel $K(s,t) = \int_0^1 G_m(t,u)G_m(s,u)du$ [168]. The function h_1 can be expanded on the basis of $\mathbb{H}_1$, $h_1(x) = \sum_{i=1}^{n} c_i K(s_i, x)$. The generalized taylor series expansion is described as Eq. (6.16). In matrix format, $\vec{h} = Td + \Sigma c$, where T is $n \times m$ matrix , Σ is a $n \times n$ matrix, $d = (d_1, d_2, ..., d_m)'$, $c = (c_1, c_2, ..., c_n)'$. The noise observation $\vec{g} = \vec{h} + \eta$ can be smoothing by minimizing Eq. (6.17).

$$\begin{aligned} h_1(x) &= \int_0^1 \frac{(x-u)^{m-1}}{(m-1)!} h^{(m)}(u)du \\ &= \int_0^1 \frac{(x-u)^{m-1}}{(m-1)!} dh^{(m-1)}(u) \\ &= G_m(x,1)h^{(m-1)}(1) - G_{m-1}h^{(m-2)}(1) + \\ &G_{m-2}(x,1)h^{(m-3)}(1) + .. + G_1(x,1)h^{(0)}(1) \end{aligned} \tag{6.15}$$

$$h(x) = \sum_{v=1}^{m} d_v \phi_v(x) + \sum_{i=1}^{n} c_i \xi_i(x) \tag{6.16}$$

$$\min_{c,d} \frac{1}{n} \parallel \vec{g} - Td - \Sigma c \|^2 + \lambda c' \Sigma c \tag{6.17}$$

As previous discussion, the RKHS based model makes the image smooth. However, the image contrast is edge dependent and functions significantly, especially for narrow band cooled IR camera. The heaviside step function includes a singular point at $x = 0$, which describes the edge on image signal perfectly. In the proposed model, an approximated heaviside function, Eq. (6.18) is applied, where ξ controls the sharpness of signal. The ξ is selected distinctly for reflective IR image, narrow band and broadband IR image depending on the thermal image contrast requirement. The total image intensity function is the summation of edges and smooths part, as $\vec{f} = \vec{h} + \vec{g} = Td + \Sigma c + \Psi\beta$, where $\vec{g}$ is edge model, Ψ is edge matrix and β is edge coefficients.

$$\psi(x) = \frac{1}{2} + \frac{1}{\pi} arctan(\frac{x}{\xi}) \tag{6.18}$$

6.4.3 Iterative Reconstruction Algorithm

Based on proposed model, the super resolution image restoration is the minimization problem as Eq. (6.19), where desired $H = T^h d + \Sigma^h c + \Psi^h \beta$ is the target super resolution output and re-define the equivalent problem in Eq. (6.20). The Eq. (6.20) ℓ_1 problem can be easily converted to its augmented lagrangian function as Eq. (6.21), α, $\rho, b \in \mathbb{R}$. In order to solve Eq. (6.21), it is divided by two subproblems as the u problem and (d, c, β) problem. The $u-$subproblem Eq. (6.22) has the solution of Eq. (6.23) [169].

$$\min \frac{1}{n} \parallel L - DB(T^h d + \Sigma^h c + \Psi^h \beta \|^2 + \lambda c' \Sigma c + \alpha \parallel \beta \|_1 \tag{6.19}$$

$$\begin{gathered} \min \frac{1}{n} \parallel L - (T^l d + \Sigma^l c + \Psi^l \beta \|^2 + \lambda c' \Sigma^l c + \alpha \parallel u \|_1 \\ s.t., u = \beta \end{gathered} \tag{6.20}$$

$$\mathbb{L}(d, c, \beta, u) = \frac{1}{n} \parallel L - (T^l d + \Sigma^l c + \Psi^l \beta \|^2 + \lambda c' \Sigma^l c + \alpha \parallel u \|_1 + \frac{\rho}{2} \parallel u - \beta + b \|^2 \tag{6.21}$$

$$\min \alpha \parallel u \|_1 + \frac{\rho}{2} \parallel u - \beta + b \|^2 \tag{6.22}$$

$$u_i = max\{\parallel \beta_i - b_i \|, \frac{\alpha}{\rho}\} \frac{\beta_i - b_i}{\parallel \beta_i - b_i \|} \tag{6.23}$$

The $(d, c, \beta)-$subproblem is more complicated than $u-$ subproblem. The Eq. (6.24) subproblem can be employed by least square method [167] so that d, c, β can be obtained. The main

Table 6.1 RKHS based single image super resolution algorithm.

```
Algorithm
for i = 1: n
    a. Compute the $u^{(i)}$ by Eq. (6.23).
    b. Update the coefficients: $(d^{(i)}, c^{(i)}, \beta^{(i)})$ =
    argmin $\frac{1}{n} \| L^{(k)} - (T^l d + \Sigma^l c + \Phi^l \beta)\|^2 + \lambda c^{'} \Sigma^l c + \alpha \| \beta \|_1$
    c. Calculate the high-resolution image:
    $H^{(k)} = T^h d^{(k)} + \Sigma^h c^{(k)} + \Phi^h \beta^{(k)}$.
    d. Down sampling to $\tilde{L} = DBH^{(k)}$.
    e. Residual update: $L^{(i+1)} = L^{(k)} - \tilde{L}$
end
```

iteration is shown in Table 6.1.

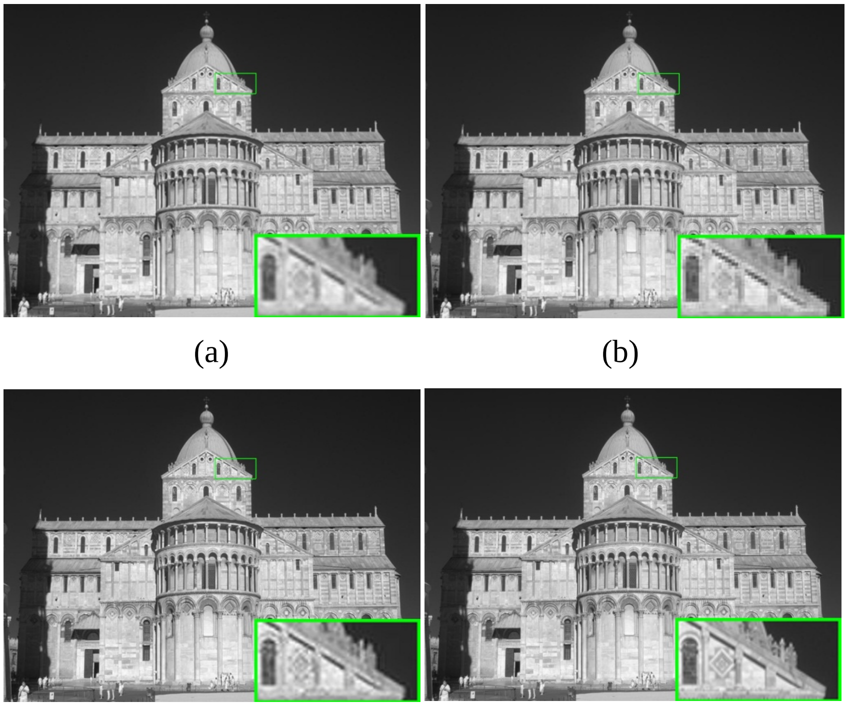
$$\min_{d,c,\beta} \frac{1}{n} \| L - (T^l d + \Sigma^l c + \Psi^l \beta \|^2 + \lambda c^{'} \Sigma^l c + \frac{\rho}{2} \| u - \beta + b \|^2 \tag{6.24}$$

6.4.4 Simulations and Experiments

In this section, there are three groups of numerical experiments discussed: reflective IR images, narrow band cooled and broadband uncooled IR emissive images. The ground truth image is from IR camera. In order for comparison and algorithm evaluation, a low resolution image is generated from downsampling of HR ground truth at scale of 2 or 3. All experiments are run in MATLAB(R2010a) on a laptop of 3.25Gb RAM and Intel(R)Core(TM) i3-2370M CPU: 2.40 GHz.

The reflective IR sensors work similar as SLR camera but different wavelength, which manifests the object shape and material reflectance. The image detail (high frequency components, the edges) and base (low frequency, smooth variations) are equivalent to each other, so that a medium edge part is chosen for modeling near IR image.

As shown in Figure 6.10, the standard bicubic and nearest neighbor approaches are compared



(a) (b)

(c) (d)

Figure 6.10 Near IR image of building. a) bicubic method; b) nearest neighbor; c) proposed method; d) ground truth image.

to proposed method (Figure 6.10 (c)). From qualitative images, it is obviously that the bicubic method smooths the target too much and generates blur at edge, while the jaggy effect happens on nearest neighbor. The small arc door shape and triangular edge in selected green box area are very close to ground true image (Figure 6.10(d)). In quantitative viewpoint, the RMSE value is bicubic: 6.8170, nearest neighbor: 8.7848 and proposed model: 6.0488 at scale of 2. The image spacial resolution increases 4 times with 2.37% normalized RMSE error.

The cooled emissive IR images have ultra high sensitivity and background filter so that it outputs a high contrast with great sharpness images. However, the smooth part will dominate super resolution result since the edge can be sampled in low resolution image. The proposed model is modified with a smaller edge component compared to reflective IR image super resolution. Figure 6.11 and Figure 6.12 show two cooled emissive IR images and three super resolution compar-

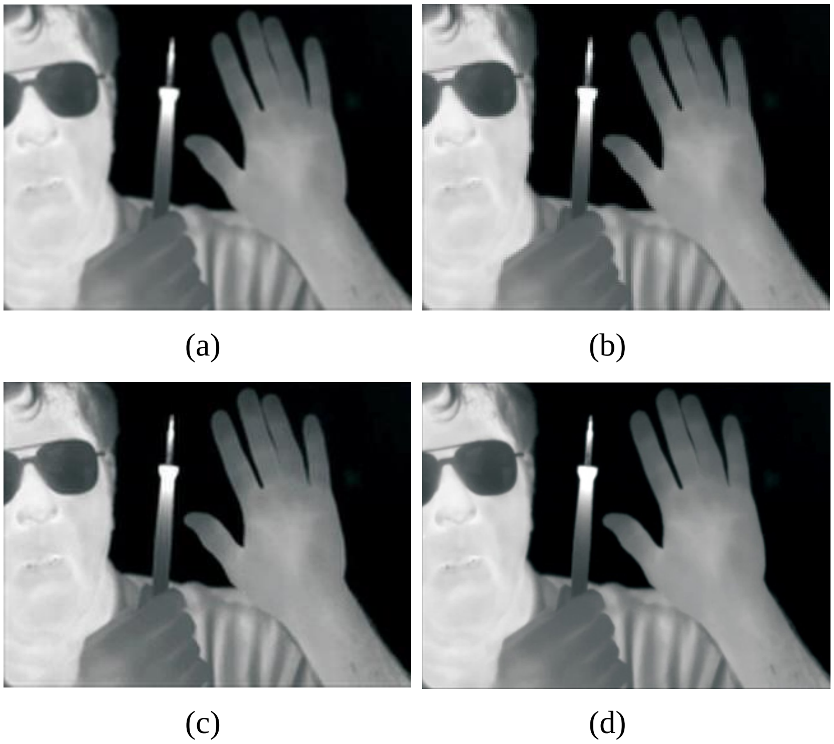

(a) (b) (c) (d)

Figure 6.11 Hand and head infrared image from superlattice (SLS) cooled FPA. a) bicubic method; b) nearest neighbor; c) proposed method; d) ground truth image .

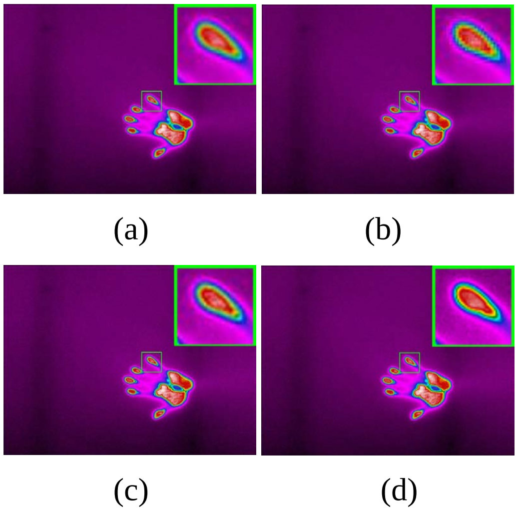

(a) (b) (c) (d)

Figure 6.12 Handprint thermal image from cooled thermal camera. a) bicubic method; b) nearest neighbor; c) proposed method; d) ground truth image.

ison.

The nearest neighbor is worst but all other three images look very similar in Figure 6.11. The

RMSE of bicubic and proposed approach have both relative small values of 2.8092 and 2.3202 respectively. The weaker edge component will improve better output on images which has higher contrast in LR format. Figure 6.12 compares the super resolution reconstruction of handprint thermal image in color style, in which the different color shows the temperature distribution. The proposed recovery is between bicubic method and ground truth image. The selected area color order is purple, blue, green, yellow and red from outside to inside. The downsampling is at scale of 2 and the RMSE is 3.7619 (bicubic) and 3.5696 (proposed).

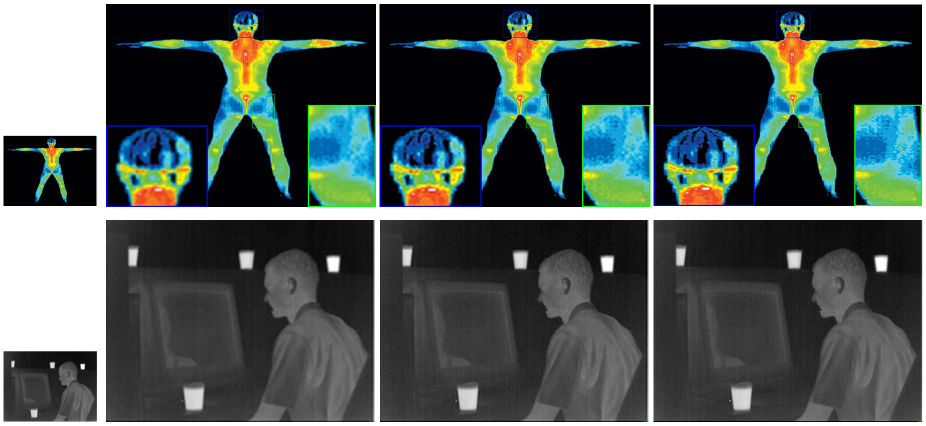
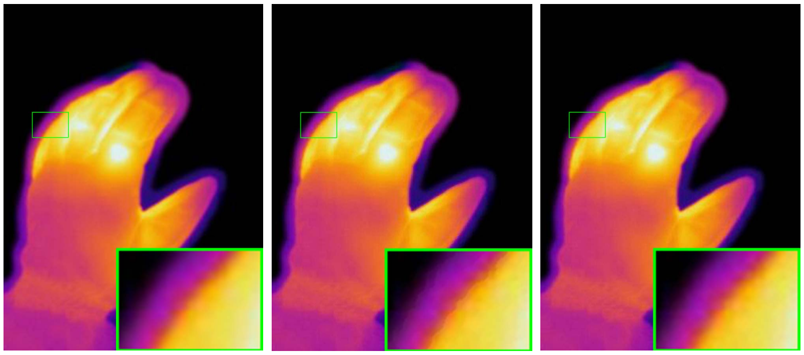


Figure 6.13 The uncooled thermal image super resolution comparison.



(a) (b) (c)

Figure 6.14 Hand infrared image from uncooled thermal IR camera, a) bicubic method; b) proposed method; c) ground truth image .

The uncooled IR images are not as good as cooled sensors. The image thermal contrast is lower than cooled images so that the edge would be lost in low resolution sampling. In the proposed model, a high edge component is applied on this image restoration. Figure 6.13 shows one medical IR image (top) and indoor IR image (bottom) where the left small image is downsampling at scale of 3 from the most right. The large size image in the middle is from proposed model while the left is from bicubic model. From viewing imaging point, the neck part in top image is becoming clear compared to left HR image. Although it is very hard to distinguish from each other due to the large raw image (top: 1200×900, bottom: 384×288), the Table 6.2 shows the RMSE. The worst normalized root-mean-square-error is 2.35% in all results.

Table 6.2 RMSE value of medical IR image and indoor IR image.

Wavelength	Bicubic	Proposed method
Medical IR	6.817	6.004
Indoor IR	5.636	5.479

Most of uncooled thermal IR camera has very low resolution. The low level civil thermal imager features around 60×80 pixels and enhances the spacial resolution to multiple times digitally. The edge information is lost as a result of such large enlargement, as shown in Figure 6.14 (c), the output thermal image is blur. However, a high edge model is introduced to strengthen image details in Figure 6.14 (b). The finger edge can be clearly seen in the reconstruct image.

6.5 Chapter summary

The signal processing based super resolution method is discussed in this chapter which enlarges the image resolution by computing. The multiple LR images and single LR image were both discussed. By taking multiple angular images, a robust high resolution reconstruction algorithm was applied to recover HR image. Meanwhile, the adaptive single image based super resolution model was also discussed for IR image enhancement. The infrared image is split by smooth part and edge details, where the 2D spline based RKHS model is used for image base (smooth component) while a modulated heaviside function is applied for image edge. The super resolution is casted into a model based estimation problem. By computing the coefficients of the redundant basis of low resolution image, it was applied as same in high resolution image. The iterative scheme was proposed to preserve more image details and the experimental results show the model is robust and easily adaptive with different IR image. The computing is still a challenging for this method due to iteration in estimation. However, the parallel computing and hardware accelerating will make the proposed approach fast and easily integrated into consumer electronics.

Chapter 7

Compressive Video Sensing

7.1 Previous Work

In plenoptic function, the light irradiation is time dependent for a dynamic scene so that the video sensing is necessary for moving visual image acquisition. The video camera brings broadcasting lively and real time images directly to a screen for immediate observation. It not only serves for live TV production, but also for security, military and industrial operations. There are numerous designs and applications, including professional video cameras, camcorders, closed circuit television (CCTV), webcams, phone camera and some special camera systems. The analog video production attracted research attention since 1970s, such as videotape recorders (VTR) on magnetic tape, although the digital video recorder was mostly pioneered by company [170].

The digital video is highly correlated with big data because it comprises a series of orthogonal frames at a constant rate. For an example of one minute duration, frame size 640×480 px, 24 bits color depth, and frame rate of 25 fps, the video size will be 1.38 GB. Therefore, the video compression is necessary for recreating the video signals and it is usually integrated in one chip [171] or software algorithm [172]. In conventional method, as shown in Figure 7.1, the compression is after sampling process, of which the most applied standard is MPEG (Moving Picture Experts Group) compression. The separated sampling and compression method will not be applicable to non focal plane sensors imaging system as single pixel camera. Otherwise, it will cost much time and makes it very slow frame rate. The compressive sensing allowes us to combine the signal

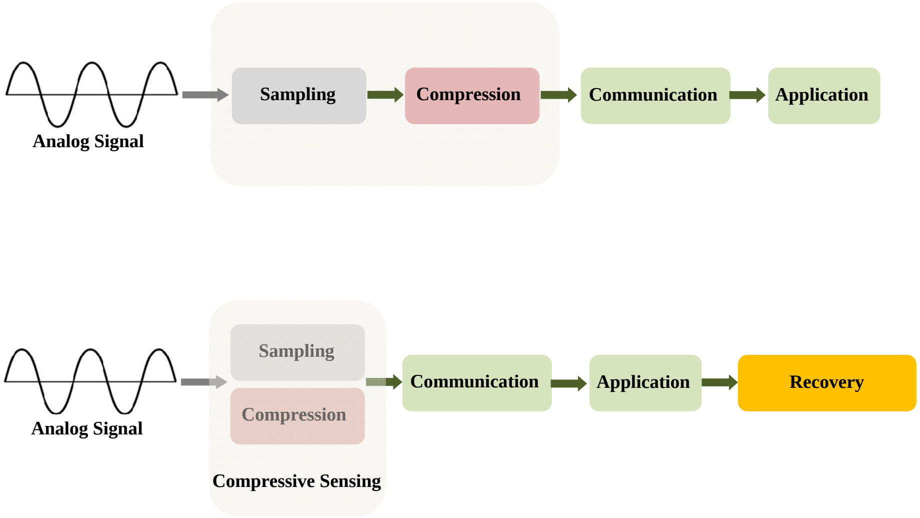


Figure 7.1 Conventional Nyquist Shannon signal sampling and compressive sampling.

acquisition and compression processes into one step [55], as shown in Figure 7.1 bottom.

Although the conventional video camera can record most of scenes, it still doesn't work when seeing more in temporal domain. For a space shuttle discovery flight deck, it is 2.74 gigapixels [173] and a bubble dynamics research needs 500 fps video microscopy [174]. More importantly, the commercialized high performance video camera is very expensive as a basic model 7,500 fps at mega pixel (1K×1K) resolution with 12 bit color depth (fastcam SA5 from photron) is around \$100,000 (quotation from techimaging.com). There are at least two factors, light limitation and readout bandwidth, to elevate the cost. As shown in Figure 7.2 and Eq. (7.1), the electrons accumulated on sensor will be reverse proportional to F-number, the ratio of focal length to aperture size ($F\backslash\#$), but proportional to exposure time (t), incident illumination (I_{src}), scene reflectivity (R), quantum efficiency (q) and pixel size (Δ^2) [175]. The high speed video camera requires less t, less Δ, but they are mutual restraint in the camera design. Another limitation is sensor readout rate. Since the readout timing will include analog to digital conversion, clear charge from parallel register and shutter delay, the 1 mega pixel, 1000 fps, 16 bit color camera requires 4 GB/s readout

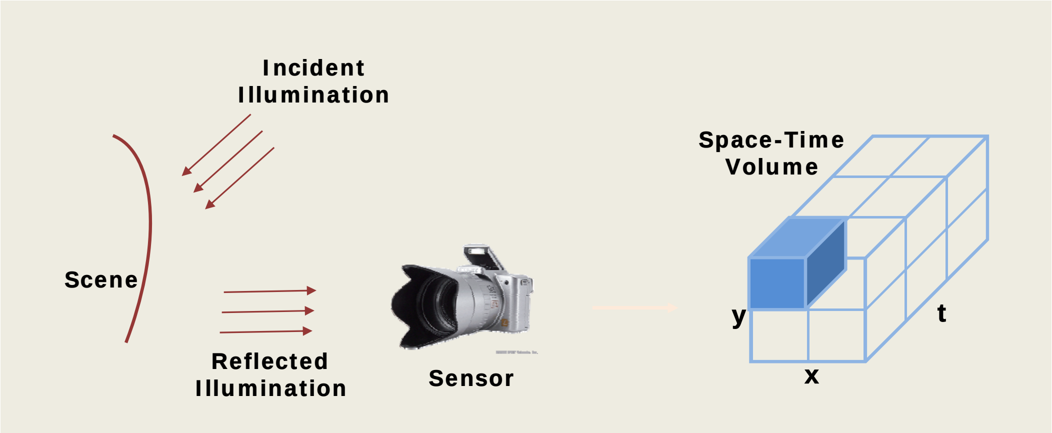


Figure 7.2 Light illumination in SLR camera.

circuitry.

$$\mathbf{J} = 10^{15}\ (F \setminus \#)^{-2}\ tI_{src}R\ q\Delta^2 \tag{7.1}$$

The compressive video sensing was proposed to solve this trade off and made infrared spectrum high speed video camera feasible. In [96], a single pixel camera was set up for video sampling with the assumption of image changing slowly across a group of snapshots. This would not be used for real time video streaming without incurring latency and delay. Another practical system [176] is developed by first splitting each video frame into non-overlapping blocks of equal size. Compressive sampling is then performed on sparse blocks, determined by predicting sparsity based on previous reference frames which are sampled conventionally. However, the remaining blocks are still sampled fully so that the compressive ratio is low. In order to fully use sparsity of image in [177], there are two frames captured the position of the moving object at two instances, by which the difference frame is obtained subtracting. It accumulates residual error and mean square error increases significantly when the difference is not sparse. For a periodic signals, the model based high speed sampling was proposed in [178], to recovery signal with temporal sparsity. Meanwhile, a novel measurement model reduces the bilinear problem to a sequence linear convex problem for

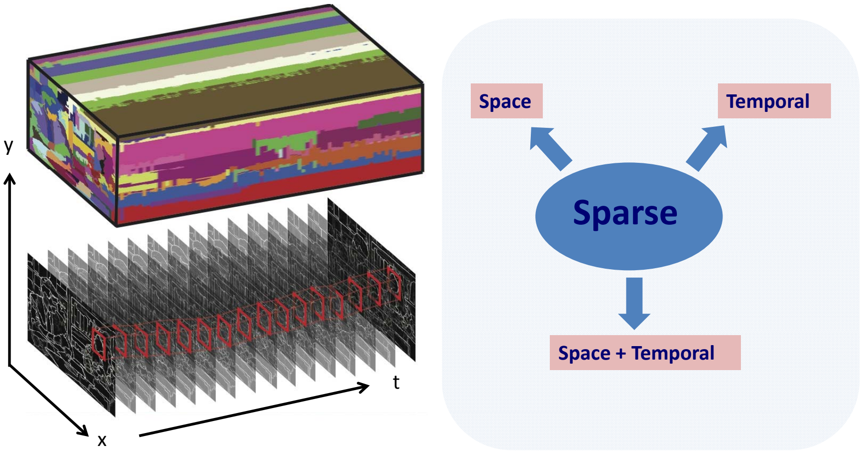


Figure 7.3 Sparsity in video signal.

a linear dynamic system in [179]. However, there is hardly any contributions on infrared spectrum due to the sensor limitation. In this chapter, an infrared high speed video sensing will be discussed.

7.2 Sparsity of Video

The video can be seen as a series of images as shown in Figure 7.3, where the coordinate space (x,y,t) is composed as spatial domain $((x,y))$ and temporal domain (t). For each uncompressed frame, it is full resolution of 2D image. Fortunately, it is proved that the natural image is sparse in specific domain so the spatial information could be redundant and intraframe compression is proposed to discuss the correlation and compression within a frame. The baseline is JPEG compression standard. The another redundant of video is between like frames. The compression, referred to interframe compression, will base on H.261 standard. The audio compression is not included in this thesis.

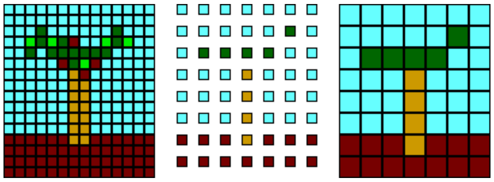


Figure 7.4 Sub-sampling of image compression.

7.2.1 Intraframe Compression

In video sequence, the individual frame can be considered as sperate file. The intraframe coding, by spatial redundancy, refers to the fact that the various lossless and lossy compression techniques are performed relative to information that is contained only within the current frame.

Sub-sampling is the most basic of intraframe compression technique, which reduces the amount of data by throwing some of it away. As shown in Figure 7.4, it samples the color component of image by 2 in both directions, in which every second line and column is completely ignored so that only quarter of pixels are sampled. It is very common to sub-sample the color data and the interpolation process makes sub sampled image appear to original resolution.

Rather than reducing the number of pixels, coarse quantization [180] method accomplishes compression by reducing the number of bits used to describe each pixel. The pixel is reassigned a lower bits color depth than the original image to save storage requirements for images. Vector quantization [181] is also introduced to image compression, where the input data stream is divided as blocks and coded using patterns in a pre-defined table. By comparing the block between input data and pre-defined block in table, the image could be compressive described.

The widely used intraframe compression is transform coding which transforms an image from the spatial domain into other domain (frequency, wavelet etc). The Discrete Cosine Transform (DCT) is the most popular transform in image coding because of the low computational complexity. As shown in Figure 7.5, the 8×8 grid is represented by DCT coefficients instead of 64 original

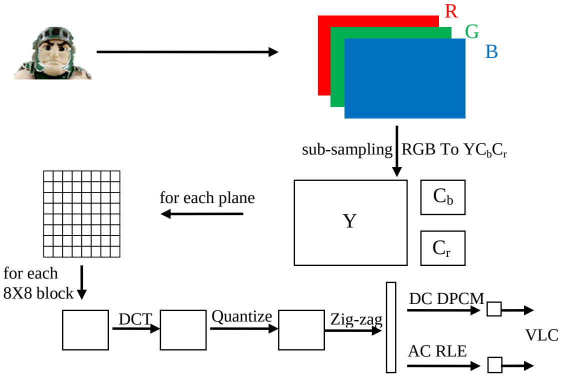


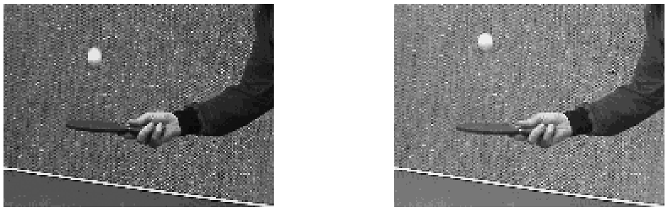
Figure 7.5 DCT based transform coding image compression.

values. The coefficients can be coarsely quantized without seriously affecting the quality of the image which results from an inverse DCT of the quantized coefficients. Most of lossy image coding techniques are extension of these basics but only slight implementation detail differences.

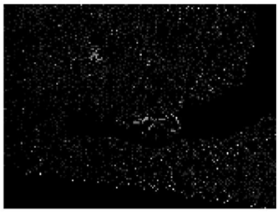
7.2.2 Interframe Compression

In video streams, the adjacent frames are highly correlated and relatively little changes happen from one frame to the next in general. The interframe compression is compression method applied to a sequence of video frames rather than a single image, which exploits the similarities between successive frames, known as temporal redundancy, to reduce the volume of data required to describe the sequence.

In most of interframe compression techniques, the basic concept is to reuse parts of frames already exist. The new frame comes from the previous frame and estimation. The simplest interframe compression is sub-sampling which originates from sub-intraframe sampling. It only transmits some of frames, e.g. every second frame and the decoder (viewer) would be required to interpolate the missing frames at the receiving end. The first interframe compression using adjacent frame is conditional replenishment [182], also named difference coding. It bases on the premise of small



frame 1 frame 2



difference between
frame 1 and frame 2

Figure 7.6 Difference between adjacent frame.

difference between successive frames. As shown in Figure 7.6 (160×120 px) and Figure 7.7, more than 85% pixels have no changes. In this method, each frame of a sequence is compared with its predecessor and only changed pixels have been transmitted. The lossless compression will require to sample every changed pixels. However, only pixels that change significantly will be updated at the cost of introducing some loss.

Since the difference coding compression considers the whole image as target object, the block based difference coding will be more efficiency. During this compression, the frames are divided into non-overlapping blocks and each block will be compared with its counterpart in the previous frame. The update sampling is only for blocks significantly changed. The basic difference coding block size is single pixel while the block based difference coding depends on an area. Due to the larger blocks update some pixels unnecessarily, the block size would be optimized before designing the pattern and algorithm. In general perspective, the block based difference coding is limited by its prerequisites. When there is a lot of motion or the camera itself is moving, there are lots of changes happened in each frame. To solve this problem, it is necessary to compensate the object

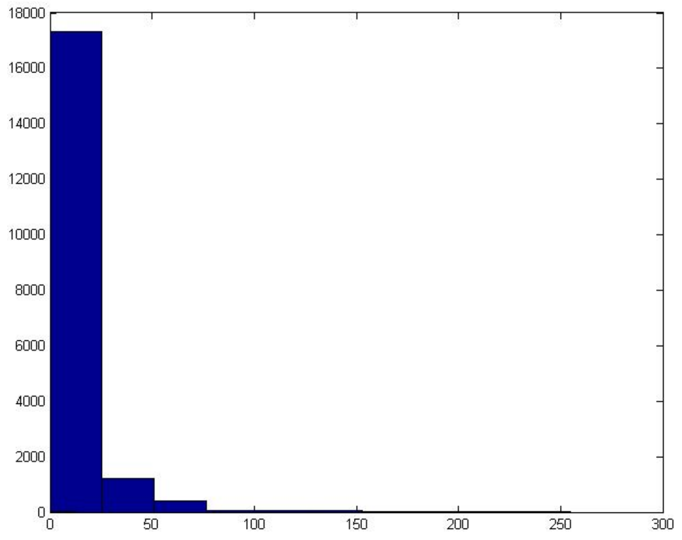


Figure 7.7 Histogram plot of adjacent frame difference.

motion.

The block based motion compensation is a practical and widely used method to compensate movements of blocks for future frame. Firstly, it has frame segmentation, as shown in Figure 7.8. The frame to be approximated (current frame) is divided into uniform non-overlapping rectangular blocks. In [183], it claims that the block is rectangular shape (16×8 pixels) based on the fact that motion within image sequences is more often in the horizontal direction than the vertical. For larger block size, there will be fewer motion vectors but more correction data to transmit [184]. On the opposite, the smaller blocks will result in a greater number of motion vectors, but very sensitive to noise [185]. It is a trade-off between minimizing the number of motion vectors and maximizing the quality of matching blocks in frame segmentation. Secondly, the block matching is the most time consuming of encoding compression because each target block of the current frame will be compared with a past frame so as to find a matching/corresponding block. Although there is a search area in blocking matching, which could be on all of the past frame, it is usually restricted to a smaller area centered around the position of the target block in current frame. If

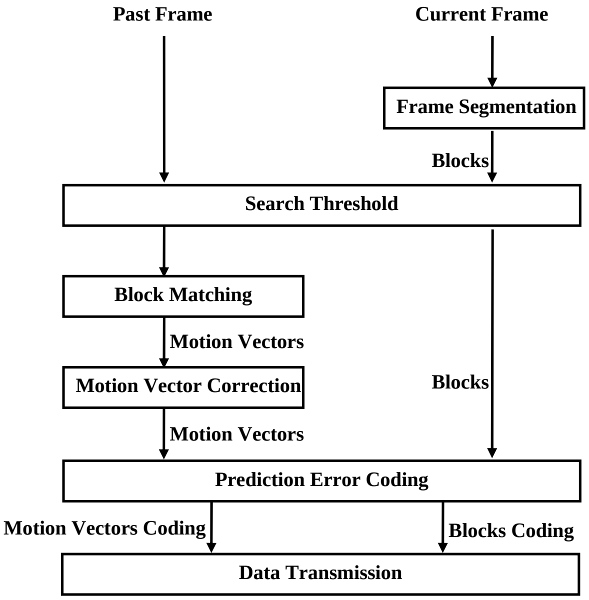


Figure 7.8 Flow chart of motion compensation process.

the block size is b and the maximum displacements in the horizontal, vertical directions are dx, dy respectively, the search area will be of size $(2dx + b)(2dy + b)$. The Exhaustive Search, Brute Force Search, and Full Search are proposed in research for block matching [183]. The matching criterion affects quality of compression, the runtime of algorithm. The whole process of finding the best match is known as motion estimation. Thirdly, a motion vector will be calculated based on relative difference between target block and past frame, which is quantized by efficient arithmetic, adaptive Huffman and Lewpel-Ziv coding [185]. If the target block and matching block are found at the same location in their respective frames, the motion vector that describes their difference is known as a zero vector. In the decompression process, the motion vectors and past frame will be used reconstruct current frame.

However, the block motion compensation has at least two limitations. It is the assumption that the target is moving parallel to camera plane and the illumination is uniform over time.

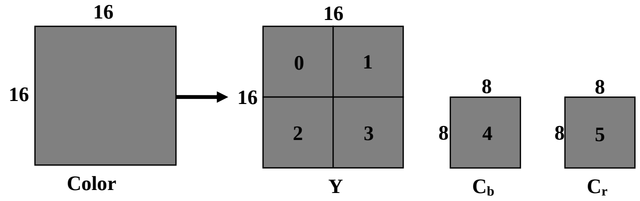


Figure 7.9 Macroblock (4:2:0).

7.2.3 Video Compression

The video compression can be classified as lossless and lossy method. The former compression retains the original data and the algorithm saves some space by just removing image areas that use the same color. The lossy compression methods remove image and sound information that is unlikely to be noticed by the viewer or not differentiated by the human perception. Therefore, the quality perceived is still the same, but the volume is dramatically decreased.

The most popular compression scheme is MPEG-X (Moving Picture Experts Group) series and H.26x series. The common of these visual coding standards are similar macroblock based motion compensated prediction, as shown in Figure 7.9. It is corresponding to a 16×16 pixel region of a frame, including 16×16 pixels for Y frame, 8×8 for C_b and C_r frames. The 4:2:0 chroma subsampling will have four Y, one C_b and one C_r.

In H.261 standards, there are two types of images frames: Intra-frames (I-frames) and Inter-frames (P-frames), as shwon in Figure 7.10. The I-frames are coded using intraframe compression, while the P-frames are coded by a forward predictive coding method. In Motion compensation, the encoder will search the image surrounding marching block to discover where it comes from. The Brute Force, Hierarchical Search are applied to encode motions. The motion vector depending on the difference between current macroblock and best match block will be encoded. The DCT coefficients will be quantized for entropy coding.

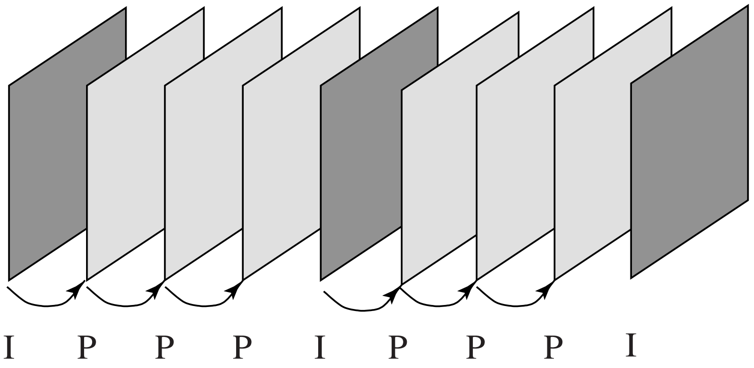


Figure 7.10 H.261 frame sequence.

MPEG-1 is another international standards employing two basic techniques: block based motion compensation for temporal redundancy and transform coding, such as DCT for spatial redundancy. The video sequence is divided into a group of pictures, as Intracoded Frame (I-frame), Predictive Frame (P-frame) and Bidirectional predictive Frame (B-frame), as shown in Figure 7.11. For forward prediction, the macroblock in P-frame is assigned a best matching block from previous I-frame or P-frame. However, the target macroblock may not have a good matching in the previous frame due to unexpected movements and occlusion, but it might be best match from macroblock in next frame. In MPEG-1 search matching algorithm, a backward prediction is also performed in addition to forward prediction so that the best matching macroblock is from a next P-frame or B-frame. Consequently, each macroblock from a B-frame will be specified to two motion vectors, one from forward and one from backward prediction. Figure 7.12 shows the whole MPGE-1 encoding process.

The MPEG-X series achieves a high compression rate by storing only the changes from one frame to another. MPEG-1 provides a 352×240 resolution at 30 fps. The MPEG-2 provides a 720×480 and 1280×720 resolutions at 60 fps with full CD-quality audio. Although there are many standards from MPEG-1, they all have similar motion compression with little difference on coding or motion estimation.

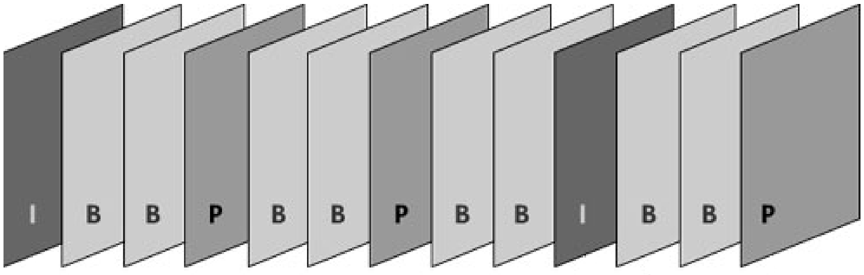


Figure 7.11 MPEG-1 frame sequence.

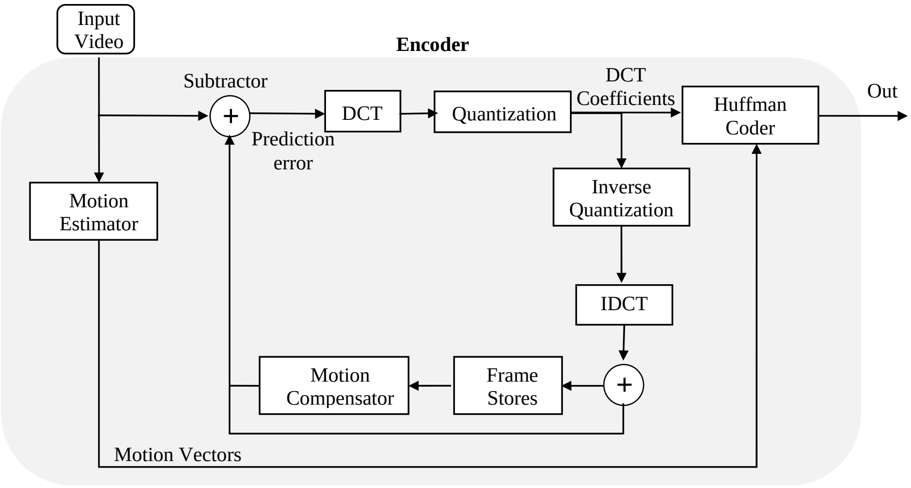


Figure 7.12 MPEG compression.

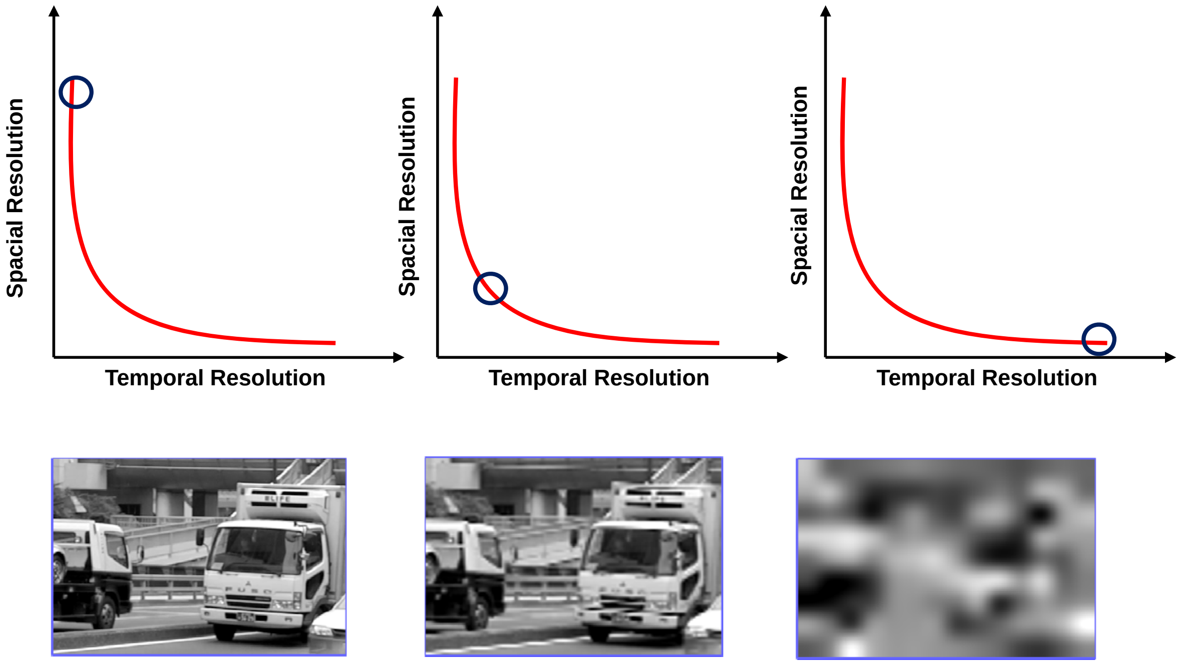


Figure 7.13 Spatial and temporal resolution trade-off in video stream.

7.3 Compressive Video Sensing

7.3.1 Introduction

As discussed in this chapter, the temporal resolution and spacial resolution are a trade-off due to Eq. (7.1) in conventional video sampling. As shown in Figure 7.13 left, the image is high spacial resolution but low temporal resolution and the spacial resolution will go worse when temporal resolution increase, shown in Figure 7.13 right. However, there is no trade off between spacial and temporal resolution in compressive video sensing. The spacial compressive sensing could be implemented using spacial light modulator, DMD (Digital Micro Device), LCOS-SLM (Liquid Crystal on Silicon-Spatial Light Modulator), discussed in Chapter 3 and Chapter 4. The combined spatial and temporal compressive sensing will be discussed in this chapter.

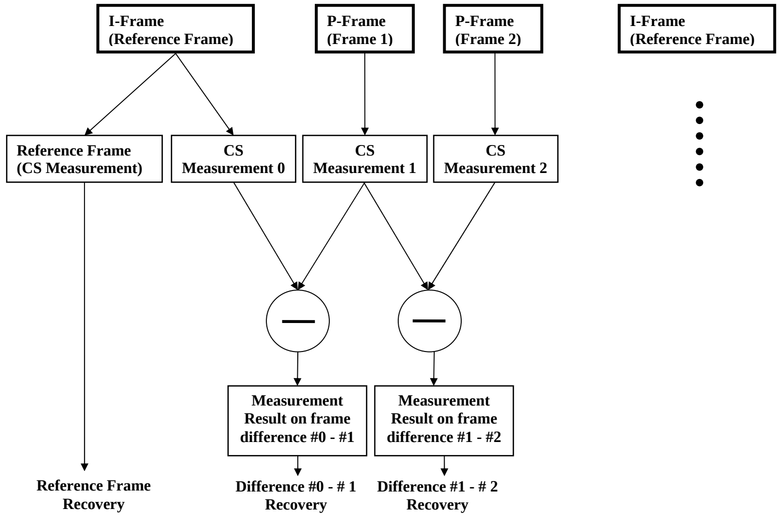


Figure 7.14 Frame difference sampling.

7.3.2 Combined Sparsity Sampling for Video

In the proposed design, the temporal redundancy will be the first priority. As H.263 standards, the image sequence is divided by groups as three frames in each. The first frame is reference, which only samples with spacial redundancy, as shown in Figure 7.14. Another two frames following reference frame, name P-frame, will be sampled by temporal redundancy and reconstructed by reference frame and the difference recovery.

7.3.3 Non-convex Problem Solver

Although ℓ_1 minimization is fully understood and stable with a number of theoretical results, it is not the best sparest solutions in special applications, e.g. computed tomography, video stream difference in which the recovered signal is very sparse, as shown in Figure 7.6. The non-convex optimization, also referred to ℓ_p minimization, requires weaker conditions and enhances the sparsity of recovery solutions.

The basic non-convex optimization is $\ell_p(0<p<1)$, between ℓ_1 and ℓ_0 problem [186] [187].

Figure 7.15 shows the feasible set of $\boldsymbol{y} = \Phi\boldsymbol{x}$ at $0, 1/2, 1$ and 2 norm. The norm-2 has a bad intersect point which is not on x_1 or x_2. Relatively speaking, the $\ell_{1/2}$ is closer to non-convex problem ℓ_0 and the smaller p will recover more sparse result. More importantly, the ℓ_p minimization problem reconstructs results from fewer linear measurements than ℓ_1 optimization [188]. In video streams, the difference between the adjacent frames are very sparse, which need nonconvex penalty to improve the recovery performance.

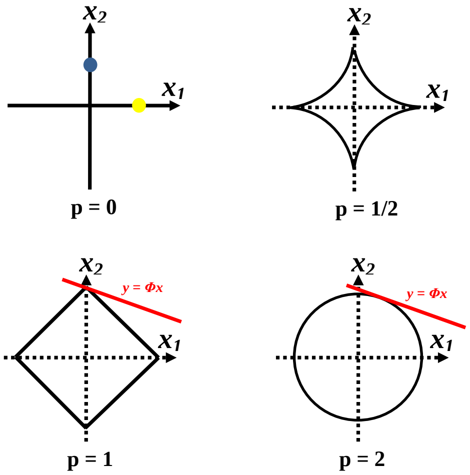


Figure 7.15 Countour maps of different penalties and feasible set of $y = \Phi x$ at $p = 0, 1/2, 1$, and 2.

In order for understanding, the bold font $\boldsymbol{x}$ denotes a vector and $\boldsymbol{x} = (x_1, x_2, ..., x_n) \in \mathbb{R}^n$. The truth value is $\boldsymbol{x}_0$ and the variable at its i_{th} iteration is denoted as $\boldsymbol{x}^{(i)}$. $\tilde{\boldsymbol{x}}$ represents the converged value. In general, there are three categories of ℓ_p problem algorithms for Eq. (7.2). Iteratively Reweighted ℓ_1 minimization (IRL1) is the first method to solve problem Eq. (7.2). The weighted ℓ_1 minimization is demonstrated as Eq. (7.3). Candes et al. [189] suggests the weights as inversely proportional to the the ground truth value. In implementation, it is suggested to choose according to the current iterates, as Eq. (7.4). The iteratively reweighted minimization problem can be simplified as Eq. (7.5), by iteratively solving it.

There are a lot of extending algorithms by IRL1 and the weights are suggested as Eq. (7.6) [187] [190], where $\varepsilon > 0$ to avoid zeros under denominator. The IRL1 is analyzed and proved that the iterates will converge to the sparsest solution when measurements are sufficient.

$$\begin{aligned} \tilde{\boldsymbol{x}} &= \operatorname{argmin} \| \boldsymbol{x} \|_p \\ \text{s.t.} \quad & \Phi \boldsymbol{x} = \boldsymbol{y} \end{aligned} \tag{7.2}$$

$$\begin{aligned} &\min_{\boldsymbol{x} \in \mathbb{R}^n} \sum_{i=1}^{n} w_i \parallel x_i \parallel \\ &\text{s.t.} \quad \boldsymbol{y} = \Phi \boldsymbol{x} \end{aligned} \tag{7.3}$$

$$w_i^{l+1} = \frac{1}{| x_i^l | + \varepsilon} \tag{7.4}$$

$$\boldsymbol{x}^{l+1} = \operatorname{argmin}\{\sum_{i=1}^{n} w_i^l |x_i|, \quad \text{s.t.} \quad \Phi \boldsymbol{x} = \boldsymbol{y}\} \tag{7.5}$$

$$w_i^{l+1} = \frac{1}{(| x_i^l | + \varepsilon)^{1-p}} \tag{7.6}$$

The parallel approach, Iteratively Reweighted Least Squares (IRLS) was proposed in [191] for ℓ_p minimization. It is very similar as IRL1 but completely different, as Eq. (7.7), where the weights are set by Eq. (7.8). Although it is shown that IRLS is theoretically better than IRL1, the convergence is still uncertain.

$$\boldsymbol{x}^{l+1} = \operatorname{argmin}\{\sum_{i=1}^{n} w_i^l x_i^2, \quad \text{s.t.} \quad \Phi \boldsymbol{x} = \boldsymbol{y}\} \tag{7.7}$$

$$w_i^{l+1} = ((x_i^{(l)})^2 + \varepsilon)^{\frac{p}{2}-1} \tag{7.8}$$

The iteratively thresholding method [188] has been established for unconstrained problem by introducing $F(\boldsymbol{x}, \lambda)$ as penalty function. Eq. (7.9) can be solved by alternating minimization of Eq. (7.10), where $\boldsymbol{z}^l$ is an auxiliary variable. By simply assigning $\boldsymbol{z}^{(l+1)} = \boldsymbol{x}^{(l+1)}$, the solution will be given by Eq. (7.11), where Θ is referred to thresholding function. This algorithm will search a local minimization sequence for large scale problem, although it only applies to ℓ_p problem at $p =$ 0, 1/2, 2/3 and 1.

$$\min_{\boldsymbol{x}\in\mathbb{R}^n} \quad \frac{1}{2} \parallel \boldsymbol{y} - \Phi\boldsymbol{x}\|^2 + F(\boldsymbol{x}, \lambda) \tag{7.9}$$

$$\begin{aligned} \boldsymbol{x}^{(l+1)} = \operatorname{argmin}_{\boldsymbol{x}\in\mathbb{R}^n} \frac{1}{2} \parallel \boldsymbol{x} - [(I - \Phi^T\Phi)\boldsymbol{z}^{(l)} + \Phi^T\boldsymbol{y}]\|^2 \\ + F(\boldsymbol{x}, \lambda) \end{aligned} \tag{7.10}$$

$$\boldsymbol{x}^{(l+1)} = \Theta((I - \Phi^T\Phi)\boldsymbol{z}^l + \Phi^T\boldsymbol{y}; \lambda) \tag{7.11}$$

7.3.4 Non-convex Sorted ℓ_1 Method

In order to get closer to ℓ_0, a non-convex sorted ℓ_1 is introduced. Let the coefficient ω_i be a nondecreasing sequence of nonnegative numbers, where $\omega_n \geqslant ... \geqslant \omega_2 \geqslant \omega_1 \geqslant 0$. The objective function is defined as Eq. (7.12), in which the higher weights are assigned on components with smaller absolute values. The contour map of proposed ℓ_p is shown in Figure 7.16. An additional variable P is introduced in Eq. (7.13) to solve the nonconvex sorted ℓ_1 problem, where $\Omega = (\omega_1, \omega_2, ..., \omega_n)^T$, $(\boldsymbol{x})_i$ is the i^{th} element of vector $\boldsymbol{x}$. As assumed the absolute values are ranking in decreasing order for generality, $|x_1| \geqslant |x_2| \geqslant ... \geqslant |x_n|$.

$$R_{\omega}(x_1, x_2, ..., x_n) = \omega_1|x_1| + \omega_2|x_2| + ... + \omega_n|x_n| \tag{7.12}$$

$$F(\boldsymbol{x}, P) = \sum_{i=1}^{n} (P\Omega)_i |x_i| \tag{7.13}$$

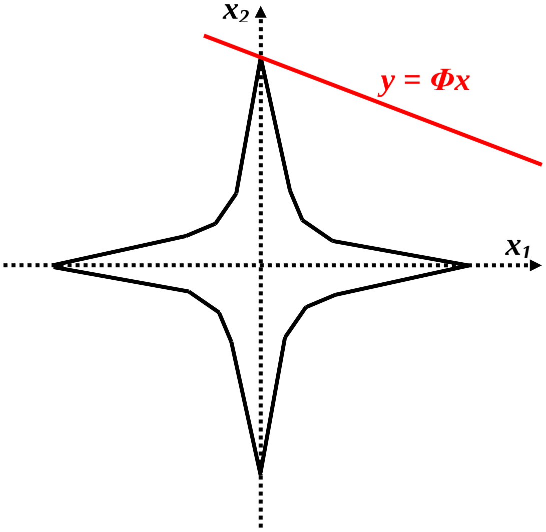


Figure 7.16 Countour maps of proposed nonconvex sorted ℓ_1 with M (M = 4) values.

Given any $P \subseteq R^{n\times n}$, if $(P\Omega)_1 \neq \omega_1$, let $(P\Omega)_k = \lambda_1$, and $k > 1$. By row switching of P, let 1^{st} row exchange with k^{th} row and obtain $P^{(1)}$ such that $(P^{(1)}\Omega)_1 = \omega_1$, otherwise $P^{(1)} = P$. It is easy to derive the relation of Eq. (7.14), and $F(\boldsymbol{x}, P^{(1)}) \leqslant F(\boldsymbol{x}, P)$. For an arbitrarily $j > 1$ and $i = 1, 2, ..., j$, it is similarly to find $P^{(j)}$ such that $(P^{(j)})_i = \omega_i$, and $F(\boldsymbol{x}, P^{(j)}) \leqslant F(\boldsymbol{x}, P^{(j-1)}) \leqslant ... \leqslant F(\boldsymbol{x}, P)$. After n times ordering, $P^{(n)}\Omega = I_n\Omega$, where I_n is identity matrix, and $F(x, I_n) \leqslant F(x, P)$. It means $\min_P \quad F(\boldsymbol{x}, P) = F(\boldsymbol{x}, I_n) = R_{\omega}(\boldsymbol{x})$, shown in Eq. (7.12).

$$\begin{aligned} F(\boldsymbol{x}, P^{(1)}) - F(\boldsymbol{x}, P) &= \omega_1|x_1| + (P\Omega)_1|x_k| - (P\Omega)_1|x_1| - \omega_1|x_k| \\ &= [\omega_1 - (P\Omega)_1](|x_1| - |x_k|) \leqslant 0 \end{aligned} \tag{7.14}$$

Since the nonconvex sorted ℓ_l has the same convergence as minimization of $F(\boldsymbol{x},P)$, the equivalent basis pursuit problem Eq. (7.3) will become Eq. (7.15), where $L(\boldsymbol{x})$ is lagrange multiplier of unconstrained function. It is proven that by a given fixed $\boldsymbol{x}^*$, when $\boldsymbol{x}^*$ minimizes $F(\boldsymbol{x},P)$, the $\boldsymbol{x}^*$ is a local minimizer of $F(\boldsymbol{x})$ in [188]. In Eq. (7.15), there are two variables to solve the problem. The alternating minimization procedure will be the best option because it is easy to solve problem with any one of variables fixed. In this approach, the optimization is divided by variable P and $\boldsymbol{x}$. The problem for $\boldsymbol{x}$ with fixed P is formulated into a weighted ℓ_1 minimization, such as basis pursuit. The updating of P is to change the weights, referred to iteratively reweighted ℓ_1 problem.

$$\begin{gathered}\min_{\boldsymbol{x}} \sum_{i=1}^{n} \omega_i |x_i|, \quad \text{s.t.} \quad \boldsymbol{y} = \Phi \boldsymbol{x} \\ or \\ \min_{\boldsymbol{x}} R_\lambda(\boldsymbol{x}) + L(\boldsymbol{x})\end{gathered} \tag{7.15}$$

The nonconvex sorted ℓ_1 can be divided as different levels based on the number of weights. It becomes 2-Level when giving two different weights, e.g. $\omega_1 = \omega_2 = ... = \omega_k = a_1$ and $\omega_{k+1} = ... = \omega_n = 1$. It will turn into iterative support detection when $a_1 = 0$. The proposed M-Level sorted ℓ_1 minimization has $m(m > 1)$ number of weights and the value is generated by Eq. (7.16), where r controls the decreasing rate form 1 to 0, K depends on support detection discussed later. The parameters a_1 and r are closely related to signal sparsity and the algorithm nonconvexity. The smaller a_1 is, the more nonconvexity becomes. In order to converge fast, it starts from ℓ_1 to get initial value x_0 in the beginning and increases the nonconvexity by updating K.

$$\omega_i = \begin{cases} 1 & \text{if } i > K \\ e^{-r(K-i)/K} & \text{otherwise} \end{cases} \tag{7.16}$$

The weights generating rule is critical on converging speed. In the proposed method, K in Eq. (7.16) is updated by reliable support detection and there are two categories signals included, sparse Gaussian signals and certain power law decaying signals. In each iteration, a threshold (ε) is generated to compare with x_i and then K is determined. The threshold (ε) is redifined in outer loop by counting the signal fast decaying.

The most simple threshold generating rule is $\varepsilon^{(l)} = \mid x^{(l)} \mid_\infty / \beta^{(l+1)}$, $\beta > 0$ [192]. This is an effective rule when selecting an appropriate β because the large β introduces a small number of iteration but low solution quality while the small β causes a large number of iterations. In this work, the rule is based on location of "*first significant jump*" in the increasingly ordered sequence $\mid x^{(l)} \mid$. Finding the smallest i such that $\mid x^{(l)}_{[i+1]} - x^{(l)}_{[i]} \mid > \varepsilon^{(l)}$, then set $\varepsilon^{(s)} = \mid x^{s}_{[i]} \mid$, where $x_{[i]}$ represents the i^{th} largest component in $x^{(l)}$ by magnitude. In sparse signal, the true nonzeros are large value but small in number, while the false signal are large in number but small value as noise. By this method, the true ones are spread out while the false elements are clustered. This is proved and experimentally verified in [192].

Besides, the accumulated residual errors within interframes is another challenge in proposed compressive video sampling. Since the noise error only happens on the significant changes part, especially on the image edge due to frame difference sensing mechanism, the error of frame difference reconstruction can be converted into edge detection problem. The edges has strong intensity contrast and it is a jump intensity from one pixel to the next, so that the edge detection will remove background noise even the residual errors.

The majority of distinct edge detection could be grouped into two categories, gradient and laplacian. The gradient method detects the edges by looking for the maximum and minimum in the first derivative of the image, while the laplacian method searches for zero crossings in the second derivative of the image to find edges [193]. The gradient edge detection shows a maximum located

Table 7.1 Iteratively reweighted ℓ_1 minimization with thresholding.

Algorithm

```
Initialize $\boldsymbol{x}_0, \omega$
for i = 1: maxit
    a. Compute threshold $\varepsilon$.
    b. Update $\omega^{(i)}$ by support detection, and check stopping rules.
    c. Update $\boldsymbol{x}^{(l+1)}$ with fixed weights. Yall1 solver [195] for $\ell_1$-minimization models
end
d. Sobel edge detection of $\tilde{\boldsymbol{x}}$ and select significant changes part
e. Reconstruct next frame image by current frame and the denoised frame difference.
```

at the center of the edge in the original signal [194]. In this work, Sobel operator, using a pair of 3×3 convolution mask on *x*-*y*, is used to detect frame difference, which finds the approximate absolute gradient magnitude at each point in grayscale image and enhances frame difference to improve image quality.

7.3.5 Numerical Analysis

In this section, a group of numerical experiments are analyzed and compared to illustrate the performance of proposed algorithm. In order for comparison, iterative reweighted ℓ_1 (IRL1), two different weights method including ISD (ISD), 2-Level (2-Level) and multiple weights ℓ_1 minimization (M-Level) are run for same object. In IRL1, the weights are updated by Eq. (7.17). The difference between ISD and 2-Level is the weight value (0, 1) pair or (a, 1) $a \in (0,1)$ pair. In this analysis, a is selected as 0.6.

$$\omega_i^{(l)} = \frac{1}{|x_i| + max\{0.5^{l-1}, 0.8^8\}} \tag{7.17}$$

The nonconvex compressive sensing has advantage on sparse signal recovery compared to convex problem. In compressive video recovery, the frame difference is quite sparsity such that the nonconvex algorithm has better recovery. Firstly, it examines the relation between outcome of reconstruction and signal sparsity. As shown in Figure 7.17, the target is to reconstruct a sparse signal (4096 in length) which has k nonzero value ($k \in [100, 1000]$). The index of k is pseudorandom values selected within 4096 and signal intensity is also pseudorandom, as $x = (0, 0, ..., -1.5293, 0, 0..., -0.9123, 0, ...)$. In order for sampling hardware compatible, the measurement matrix is Bernoulli matrices with ±1 entries and the dimension of matrix is 614×4096 (15%).

There are two criteria to evaluate recovery algorithm, runtime and root-mean-square error (RMSE). As shown in Figure 7.17, under same measurement matrix, when setting 500 nonzero values in observe signal, the recovery by IRL1 cost most time around 28 s and RMSE is close to 0.32, largest error among four algorithms. In general, IRL1 algorithm has the worst results with longest runtime and largest RMSE, while the proposed 2-Level and M-Level have smaller RMSE and less time. In the compressive video sensing, the less RMSE will fulfill the goal to reduce the number of samplings and increase video speed.

As discussed in previous section, nonconvex approach will reduce the number of measurements to achieve fully recovery. The less measurements are required, more powerful algorithm is deployed. The second analysis is to explore the outcome of reconstruction and sampling rate. As shown in Figure 7.18, there are two adjacent frames (particle moving) and its difference by simply subtraction. In this simulation, the number of measurements is adjusted from 6% to 35% (×4096), as shown in Figure 7.19. Each row of measurement matrix will keep same on row index and the newly added line is generated by pseudorandom bernoulli distribution with ±1 entries. In Figure 7.19, the x-axis represents the percentage between number of measurements and signal length (4096). From the analysis results, the IRL1 still has worst output with longest runtime and largest RMSE. The M-Level algorithm has least RMSE when measurement percentage is less than

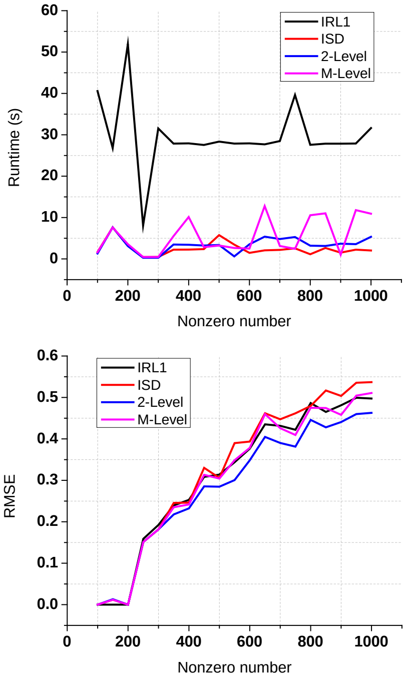


Figure 7.17 Signal recovery on distinct sparsity, 4096 in length.

0.1 (10%). Meanwhile, the Peak Signal to Noise Ratio (PSNR) which describes the quality of recovery image and its ground truth is also discussed. The M-Level proposed algorithm has largest PSNR on above 8% percentage measurement.

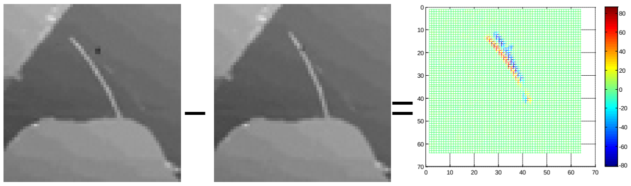


Figure 7.18 Adjacent frame intensity difference.

The third analysis discusses the interframe error. Accumulation residual error deteriorates image reconstruction in far behind, e.g. there are 10 frames in one group, 1 reference frame (I-frame) and 9 interframes (P-frames, named $P_1, P_2, \ldots, P_9$), P_9 frame will have largest error since all errors happening before will be accumulated on this frame. As shown in Figure 7.20, there are five frames (particle moving) recovery based on one I-frame. P_{01} - P_{05} shows five ground truth

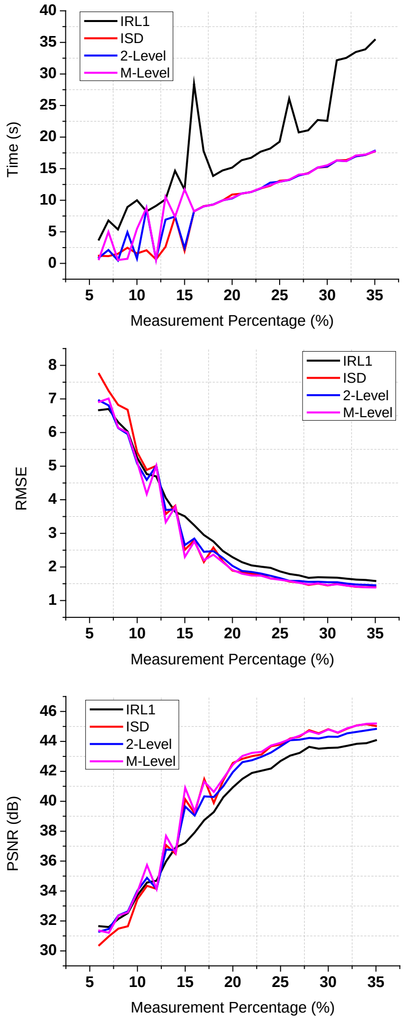


Figure 7.19 Signal recovery on different sampling rate.

frames. P_{11} - P_{15} are reconstruct images directly using M-Level sorted algorithm, while P_{21} - P_{25} come after edge detection denoising. Since the image dimension is only 64×64, it is hardly seeing difference in visual. The RMSE and PSNR are listed in Table 7.2 for comparison. The residual errors increase when recovering more interframes. However, the edge detection denoising can improve frame reconstruction by 1.5 dB in PSNR because most error happens on the frame difference in proposed method. The edge detection removes noise by selecting the most significant changing part and improves the interframe image quality.

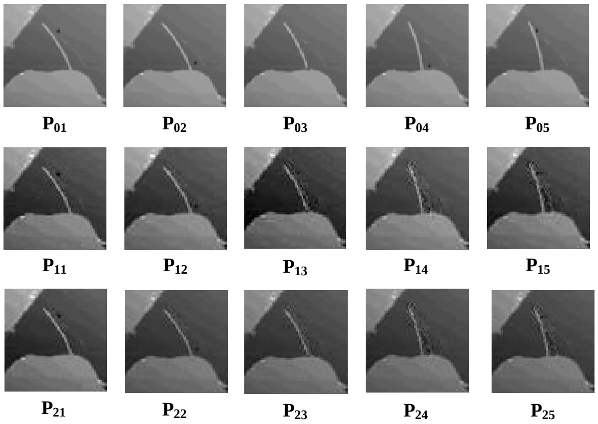


Figure 7.20 Accumulation residual error.

Table 7.2 Characterizing video frames accumulation residual error by PSNR and RMSE.

	P_{01}	P_{02}	P_{03}	P_{04}	P_{05}
$PSNR_{M-Level}$	40.8987	37.4587	36.2745	35.6323	35.0012
$PSNR_{Denoising}$	42.3382	37.5839	36.6928	35.9371	35.0856
$RMSE_{M-Level}$	2.2994	3.4167	3.9157	4.2163	4.5340
$RMSE_{Denoising}$	1.9482	3.3678	3.7317	4.0709	4.4901

7.3.6 Experimental Implementation and Results

The compressive video sampling is implemented on a spatial light modulator, digital micromirror device from Texas Instruments. The target dynamic scene is projected onto DMD plane by objective lens. After projection with measurement matrix (on DMD), another lens focuses the projected image to a single sensor. In the experiments, the irradiator (THORLABS LIU850A) is 850 nm near IR source and commercial silicon photodiode (THORLABS FDS1010) is selected as receiver sensor. The DMD sampling rate (projection) is 6000 Hz.

There are two experiments including linear moving object and rotating object to validate the proposed compressive video system. Figure 7.21 shows one airplane moving with the frame rate 10 fps. In this experiments, there is only one I-frame and one P-frame, e.g. $t_{01}, t_{03}, ...t_{17}$ are P-frames and $t_{00}, t_{02}, ...t_{16}$ are I-frames. The sampling ratio is 18% and 8.5% for reference frame and P-frame respectively. The proposed video system clearly records whole scene on realtime.

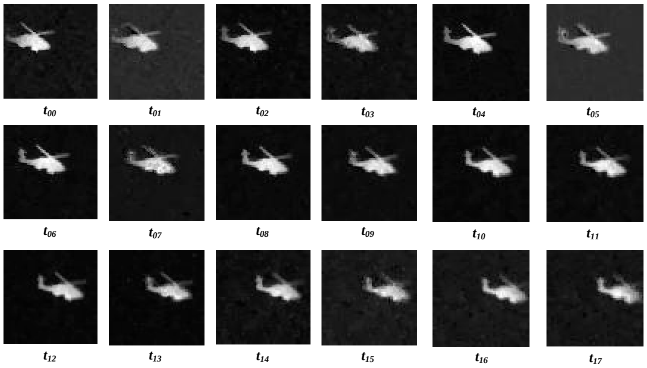


Figure 7.21 Moving object video recording.

The second experiment is to capture the fan rotation. As shown in Figure 7.22, each blade is designed with different length for easily identification. There is one reference frame (I-frame) and three P-frames in testing. Each line shows one group of frames, including one I-frame and three P-frames. The video frame is demonstrated from first to seventh line on time sequence. The frame rate is 18 fps and the sampling ratio of reference frame is 20%, 9% for P-frame.

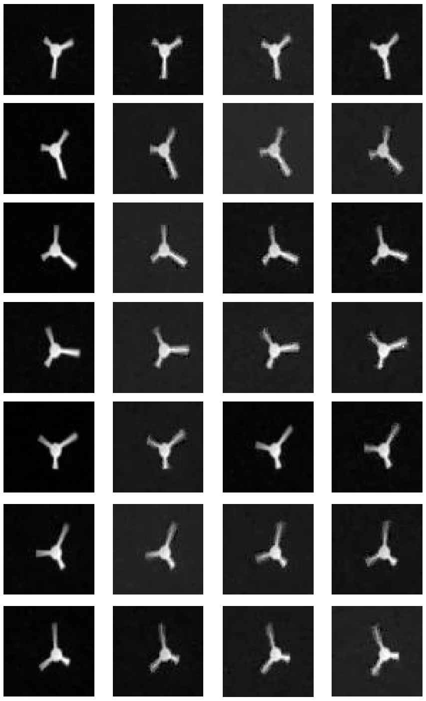


Figure 7.22 Rotating object video recording.

7.4 Chapter summary

In this chapter, the video compression was discussed by spacial and temporal redundancy. The combined method uses H.263 standard to divide video frame as intra-frame (reference frame) and inter-frame. The non-convex compressive video sensing has advantage on requiring very few number of measurements to record a real time dynamic scene. The proposed non-convex sorted ℓ_1 approach has fast convergence to local minimizer and achieves a high accuracy. Besides, the edge detection based denoising reduces the accumulated residual error on frame difference to increase frame rate. Compared with the conventional compressive video sensing, there is no image analysis during sampling process. Although the experimental video frame is only 18 fps, the sampling frame rate can reach to 105 fps based on current DMD mirror limitation (maximum 32,000Hz). Moreover, this non-convex compressive video sensing gives a real-time video and makes single pixel compressive video camera available.

Chapter 8

Conclusions and Future Work

8.1 Conclusions

Infrared spectrum expands visual to explore more stories beyond visible imaging. In the state of art infrared detection, the higher sensitivity sensor mostly requires the larger bucky and expensive cryogenic cooling, although this provides better spatial resolution and determines longer detection distance. The low sensitivity uncooled IR imaging system needs a small f-number lens to increase the light signal transmitted through it. Both are subject to the critical component of photodetector which determines the overall system performance.

The low dimensional materials, e.g. carbon nanotube, graphene, have attracted attention since its discovery because of ultra high surface to volume ratio and near quasi 1D ballistic electronic transport. The novel materials have promising optoelectronic and optical properties. Meanwhile, single pixel compressive sensing based camera samples and reconstructs target signals overcoming the limits of Nyquist-Shannon theorem. This mechanism uses fewer sensing cells and measurements to reconstruct image using iterative algorithm. In this research, the CNT based infrared photodetector and infrared imaging system are built. Based on the results and discussions presented in previous chapters, the following conclusions can be made:

- **Nano Fabrication:** the real-time electrical feedback system was introduced into DEP assembly. By measuring the impedance changes, the system can quantitatively count the number of carbon nanotubes bridged between two electrodes. The experimental results show that the system response speed is fast enough to single wall and multi-wall CNTs alignment, although the impedance differs much. It reduces the time of CNT selection and sample clean compared to conventional method and will also be applicable to thin layer graphene deposition control and other nanomaterials deposition and localization.

- **Nano Sensor Characterization:** the in-situ measurement method using digital microscope and precise five axis substage localize the relative position between nanoscale sensor and IR beam by mapping the photocurrent on laser spot. The experimental results indicate that photovoltaic dominates photoresponse on CNT-metal Schottky photodetector although the photovoltaic voltage can not be sampled due to its tiny value merged in noise. The detector IR response are dependent on CNT-metal contact size and metal work function. The responsivity can be calculated by photocurrent mapping. The proposed measurement method provides a robust and precise approach to characterize sub-micro and nano scale photodetector which is important for sub-wavelength scale photodetector characterization.

- **Nano Sensor Reliability:** the reliability related characteristics of nanoelectronics and nano sensors could be summarized as multi-scale in both geometric and time domain. The former is highly dependent on sensor array uniformity while the latter nano-reliability measures the probability that a nano-scaled product performs its intended functionality without failure under given conditions for a specified period of time. The sandwich structured sensor was fabricated within two polymer layers, in which the substrate polyimide provided sensor with high bending stiffness, and top parylene packing blocked humidity environmental noise. The designed structure isolates stimulus except selected IR wavelength, especially from substrate charges. The proposed design improves sensor response and avoids failure under different environmental conditions. The fabrication, substrate material, and sensor structure are critical to CNT-IR photodetector reliability which is essential for the transition from nano science to nano engineering.

- **Weak Signal Readout:** the electrical properties of CNT IR photodetector are such complicated due to unknown internal structure. The weak signal intensity (nA), bias sensitive (zero bias) limit the readout design. In this research, the zero bias, high gain, two stage readout is designed to test the low current to sub nano ampere in CNT based IR detector. Meanwhile, it also works in CNT based single pixel IR imaging system. Although it works in hundreds hertz , that is not enough for fast high resolution image recovery. The ultra fast readout method is still a challenging project in nano electronics.

- **IR Compressive Imager & 3D:** a compressive sensing based IR camera system was developed. In the camera system, a single CNT field effect transistor integrated with a photonic cavity was employed to measure the compressed signals. This camera system was capable of observing the dynamic movement of a laser spot. The binocular 3D IR camera is also implemented in this system. This camera architecture provides a novel and alternative platform for future IR cameras, particularly the cameras using nano-photodetectors.

- **Compressive Light Field:** by taking multiple angular images through coded aperture, we can use single IR sensor to reconstruct a large object. Synthetic aperture photography will make the IR focus problem go away and also get 3D imaging using similar optics. The experimental results show that the more angular images will achieve the larger object recovery, better refocus in synthetic aperture photography. The double compressive sensing reduces the sampling ratio by angular image redundancy. It improves sampling speed with high accuracy.

- **Non Convex Compressive Video Sensing:** the non-convex compressive video sensing has advantage on the trade off problem between temporal and spatial resolution in video capturing. Besides, it requires very few number of measurements to record a real time dynamic scene. The non-convex sorted ℓ_1 approach has fast convergence to local minimizer and achieve a high accuracy. The edge detection based denoising reduces the accumulated residual error on frame difference so as to increase frame rate. Compared with the conventional compressive video sensing, the proposed ℓ_p solver will generate video frame in real time. This is the first reported real-time single pixel compressive video camera.

- **Infrared Super Resolution:** the super resolution is a special approach to solve infrared image low resolution problem. The 2D spline based RKHS model is used for image base (smooth component) while a modulated heaviside function is applied for image edge. The super resolution is casted into a model based estimation problem. By computing the coefficients of the redundant basis of low resolution image, it was applied as same in high resolution image for computing. The experimental results show the model is robust and easily modified with different IR image. The single low resolution image based super resolution will bring infrared technology into consumer electronics.

8.2 Future Research

There are three parts of future works from infrared sensor and its broadly imaging application.

- **Ultra High Sensitivity IR Sensor:** the sensitivity is a critical parameter to characterize infrared sensor. The conventional materials based technology has its limitation on thermal noise and the noise-equivalent temperature difference (NETD) is more than 20 mK. Another related parametric value is sensing area which dominates quantum efficiency. The single

CNT based IR photodetector has such tiny sensing area that it might not applicable for large area IR imager. Besides, the sensing area also depends on IR wavelength due to light diffraction. The 2D low dimensional materials, e.g. graphene and graphene derivative (rGO), Molybdenum disulfide (MoS_2), Phosphorene or its hibrid materials, have large area and good semiconductor performance. They will open a new approach for high sensitivity, high quantum efficiency IR photodetector.

- **Compressive IR Light Field Application:** infrared wavelength discovers distinct characteristics of object compared to visible imaging because the material reflection has unique distribution along whole electromagnetic spectrum especially on material, biology, plants and animals science. The broad infrared light field includes 3D, general light field and fast infrared video. These three technology will explore more distinct details for all applications above. The infrared microscopy together with 3D, light field and fast compressive video is a promising technology for the cutting-edge research.
- **IR Super Resolution:** there are two concepts on super resolution in scientific research. The first is algorithm based in computer science. It increases image spatial resolution from a low resolution image. The basic idea is to model the target and estimate the high resolution using signal processing. However, another super resolution is to resolve light diffraction limitation in optics and science research. The infrared wavelength is sub micro to tens micrometer so that it is extremely difficult to discover feathers underneath the wavelength. An extra hardware or firmware is required to increase sampling resolution. This super resolution is more challenging and more meaningful in microscopy imaging.

BIBLIOGRAPHY

BIBLIOGRAPHY

[1] Michael Rowan-Robinson. *Night Vision: Exploring the Infrared Universe*. Cambridge University Press, 2013.

[2] F. Szmulowicz, Gail J. Brown, Hui C. Liu, Aidong Shen, Zbigniew R. Wasilewski, and Margaret Buchanan. GaAs/AlGaAs p-type multiple quantum wells for infrared detection at noimal incidence: model and experiment. *Opto electronics review*, 9:164–173, 2001.

[3] Barbara H. Stuart. *Infrared Spectroscopy: Fundamentals and Applications*. Wiley, 2004.

[4] Christian W. Freudiger, Wenlong Yang, Gary R. Holtom, Nasser Peyghambarian, X. Sunney Xie, and Khanh Q. Kieu. Stimulated raman scattering microscopy with a robust fibre laser source. *Nature Photonics*, 8:153 – 159, 2014.

[5] H.Z. Chen, N. Xi, K.W. Lai, L.L. Chen, R.G. Yang, and B. Song. Gate dependent photoresponses of carbon nanotube field effect phototransistors. *Nanotechnology*, 23:385203, 2012.

[6] C. Hildebrandt, C. Raschner, and K. Ammer. An overview of recent application of medical infrared thermography in sports medicine in Austria. *IEEE Engineering in Medicine and Biology*, 1:21–57, 2010.

[7] J. F. Head, F. Wang, C. A. Lipari, and R. L. Elliott. The important role of infrared imaging in breast cancer. *Inegrative and Comparative Biology*, 48(365):50–59, 2000.

[8] M. Massoud. *Blackbody radiation. Engineering thermofluids: thermodynamics, fluid mechanics, and heat transfer*. Springer, 2005.

[9] J. Mehra and H. Rechenberg. *The Historical Development of Quantum Theory*. Springer-Verlag., 1982.

[10] A.P. French and E.F. Taylor. *An Introduction to Quantum Physics*. Norton Company, 1978.

[11] P. Madejczyk, W. Gawron, A. Piotrowski, K. Klos, J. Rutkowski, and A. Rogalski. Improvement in performance of high operating temperature HgCdTe photodiodes. *Infrared Physics and Technology*, 54:310–315, 2011.

[12] A. Rogalski. Infrared detectors: an overview. *Infrared Physics & Technology*, 43:187–210, 2001.

[13] G. L. Hansen, J. L. Schmit, and T. N. Casselman. Energy gap versus alloy composition and temperature in $Hg_{1-x}Cd_xTe$. *J. Appl. Phys*, 53:7099, 1982.

[14] Chinedu E. Ekuma, C. E. Singh, and D. J. Moreno. Optical properties of PbTe and PbSe. *Physical Review B*, 85:085205, 2012.

[15] Roger K. Richards, Donald P. Hutchinson, and Charles A. Bennett. Room-temperature QWIP detection at 10 μm. In *Proc. SPIE, Infrared Technology and Applications*, volume 4820, pages 1–4. SPIE, 2003.

[16] Yole. Uncooled infrared imaging technology and market trends 2014, January 2014.

[17] Markets. Infrared imaging market worth $8450 million by 2020, January 2014.

[18] FLIR. Infrared cameras from FLIR, January 2014.

[19] Fluke. Fluke infrared cameras are on the job because they do the job, January 2014.

[20] P.L. Richards. Bolometers for infrared and millimeter waves. *Journal of Applied Physics*, 76(1):1–24, 1994.

[21] S. Nudelman. The detectivity of infrared photodetectors. *Applied Optics*, 1(05):627–636, 1962.

[22] R. C. Jones. Proposal of the detectivity for detectors limited by radiation noise. *Journal of the Optical Society of America*, 50(11):1058–1059, 1960.

[23] Michael A. Kinch. *Fundamentals of Infrared Detector Materials*. SPIE Press, 2007.

[24] Van Thanh Daua, Takeo Yamadab, Dzung Viet Daoa, Bui Thanh Tunga, and Kenji Hatab. Integrated CNTs thin film for MEMS mechanical sensors. *Microelectronics Journal*, 41(12):860–864, 2010.

[25] Yan Wang, Zhiqiang Shi, Yi Huang, Yanfeng Ma, Chengyang Wang, Mingming Chen, and Yongsheng Chen. Supercapacitor devices based on graphene materials. *J. Phys. Chem*, 113:13103–13107, 2009.

[26] Maogang Gong, Tejas A. Shastry, Yu Xie, Marco Bernardi, Daniel Jasion, Kyle A. Luck, Tobin J. Marks, Jeffrey C. Grossman, Shenqiang Ren, and Mark C. Hersam. Polychiral semiconducting carbon nanotube fullerene solar cells. *Nano. Lett*, 14:5308–5314, 2014.

[27] Vera Sazonova, Yuval Yaish, Hande Ustunel, David Roundy, Tomas A. Arias, and Paul L. McEuen. A tunable carbon nanotube electromechanical oscillator. *Nature*, 431:284–287, 2004.

[28] M. Mittala and A. Kumara. Carbon nanotube (CNT) gas sensors for emissions from fossil fuel burning. *Sensors and Actuators B: Chemical*, 203:349–362, 2014.

[29] M. S. Marcus, J. M. Sinnons, O. M. Castellini, R. J. Hammers, and M.A. Eriksson. Photogating carbon nanotube transistors. *Journal of Applied Physics*, 100(07):084306, 2006.

[30] I. A. Levitsky and W. B. Euler. Photoconductivity of single-wall carbon nanotubes under continuous-wave near infrared illumination. *Applied Physics Letters*, 83:1857–1859, 2003.

[31] Akihiko Fujiwara, Yasuyuki Matsuoka, Hiroyoshi Suematsu, Naoki Ogawa, Kenjiro Miyano, Hiromichi Kataura, Yutaka Maniwa, Shinzo Suzuki, and Yohji Achiba. Photoconductivity in semi-conducting single walled carbon nanotubes. *Japanese Journal of Applied Physics*, 40:L1229–L1231, 2001.

[32] Jiangbo zhang, Ning Xi, Hongzhi Chen, King Wai Chiu Lai, Guanyong Li, and Uchechukwu C. Wejinya. Design, manufacturing, and testing of single carbon nanotube based infrared sensors. *IEEE transactions on nanotechnology*, 8(2):245–251, 2009.

[33] Jiangbo Zhang, Ning Xi, King W.C. Lai, Hongzhi Chen, and Yilun Luo. Single carbon nanotube based photodiodes for infrared detection. In *Proceeding of the 7th IEEE International Conference on Nanotechnology*, pages 1156–1160. IEEE, 2009.

[34] Jiangbo Zhang, Ning Xi, Hongzhi Chen, King W.C. Lai, Guanyong Li, and Uchechukwu Wejinya. Photovoltaic effect in single carbon nanotube-based schottky diodes. *Int. Journal Nanoparticles*, 1(2):108–118, 2008.

[35] Hongzhi Chen, Ning Xi, King W.C. Lai, Carmen K.M. Fung, and Ruiguo Yang. Development of infrared detectors using single carbon nanotubes based field effect transistors. *IEEE Transaction on Nanotechnology*, 5(9):582–589, 2010.

[36] Javey A., Guo J., Wang Q., Lundstrom M., and Dai H. Ballistic carbon nanotube transistors. *Nature*, 424:654–657, 2003.

[37] Collins P.G., Arnold M.S., and Avouris P. Engineering carbon nanotubes and nanotube circuits using electrical breakdown. *Science*, 292:706–709, 2001.

[38] Bockelmann U. and Bastard G. Phonon scattering and engergy relaxation in two one and zero dimensional electron gas. *Phys. Rev. B*, 42:8947–8951, 1990.

[39] Kuo D., Fang A., and Chang Y. Theoretical modeling of dark current and photo response for quantum well and quantum dot infrared detectors. *Infared Phys. Technol.*, 42:422–433, 2001.

[40] Da Jiang Yang, Qing Zhang, and George Chen. Thermal conductivity of multiwalled carbon nanotubes. *PHYSICAL REVIEW B*, 66:165440, 2002.

[41] Zhidong Hana and Alberto Finab. Thermal conductivity of carbon nanotubes and their polymer nanocomposites: A review. *Progress in Polymer Science*, 36:914–944, 2011.

[42] Rahat Afrin, Jibran Khaliq, Mohammad Islam, Iftikhar Hussain Gul, Arshad Saleem Bhatti, and Umair Manzoor. Synthesis of multiwalled carbon nanotube-based infrared radiation detector. *Sensors and Actuators A*, 187:73–78, 2012.

[43] M. Freitag, Y. Martin, J. A. Misewich, R. Martel, and Ph. Avouris. Photoconductivity of single carbon nanotubes. *Nano Letters*, 3:1067–1071, 2003.

[44] Basudev Pradhan, Kristina Setyowati, Haiying Liu, David H. Waldeck, and Jian Chen. Carbon nanotube polymer nanocomposite infrared sensor. *Nano Letters*, 8:1142–1146, 2008.

[45] R. Lukac. *Computational Photography: Methods and Applications*. CRC Press, 2010, 2010.

[46] S.D. Babacan, R. Ansorge, M. Luessi, P.R. Mataran, R. Molina, and A.K. Katsaggelos. Compressive light field sensing. *IEEE Transactions on image processing*, 21:4746–4758, 2012.

[47] S.K. Nayar and T. Mitsunaga. High dynamic range imaging: Spatially varying pixel exposures. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 472–479. IEEE, 2000.

[48] P. Sen, N.K. Kalantari, M. Yaesoubi, S. Darabi, D.B. Goldman, and E. Shechtman. Robust patch-based HDR reconstruction of dynamic scenes. *ACM Transactions on Graphics*, 31:203:1–203:12, 2012.

[49] J. Im, J. Jeon, M.H. Hayes, and J. Paik. Single image-based ghost-free high dynamic range imaging using local histogram stretching and spatially-adaptive denoising. *IEEE Transactions on Consumer Electronics*, 57:1478–1485, 2011.

[50] M. Levoy. Light fields and computational imaging. *IEEE Computer*, 8:45–56, 2006.

[51] Z. Xu, J. Ke, and E.Y. Lam. High-resolution lightfield photography using two masks. *Opt. Express*, 20:10971–10983, 2012.

[52] R.F. Marcia, Z.T. Harmany, and R.M. Willett. Compressive coded aperture imaging. In *in Computational Imaging VII, Proc. SPIE*, volume 7246, pages 1–13. SPIE, 2009.

[53] R.F. Marcia and R.M. Willett. Compressive coded aperture supperresolution image reconstruction. In *ICASSP 2008. IEEE International Conference on*, pages 833–836. IEEE, 2008.

[54] M.F. Duarte, M.A. Davenport, D. Takhar, J.N. Laska, K.F. Kelly T. Sun, and R.G. baraniuk. Single-pixel imaging via compressive sampling. *IEEE SIGNAL PROCESSING MAGAZINE*, 8:82–90, 2008.

[55] E. Candes, J. Romberg, and T. Tao. Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information. *IEEE Trans. Inform. Theory*, 52:489–509, 2006.

[56] E. Candes and J. Romberg. Quantitative robust uncertainty principles and optimally sparse decompositions. *Found. Comput. Math.*, 6:227–254, 2006.

[57] E. Candes and T. Tao. Decoding by linear programming. *IEEE Trans. Inform. Theory*, 51:4203–4215, 2005.

[58] Daniel Dregely, Richard Taubert, Jens Dorfmuller, Ralf Vogelgesang, Klaus Kern, and Harald Giessen. 3D optical Yagi-Uda nanoantenna array. *Nature Communications*, 2(267):1–7, 2011.

[59] J.R. Bergen E.H. Adelson. The plenoptic function and the elements of early vision. *In Computation Models of Visual Processing*, 51:3–20, 1991.

[60] Ren Ng. *Digital light field photography*. Stanford University, Stanford, CA, 2006.

[61] Sumio Iijima. Helical microtubules of graphitic carbon. *Nature*, 354(07):56–58, 1991.

[62] Kataura H., Kumazawa Y., Maniwa Y., Umezu I., Suzuki S., Ohtsuka Y., and Achiba Y. Optical properties of single wall carbon nanotubes. *Synthetic Metals*, 103(01):2555–2558, 1999.

[63] Collins P.G., Arnold M.S., and Avouris P. Engineering carbon nanotubes and nanotube circuits using electrical breakdown. *Science*, 292:706–709, 2001.

[64] Y. Matsuda, W. Deng, and W. Goddard. Contact resistance properties between nanotubes and various metals from quantum mechanics. *J. Phys. Chern*, I11:1113–11116, 2007.

[65] I. Deretzis and A. Magna. Role of contact bonding on electronic transport in metal-carbon nanotube-metal systems. *Nanotechnology*, 17:5063, 2006.

[66] Y. Chai, A . Hazeghi, K . Takei, Hong-Yu Chen, Chan Philip C.H., Javey A., and Wong H.S.P. Low-resistance electrical contact to carbon nanotubes with graphitic interfacial layer. *IEEE Trans. on Electron Devices*, 59:12–20, 2012.

[67] Carla M. Aguirre, Pierre L. Levesque, Matthieu Paillet, Francois Lapointe, Benoit C. St-Antoine, Patrick Desjardins, and Richard Martel. The role of the oxygen/water redox couple in suppressing electron conduction in field-effect transistor. *Advanced Materials*, 21:3087–3091, 2009.

[68] Harish M. Manohara, Eric W. Wong, Erich Schlecht, Brian D. Hunt, and Peter H. Siegel. Carbon nanotube schottky diodes using Ti schottky and Pt ohmic contacts for high frequency applications. *Nano Letters*, 5(7):1469 – 1474, 2005.

[69] Hongzhi Chen, Ning Xi, Bo Song, Liangliang Chen, Jianguo Zhao, King Wai Chiu Lai, and Ruiguo Yang. Infrared camera using a single nano-photodetector. *IEEE Sensors Journal*, 13:949–958, 2013.

[70] Vikrant Jayant Gokhale, Olga A. Shenderova, Gary E. McGuire, and Rais-Zadeh M. Infrared absorption properties of carbon nanotube/nanodiamond based thin film coatings. *JOURNAL OF MICROELECTROMECHANICAL SYSTEMS*, 23(1):191–197, 2014.

[71] M. Monthioux and V. Kuznetsov. Who should be given the credit for the discovery of carbon nanotubes. *Carbon*, 44(9):1621, 2006.

[72] Feng Yang, Xiao Wang, Daqi Zhang, Juan Yang, Da Luo, Ziwei Xu, Jiake Wei, Jian-Qiang Wang, Zhi Xu, Fei Peng, Xuemei Li, Ruoming Li, Yilun Li, Meihui Li, Xuedong Bai, Feng Ding, and Yan Li. Chirality-specific growth of single-walled carbon nanotubes on solid alloy catalysts. *Nature*, 510:522–524, 2014.

[73] Mukul Kumar and Yoshinori Ando. Chemical vapor deposition of carbon nanotubes: A review on growth mechanism and mass production. *Journal of Nanoscience and Nanotechnology*, 10:3739–3758, 2010.

[74] Martial Duchamp, Kyumin Lee, Benjamin Dwir, Jin Won Seo, Eli Kapon, Laszlo Forro, and Arnaud Magrez. Controlled positioning of carbon nanotubes by dielectrophoresis: Insights into the solvent and substrate role. *ACS nano*, 4(1):279–284, 2010.

[75] Zhifeng Deng, Erhan Yenilmez, Amy Reilein, Joshua Leu, Hongjie Dai, and Kathryn A. Moler. Nanotube manipulation with focused ion beam. *Applied Physics Letters*, 88(1):023119, 2006.

[76] Libao An and Craig R. Friedrich. Real-time gap impedance monitoring of dielectrophoretic assembly of multiwalled carbon nanotubes. *Applied Physics Letters*, 92:173103–173106, 2008.

[77] Sebastian Sorgenfrei, Inanc Meric, Sarbajit Banerjee, Austin Akey, Sami Rosenblatt, Irving P. Herman, and Kenneth L. Shepard. Controlled dielectrophoretic assembly of carbon nanotubes using real-time electrical detection. *Applied Physics Letter*, 94:053105 – 053108, 2009.

[78] Ming Zheng, Anand Jagota, Michael S. Strano, Adelina P. Santos, Paul Barone, S. Grace Chou, Bruce A. Diner, Mildred S. Dresselhaus, Robert S. Mclean, G. Bibiana Onoa, Georgii G. Samsonidze, Ellen D. Semke, Monica Usrey, and Dennis J. Walls. Structure-based carbon nanotube sorting by sequence-dependent DNA assembly. *Science*, 302(5650):1545 – 1548, 2003.

[79] Didi Xu, A. Subramanian, Lixin Dong, and B.J. Nelson. Shaping nanoelectrodes for high-precision dielectrophoretic assembly of carbon nanotubes. *IEEE TRANSACTIONS ON NANOTECHNOLOGY*, 8(4):449–460, 2009.

[80] S. E. Harding, M. Dampier, and A. J. Rowe. Modeling biological macromolecules in solution: The general tri-axial ellipsoid. *IRCS Med. Sci.*, 7:1813–1829, 1979.

[81] E.W. Small and I. Isenberg. Hydrodynamic properties of a rigid molecule: Rotational and linear diffusion and fluorescence anisotropy. *Biopolymers*, 16:1907–1928, 1977.

[82] King W. C. Lai, Ning Xi, and Uchechukwu C. Wejinya. Automated process for selection of carbon nanotube by electronic property using dielectrophoretic manipulation. *Journal of Micro-Nano Mechatronics*, 4:37–48, 2008.

[83] Chechukwu C. Wejinya, Yantao Shen, Ning Xi, King W. C. Lai, and Jiangbo Zhang. An efficient approach of handling and deposition of micro and nano entities using sensorized microfluidic end-effector system. *Sensors and Actuators A: Physical*, 147:6–16, 2008.

[84] Tsung-Tsong Wu, Yung-Yu Chen, and Tai-Hsu Chou. A high sensitivity nanomaterial based SAW humidity sensor. *Journal of Physics D: Applied Physics*, 41(8):085101, 2008.

[85] J. Svensson and E. Campbell. Schottky barriers in carbon nanotube-metal contacts. *Journal of Applied Physics*, 110:111101, 2011.

[86] H.L. Skriver and N.M. Rosengaard. Surface energy and work function of elemental metals. *Physical Review B*, 46:7157–7168, 1992.

[87] Mitsumasa IWAMOTO, Atsushi FUKUDA, and Fuminori HIRAYAM. Charge storage phenomena observed in ultra thin polyimide langmuir blodgett films deposited on metal electrodes. In *Proceeding of the 3rd internaltional conference on properties and applications of dielectric materials*, pages 193–196, 1991.

[88] Miao Yu, Liangliang Chen, K. W. C. Lai, C. K. M. Fung, N. Xi, L. Zhang, and L. Dong. 3D assembly and simulation of helical optical antenna-enhanced carbon nanotube IR sensors. In *Nanotechnology (IEEE-NANO), 2014 IEEE 14th International Conference on*, pages 849–852. IEEE, 2014.

[89] D. J. Bell, L. Dong, and B. J. Nelson. Fabrication and characterization of three-dimensional InGaAs/GaAs nanosprings. *Nano Letters*, 6(4):725–729, 2006.

[90] L. Chen, N. Xi, Z. Zhou, and B. Song. Quantitatively control of carbon nanotubes using real time electrical detection dielectrophoresis assembly. In *Nanotechnology (IEEE-NANO), 2015 15th IEEE Conference on*, pages 1029–1032. IEEE, 2015.

[91] X. Zhang and Z. Liu. Superlenses to overcome the diffraction limit. *Nature*, 7:435–441, 2008.

[92] J. U. Lee. Photovoltaic effect in ideal carbon nanotube diodes. *Applied Physics Letters*, 87:073101–1C073101–3, 2005.

[93] A. ROGALSKI and K. CHRZANOWSKI. Infrared devices and techniques. *Opto-Electronics Review*, 10:111–136, 2002.

[94] K. Besteman, J.-O. Lee, F. G. M. Wiertz, H. A. Heering, and C. Dekker. Enzyme-coated carbon nanotubes as single-molecule biosensors. *Nano Letter*, 3:727–730, 2003.

[95] S. J. Wind, J. Appenzeller, R. Martel, V. Derycke, and P. Avouris. Vertical scaling of carbon nanotube field-effect transistors using top gate electrodes. *Applied Physics Letters*, 80:3817–3819, 2002.

[96] Michael B. Wakin, Jason N. Laska, Marco F. Duarte, Dror Baron, Shriram Sarvotham, Dharmpal Takhar, Kevin F. Kelly, and Richard G. Baraniuk. Compressive imaging for video representation and coding. In *2006 Proceedings of Picture Coding Symposium (PCS)*, pages 1–6. IEEE, 2006.

[97] Yanfei Wang, Jingjie Cao, and Changchun Yang. Recovery of seismic wavefields based on compressive sensing by an ℓ_1-norm constrained trust region method and the piecewise random sub-sampling. *Geophys. J. Int.*, 000(1):1–9, 2010.

[98] Bo Song, Jianguo Zhao, Ning Xi, Hongzhi Chen, King Wai Chiu Lai, Ruiguo Yang, and Liangliang Chen. Compressive feedback-based motion control for nanomanipulation: theory and applications. *IEEE Transactions on Robotics*, 30(1):103–114, 2014.

[99] Peng Zhang, Zhen Hu, R.C. Qiu, and B.M. Sadler. A compressed sensing based ultra-wideband communication system. In *IEEE International Conference on Communications, 2009.*, pages 1–5. IEEE, 2009.

[100] Michael Lustig, David L. Donoho, Juan M. Santos, and John M. Pauly. Compressed sensing MRI. *IEEE SIGNAL PROCESSING MAGAZINE*, 3:72–82, 2008.

[101] M.A. Davenport, P.T. Boufounos, M.B. Wakin, and R.G. Baraniuk. Signal processing with compressive measurements. *IEEE Journal of Selected Topics in Signal Processing*, 4(2):445–460, 2010.

[102] E. Candes and M. Wakin. An introduction to compressive sampling. *IEEE Signal Processing Magazine*, 25(2):21–30, 2008.

[103] Erwan Le Pennec and Stphane Mallat. Sparse geometric image representation with bandelets. *IEEE Transactions on Image Processing*, 14(4):423–438, 2005.

[104] David L. Donoho. Compressed sensing. *IEEE Transactions on Information Theory*, 52(4):1289–1306, 2006.

[105] A. Cohen, W. Dahmen, and R. DeVore. Compressed sensing and best K-term approximation. *J. Am. Math. Soc.*, 22(1):211–231, 2009.

[106] Joel A. Tropp. Greed is good: algorithmic results for sparse approximation. *IEEE Transactions on Information Theory*, 50(10):2231C2242, 2004.

[107] G. M. Davis, S. Mallat, and M. Avellaneda. Adaptive greedy approximations. *Journal of Constructive Approximation*, 13(1):57–98, 1997.

[108] P. L. Combettes and V. R. Wajs. Signal recovery by proximal forward-backward splitting. *Multiscale Modeling and Simulation*, 4(4):1168–1200, 2005.

[109] D. Needell and J.A. Tropp. COSAMP: iterative signal recovery from incomplete and inaccurate samples. *Appl. Comput. Harmon. Anal.*, 26(3):301–321, 2009.

[110] W. Dai and O. Milenkovic. Subspace pursuit for compressive sensing signal reconstruction. *IEEE Trans. Inf. Theory*, 55(5):2230–2249, 2009.

[111] A. Beck and M. Teboulle. A fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM Journal on Imaging Sciences*, 2(1):183–202, 2009.

[112] J. Yang and Y. Zhang. Alternating direction algorithms for ℓ_1-problems in compressive sensing. *SIAM Journal on Scientific Computing*, 33(1):250–278, 2011.

[113] J. Yang and X. Yuan. Linearized augmented lagrangian and alternating direction methods for nuclear norm minimization. *Mathematics of Computation*, 82(281):301–329, 2013.

[114] L. Chen, N. Xi, H. Chen, and King Wai Chiu Lai. Readout system design for MWCNT infrared sensors. In *Proceeding of the IEEE Nanotechnology Conference, 2011*, pages 1–4. IEEE, 2011.

[115] J. G. Graeme. *Photodiode Amplifiers, OPAMP Solutions*. McGraw-Hill Press, New York, 1995.

[116] Donald A. Neamen. *Semiconductor Physics And Devices: Basic Principles*. McGraw-Hill Press, New York, 2003.

[117] Michael Faraday. Thoughts on ray vibrations. *Experimental Researches in Electricity*, 3:447–452, 1846.

[118] A. Gershun. The light field. *Translated by P. Moon and G. Timoshenko in Journal of Mathematics and Physics*, 18:51–151, 1939.

[119] R. Ng. Fourier slice photography. *ACM Trans. Graph.*, 24(3):735–744, 2005.

[120] Gordon Wetzstein, Douglas Lanman, Matthew Hirsch, and Ramesh Raskar. Tensor displays: Compressive light field synthesis using multilayer displays with directional backlighting. *ACM Trans. Graph.*, 31:1–11, 2012.

[121] A. Lumsdaine and T. Georgiew. The focused plenoptic camera. In *In Proc. ICCP.*, pages 1–8. ACM, 2009.

[122] Chia-Kai Liang, Tai-Hsu Lin, Bing-Yi Wong, Chi Liu, and Homer H. Chen. Programmable aperture photography: Multiplexed light field acquisition. *ACM Trans. Graph.*, 27(3):1–10, 2008.

[123] M. Kamal, M. Golbabaee, and P. Vandergheynst. Light field compressive sensing in camera arrays. In *In Proc. ICASSP.*, pages 5413–5416. ACM, 2008.

[124] Edward H. Adelson and John Y.A. Wang. Single lens stereo with a plenoptic camera. *IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE*, 14(2):99–106, 1992.

[125] Marc Levoy, Billy Chen, Vaibhav Vaish, Mark Horowitz, Ian McDowall, and Mark Bolas. Synthetic aperture confocal imaging. *Acm Transactions on Graphics(Proceedings of siggraph 2004)*, 23(3):822–831, 2004.

[126] M. Levoy and P. Hanrahan. Lightfield rendering. In *Proc. SIGGRAPH*, pages 31–42. ACM Press, 1996.

[127] S. Gortler, R. Grzeszczuk, R. Szeliski, and M. F. Cohen. The lumigraph. In *Proc. SIGGRAPH*, pages 43–54. ACM Press, 1996.

[128] Chengbo Li. *An Efficient Algorithm For Total Variation Regularization with Applications to the Single Pixel Camera and Compressive Sensing*. Rice University, 2009, 2009.

[129] D. Goldfarb and W. Yin. Second-order cone programming methods for total variation based image restoration. *SIAM Journal on Scientific Computing*, 27:622–645, 2005.

[130] J. Bioucas-Dias and M. Figueiredo. A new TwIST: Two-step iterative thresholding algorithm for image restoration. *IEEE Trans. Imag. Process.*, 16:2992–3004, 2007.

[131] Junfeng Yang, Yin Zhang, and Wotao Yin. A fast alternating direction method for TVL1-L2 signal reconstruction from partial fourier data. *IEEE JOURNAL OF SELECTED TOPICS IN SIGNAL PROCESSING*, 4(2):288–297, 2010.

[132] J. Harguessa and S. Strangeb. Infrared stereo calibration for unmanned ground vehicle navigation. In *Proc. SPIE 9084, Unmanned Systems Technology XVI*, pages 1–8. SPIE, 2014.

[133] P. Chakravarty and R. Jarvis. Panoramic vision and laser range finder fusion for multiple person tracking. In *Proc. IEEE/RSJ Conf of Intelligent Robots and Systems, Beijing, China*, pages 2949–2954. IEEE, 2006.

[134] Jeremy M. Wolfe. Stereopsis and binocular rivalry. *Psychological Review*, 93(3):269–282, 1986.

[135] D. Geronimo, A.M. Lopez, A.D. Sappa, and T. Graf. Survey of pedestrian detection for advanced driver assistance systems. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 32(7):1239–1258, 2010.

[136] Reinhard Koch, Ingo Schiller, Bogumil Bartczak, Falko Kellner, and Kevin Koser. MixIn3D: 3D mixed reality with ToF-camera. *Dynamic 3D Imaging*, 5742(1):126–141, 2010.

[137] B. Huhle, P. Jenke, and W. Strasser. On-the-fly scene acquisition with a handy multi-sensor system. *International Journal of Intelligent Systems Technologies and Applications*, 5(3):255–263, 2008.

[138] M. Hansard, S. Lee, O. Choi, and R. Horaud. *Time of Flight Cameras: Principles, Methods, and Applications*. SpringerBriefs in Computer Science, 2012.

[139] M. Fahle. Wozu zwei augen? [why two eyes?]. *Naturwissenschaften*, 74:383–385, 1987.

[140] R. Blake and R. Sekuler. *Perception (5th ed.)*. New York, NY: McGraw-Hill, 2006.

[141] M. Do and M. Vetterli. The finite ridgelet transform for image representation. *IEEE Transactions on Information Theory*, 12:16–28, 2003.

[142] T. Komatsu, K. Aizawa, T. Igarashi, and T. Saito. Signal-processing based method for acquiring very high resolution image with multiple cameras and its theoretical analysis. *Proc. Inst. Elec. Eng.*, 140(1):19–25, 1993.

[143] S. Park, M. Park, and M. Kang. Super-resolution image reconstruction: a technical overview. *Signal Processing Magazine, IEEE*, 20(3):21–36, 2003.

[144] M. Irani and S. Peleg. Super resolution from image sequences. In *Proceedings of the International Conference on Pattern Recognition*, pages 115–120. IEEE, 1990.

[145] R. R. Schultz and R. L. Stevenson. Extraction of high-resolution frames from video sequences. *IEEE Transactions on Image Processing*, 5(6):996–1011, 1996.

[146] M. K. Ng, H. Shen, E. Y. Lam, and L. Zhang. A total variation regularization based super-resolution reconstruction algorithm for digital video. *EURASIP Journal on Advances in Signal Processing*, 1:1–16, 2007.

[147] M. V. Joshi, S. Chaudhuri, and R. Panuganti. A learning-based method for image super-resolution from zoomed observations. *IEEE Transactions on Systems, Man, and Cybernetics*, 35(3):527–537, 2005.

[148] A. N. Rajagopalan and S. Chaudhuri. Space-variant approaches to recovery of depth from defocused images. *Computer Vision and Image Understanding*, 68:309–329, 1997.

[149] H. Shen, L. Zhang, B. Huang, and P. Li. A MAP approach for joint motion estimation, segmentation and super-resolution. *IEEE Transactions on Image Processing*, 16(2):479–490, 2007.

[150] M. Protter, M. Elad, H. Takeda, and P. Milanfar. Generalizing the nonlocal-means to super-resolution reconstruction. *IEEE Transactions on Image Processing*, 18(1):36–51, 2009.

[151] N. Nguyen, P. Milanfar, and G. Golub. A computationally efficient superresolution image reconstruction algorithm. *IEEE Transactions on Image Processing*, 10:573–583, 2001.

[152] D. Glasner, S. Bagon, and M. Irani. Super-resolution from a single image. In *International Conference on Computer Vision (ICCV)*, pages 1–8. IEEE, 2009.

[153] X. Li and M. Orchard. New edge-directed interpolation. *IEEE Trans. Image Processing*, 10:1521–1527, 2001.

[154] N. Mueller, Y. Lu, and M. Do. Image interpolation using multiscale geometric representations. In *SPIE proceedings Computational Imaging V*, volume 6498, pages 1–11. SPIE, 2007.

[155] L. Zhang and X. Wu. An edge-guided image interpolation algorithm via directional filtering and data fusion. *IEEE Transactions on Image Processing*, 15:2226–2238, 2006.

[156] R. Fattal. Image upsampling via imposed edge statistics. *ACM Transactions on Graphics*, 26:1, 2007.

[157] J. Sun, J. Sun, Z. Xu, and H.Y. Shum. Image super-resolution using gradient profile prior. In *2008 CVPR*, pages 1–8. ACM, 2008.

[158] Q. Xie, H. Chen, and H. Cao. Improved example-based single-image superresolution. *International Congress on Image and Signal Processing*, 3:1204–1207, 2010.

[159] J. Yang, J. Wright, T. Huang, and Y. Ma. Image super-resolution via sparse representation. *IEEE transactions on image processing*, 19:2861–1873, 2010.

[160] W. Yin, S. Osher, D. Goldfarb, and J. Darbon. Bregman iterative algorithm for ℓ_1 minimization with applications to compressed sensing. *SIAM J. Imaging Sciences*, 1:143–168, 2008.

[161] E.T. Hale, W. Yin, and Y. Zhang. Fixed-point continuation for ℓ_1-minimization: Methodology and convergence. *SIAM Journal on Optimization*, 19:1107–1130, 2008.

[162] L. Bregman. The relaxation method of finding the common point of convex sets and its application to the solution of problems in convex programming. *USSR Computation Mathematics and Mathematical Physics*, 7:200–217, 1967.

[163] Y. Li and S. Osher. Coordinate descent optimization for ℓ_1 minimization with application to compressed sensing; a greedy algorithm. *Inverse Problems and Imaging*, 3:487–503, 2009.

[164] R. Chartrand. Fast algorithms for nonconvex compressive sensing: MRI reconstruction from very few data. *Biomedical Imaging: From Nano to Macro, 2009. ISBI*, 1:262–265, 2009.

[165] L. Bai, P. Maechler, M. Muehlberghuber, and H. Kaeslin. High-speed compressed sensing reconstruction on FPGA using OMP and AMP. In *in Electronics, Circuits and Systems (ICECS), 19th IEEE International Conference on*, pages 53–57. IEEE, 2012.

[166] Jinyu Chu, Ju Liu, Jianping Qiao, Xiaoling Wang, and Yujun Li. Gradient-based adaptive interpolation in super-resolution image restoration. In *Signal Processing, 2008. ICSP 2008. 9th International Conference on*, pages 1027–1030. IEEE, 2008.

[167] Liang jian Deng, Weihong Guo, and Tingzhu Huang. Single image super-resolution via an iterative reproducing kernel hilbert space method. *IEEE Transactions on Circuits and Systems for Video Technology*, 1051:1–14, 2015.

[168] G. Wahba. Spline models for observational data. *SIAM. Regional Conference Series in Applied Mathematics*, 59, 1990.

[169] Y. Wang, J. Yang, W. Yin, and Y. Zhang. A new alternating minimization algorithm for total variation image reconstruction. *SIAM Journal on Imaging Sciences*, 1:248–272, 2008.

[170] T. Vaughan. *Multimedia: Making it Work (4th ed.)*. Berkeley, CA: McGraw-Hill, 1998.

[171] M. Irvin, T. Kitazawa, and T. Suzuki. A new generation of MPEG-2 video encoder ASIC andits application to new technology markets. *Broadcasting Convention, International*, 1:391–396, 1996.

[172] G. J. SULLIVAN and T. WIEGAND. Video compression from concepts to the H.264/AVC standard. *PROCEEDINGS OF THE IEEE*, 93(1):18–31, 2005.

[173] *http://www.gigapan.com/gigapans/102753*. Gigapan, 2015.

[174] M. Hepher, D. Duckett, and A. Loening. High-speed video microscopy and computer enhanced imagery in the pursuit of bubble dynamics. *Ultrason Sonochem.*, 7:229–33, 2000.

[175] Q. Cossairt, M. Gupta, and S. Nayar. When does computational imaging improve performance? *IEEE TRANSACTIONS ON IMAGE PROCESSING*, 22(2):447, 2013.

[176] V. Stankovic, L. Stankovic, and S. Cheng. Compressive video sampling. In *16th Eur. Signal Processing Conf.*, pages 1–6, 2008.

[177] J. Zheng and E. L. Jacobs. Video compressive sensing using spatial domain sparsity. *Opt. Eng.*, 48:087006, 2009.

[178] A. Veeraraghavan, D. Reddy, and R. Raskar. Coded strobing photography: Compressive sensing of high speed periodic videos. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(4):671–686, 2011.

[179] A. Sankaranarayanan, P. Turaga, R. Baraniuk, and R. Chellappa. Compressive acquisition of dynamic scenes. In *in Proc. 11th European Conf. on Comput. Vis.*, pages 129–142, 2010.

[180] M. Lai and N. Kingsbury. Coding of image sequences using coarse quantization and feature based hierarchical block matching. In *International Conference on Image Processing and its Applications*, pages 586–589, 1992.

[181] N. Nasrabadi and R. King. Image coding using vector quantization: A review. *IEEE TRANSACTIONS ON COMMUNICATIONS*, 36(8):957–971, 1988.

[182] D. Hein and N. Ahmed. Video compression using conditional replenishment and motion prediction. *IEEE TRANSACTIONS ON ELECTROMAGNETIC COMPATIBILITY*, EMC-26(3):134–142, 1984.

[183] E. Chan, A. Rodriguez, R. Ghandi, and S. Panchanathan. Experiments on block-matching techniques for video coding. *Multimedia Systems*, 2(5):228–241, 1994.

[184] D. Lin, C. Chen, and T. Hsing. Video on phone lines : technology and applications. *Proceedings of the IEEE*, 83(2):175–193, 1995.

[185] F. Lauuauret and D. Barba. Motion compensation by block matching and vector post-processing in sub-band coding of TV signals at 15 Mbit/s. *Proceedings of SPIE*, 1605:26–36, 1991.

[186] Rayan Saab and Ozgur Yilmaz. Sparse recovery by non-convex optimization-instance optimality. *Appl.Comput. Harmon. Anal.*, 29(1):30–48, 2010.

[187] Z. Xu, X. Chang, F. Xu, and H. Zhang. $\ell_{1/2}$ regularization: a thresholding representation theory and a fast solver. *IEEE Trans. Neural Netw. Learn. Syst.*, 23(7):1013–1027, 2012.

[188] Xiao-Lin Huang, Lei Shi, and Ming Yan. Nonconvex sorted ℓ_1 minimization for sparse approximation. *J. Oper. Res. Soc. China*, 3:207–229, 2015.

[189] E.J. Candes, M.B. Wakin, and S.P. Boyd. Enhancing sparsity by reweighted ℓ_1 minimization. *J. Fourier Anal. Appl.*, 14(5):877–905, 2008.

[190] G. Gasso, A. Rakotomamonjy, and S. Canu. Recovering sparse signals with a certain family of nonconvex penalties and DC programming. *IEEE Transactions on Signal Processing*, 57(12):4686–4698, 2009.

[191] R. Chartrand. Exact reconstruction of sparse signals via nonconvex minimization. *IEEE Signaln Processing Letters*, 14(10):707–710, 2007.

[192] Yilun Wang and Wotao Yin. Sparse signal reconstruction via iterative support detection. *SIAM J. Imaging Sci.*, 3(3):462–491, 2010.

[193] O. R. Vincent and O. Folorunso. A descriptive algorithm for sobel image edge detection. In *Proceedings of Informing Science and IT Education Conference*, pages 1–11, 2009.

[194] G.T. Shrivakshan. A comparison of various edge detection techniques used in image processing. *International Journal of Computer Science Issues*, 9(1):269–276, 2012.

[195] J. Yang and Y. Zhang. Alternating direction algorithms for ℓ_1-problems in compressive sensing. *SIAM Journal on Scientific Computing*, 33(1):250–278, 2011.