NORTHWESTERN UNIVERSITY

Attribute Invariant Spatial and Spatio-Temporal Correlators

A DISSERTATION

SUBMITTED TO THE GRADUATE SCHOOL
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

for the degree

DOCTOR OF PHILOSOPHY

Field of Electric Engineering

by

Mehjabin Sultana Monjur

EVANSTON, ILLINOIS

September 2016

ABSTRACT

Attribute Invariant Spatial and Spatio-Temporal Correlators

Mehjabin Sultana Monjur

Target identification and tracking is important in many defense and civilian applications. Optical correlators provide a simple technique for fast verification and identification of data. The primary limitation of such Conventional Holographic Correlators (CHC) is the poor nature of the material used for recording. We develop the concept and design of, and demonstrate experimentally, a Hybrid Opto-electronic Correlator (HOC) using only photo-detectors, Spatial Light Modulator (SLM) and VLSI chips that overcomes the material problems of the CHC. In the HOC, the amplitude and phase information of the query as well as the reference image is recorded with photo-detectors via interference with plane waves. The output of such an HOC has four terms: two convolution signals and two cross-correlation signals. We show through simulation and experimental results, that the correlation signal depends non-trivially on both the *relative* and the *absolute* phases of these two plane waves at the planes of interference. We also demonstrate experimentally, a functional HOC for identical and different reference and query images under optimized phase conditions. We develop a technique for controlling and optimizing the output signal, such that the convolution terms can be eliminated, by using a single servo, aided by repeated calibrations with known objects and reference images, interlaced

with the actual correlator operation. In addition, To achieve the ultimate speed of such a correlator, we also propose an integrated graphic processing unit which would perform all the electrical processes in a parallel manner. We also show that the HOC is capable of detecting objects in a scale and rotation invariant manner, along with the shift invariance feature, by incorporating Polar Mellin Transform (PMT). Furthermore, we demonstrate how to carry out shift, rotation and scale invariant detection of multiple matching objects simultaneously, a process previously thought to be incompatible with the PMT based correlators.

We also develop the concept and design of an Automatic Event Recognition (AER) System based on a three-dimensional Spatio-Temporal Correlator (STC), that combines the techniques of holographic correlation and photon echo based temporal pattern recognition to match a video-clip contained in a video file, using atoms stored in a porous-glass material. By employing the nonlinear properties of inhomogenous broadened atomic media we show that it is possible to realize an AER system that can recognize rapidly the occurrence of events, the number of events, and the occurrence times. To model the response of such a system, one requires solving the Schrodinger Equation (SE), which is a computationally extensive task. We develop an analytical model to find the response of the STC and show that the analytical model agrees closely with the results obtained via explicit numerical simulation, but at a speed that is many orders of magnitude faster than the numerical model. We also show how such a practical AER system can be realized using a combination of a porous-glass based Rb vapor cell, a holographic video disc, and a lithium niobate crystal.

Acknowledgements

Firstly, I would like to express my sincere gratitude to my advisor Prof. Selim Shahriar for the continuous support in my PhD study and related research, for his patience, motivation, and immense knowledge. His guidance helped me in all the time of research and writing of this thesis. I could not have imagined having a better advisor and mentor for my PhD study.

Besides my advisor, I would like to thank the rest of my thesis committee members: Prof. Seng-Tiong Ho and Prof. Oliver S Cossairt for their insightful comments, encouragement, and helpful advices.

My sincere thanks go to my colleagues Shih, Mohamed, Minchuan, Subra, Tony, May, Zifan, Renpeng and Josh for their kind support and collaboration. I have always felt like a family from the very first day at Northwestern University. They always helped me whenever I have asked for. Special thanks to Shih who was like a mentor to teach me the experimental aspects of my research and has been an amazing help throughout this time.

I would like to extend my heartfelt gratitude to my family: my parents and my brothers for their continuous support and encouragement. It was my parent’s dream that I would pursue the ultimate degree and I am glad that I have accomplished it. They had to make enormous sacrifices to make it possible and for this, I will be grateful forever. Their constant moral support and prayers have pulled me through turbulent times. I would also like to thank my in-laws for their understanding and patience while took the challenging endeavor.

Last but not the least, I am indebted to my husband, Ashraf Alom, who decided to accompany me throughout the PhD journey. Without his continuous support and encouragement, this tremendous journey would be impossible. I am blessed to have him as my life partner. I dedicate my thesis to my parents as well as to my husband.

Table of Contents

List of Figures

List of Tables

Chapter 1 Introduction

1.1 Motivation

Target identification and tracking is important in many defense and civilian applications. Optical correlators provide a simple technique for fast verification and identification of data. Over recent years, we have been investigating the feasibility of realizing an all-optical high-speed automatic target recognition (ATR) correlator system using the inherent parallelism of optical techniques[1,2,3,4,5,6,7]. Other groups have also pursued the development of such correlators[8,9]. The simplest form of such a system is the basic Vander Lugt[10] optical correlator. The limitations of such an correlator is that the recording process is very time consuming. This constraint is circumvented in a Joint Transform Correlator (JTC), where a dynamic material such as photorefractive polymer film is used so that the recording and correlation take place simultaneously. A JTC makes use of a dynamic non-linear material, such as photorefractive thin film produced by Nitto-Denko[1,2]. The primary limitation of this system is the poor nature of the material used for making the JTC. First, it is very fragile, and gets destroyed rather easily for reasons not well understood. Second, the diffraction efficiency is rather small, and it produces a lot of scattering, leading to a very poor signal to noise ratio (SNR). Third, after carrying out some correlations, the residual gratings generated in the medium have to be erased by applying a high voltage; this process takes quite a bit of time. To our knowledge, all dynamic holographic films suffer from similar limitations.

We propose a scheme that gets around this problem by making use of concepts developed recently in the context of digital holography[10]. Namely, the nonlinearity provided by the JTC medium is replaced by the nonlinearity of high-speed detectors (since detectors measure intensities, they are inherently non-linear). Of course, this requires some modification of the architecture, as well as post-processing of signals. We propose an explicit and novel architecture that enables the process of correlating images using detectors only, and the phase information is conserved by interfering with plane waves. We would like to point out that Javidi et al.[11] demonstrated a JTC that also makes use of detectors, in a manner similar to what we described here. However, this approach has some key limitations. The reference and the query images have to be placed in the same field of view. In order to do so, it is necessary to convert each image to a digital format first, and then create a composite image, which is sent to a Spatial Light Modulator (SLM), for example. This process precludes the use of a scenario, necessary for very rapid search, where the reference images are retrieved directly from a holographic memory disc. For this scenario, it would be necessary to combine the reference image, which is in the optical domain, with the query image, also in the optical domain after generation from an SLM, using a beam splitter. When this is done, the correlation will depend sinusoidally on the relative phase difference between the two optical paths, with no obvious option for stabilizing or optimizing this phase difference. Our approach presented here circumvents both of these constraints.

The proposed Hybrid Optoelectronic Correlator (HOC) is capable of shift-invariant image recognition, in the same manner as what is achieved with a conventional holographic correlator (CHC).

However, the actual architecture is very different, requiring many intermediate steps, servos and post processing. Furthermore, the output signals are also different (e.g., it contains a cross-correlation term and an anti cross-correlation term, but no convolution nor dc term).Thus, it may not be a priori obvious whether the HOC can also perform scale and rotation invariant correlation using polar mellin transform (PMT)[12,13,14,15,16,17,18,19].

We show how, indeed, it is possible to use PMT to achieve scale and rotation invariant image recognition with the HOC architecture. Furthermore, we identify some limitations of using PMT in the CHC architecture proposed previously, and show how to overcome these constraints by proper pre-processing of images for both CHC and HOC.

Automated target recognition (ATR) has been a very active field of research for several decades. Significant advances in ATR have been made using analog approaches employing holographic correlators, as well as computational approaches using dedicated digital signal processing (DSP) chips or software. However, these techniques are inadequate for the task of automatic event recognition (AER). AER is defined as the task of identifying the occurrence of an event within a large video data base. Consider, for example, a data base that contains video surveillance gathered by a camera-equipped drone or a satellite, which is monitoring a site for suspicious activities, such as a truck of a particular size entering or exiting a facility. For data gathered over a few hours, this event may have occurred several times. The goal of an AER system is to determine if these events occurred, when they occurred, and how many times. In principle, this can be achieved by searching through each frame in the data base, and comparing them with reference images. This process is prohibitively time consuming, even with a very efficient optical image correlator, a software or a DSP based image recognition system. However, by employing the properties of atoms[20,21,22,23,24] it is possible to realize an AER that can recognize rapidly the occurrence of events, the number of events, and the occurrence times. Separate aspects of the overall technology needed for the AER have also been demonstrated by us earlier[20,21,25,26,27]. We propose the design of an AER system using Rb vapor in nano-porous glass with paraffin coating and show the simulation results of such a system. The AER will be realized by the technique of spatio-temporal holographic correlation[22,53,28,29,30]. This combines the process of translation invariant spatial holographic correlation with the process of translation invariant temporal correlation. The resulting system is a Spatio-Temporal Correlator (STC).

Shift invariant in space and time, the STC can recognize rapidly an event that may be present in a video file, and determine the temporal location of the event. In general, modeling the STC requires determining the temporal dynamics of a large number of inhomogeneously broadened atoms, multiplexed with free-space wave propagation equations. Here, along with modeling the STC using the Schrodinger equation for the temporal evolution of atoms excited by optical fields, we show that the response of the STC can be determined by modeling the response of the interaction medium as a simple, three-dimensional, multiplicative transfer function in the spatio-temporal Fourier domain. We explain the physical origin of this model, and then establish the validity of this model by comparing its prediction with that determined via the quantum mechanical dynamics. We then show some examples of the response of the STC using both methods. We also address practical issues in realizing such an STC. First, we show that optically off-resonant excitation of three-level atoms in the Λ configuration is suitable for realizing the STC. Furthermore, we show how a combination of a paraffin coated, nano-porous Rb vapor cell, a holographic video disc, and a lithium niobate crystal can be used to realize the STC for practical use.

The thesis is organized as follows: Chapter 2 reviews the concept of the translation invariant CHC. Chapter 3 presents the theoretical details of the proposed shift invariant HOC. The phase stabilization and scanning technique, which eliminates the convolution terms, is also described in this chapter. It also presents some possible future works on realizing an optoelectronic processor that would speed up the whole process of correlation. Chapter 4 describes simulation results of the proposed HOC (MATLAB has been used as a simulation software), illustrating its performance for detecting images in a shift invariant manner with various cases of interest. Chapter 5 presents a brief review of the underlying concepts of PMT for achieving scale and rotation invariance. Chapter 6 describes the proposed HOC architecture incorporating polar mellin transform (PMT) to achieve scale and rotation invariance in addition to shift invariance. Chapter 7 illustrates the simulation results using MATLAB, which shows that the proposed system is capable of detecting an image in a scale and rotation invariant manner. In this chapter, we also point out some limitations of using PMT in the CHC architecture as proposed previously, and show how to overcome these constraints. Chapter 8 describes the process of detecting multiple objects in a shift, scale and rotation invariant manner using the HOC architecture and show the simulation results. In Chapter 9 , we summarize the concept of a translation invariant temporal correlator and the basic model of an automatic target recognition system based on a translation-invariant STC. In Chapter 10, we describe the explicit architecture of an STC and introduce the analytic model for it. We also present results from numerical simulations of the stimulated photon echo, temporal correlator and automatic event recognition system, by using the quantum mechanical equations of evolution for an ensemble of inhomogeneously broadened atoms and by using the analytical model. In Chapter 11, we describe how a practical automatic target recognition system can be realized. The thesis concludes with a summary and outlook in chapter 12. In addition, some important details are presented in seven appendices. In Appendix A, we review the translation invarinace of the Fourier Transform in one and two dimensions. Appendix B describes the generation of two dimensional Fourier Transform using a lens. In Appendix C, we present the basic aspects of the holographic memory and how it works. In Appendix D, we review the relevant properties of the PMT. In Appendix E, we summarize our model for an inhomogeneously broadened two level system for implementing the STC. In Appendix F, we describe how a three level, Λ type system can be tailored to produce effectively the two level system needed for the STC. In Appendix G, we derive the details of the analytic expressions for one dimensional and three-dimensional spatio-temporal transfer functions that describe the functionality of the STC and enable very efficient simulation thereof.

Chapter 2 Conventional Holographic Correlator

2.1 The Basic Architecture of the Conventional Holographic Correlator

The basic optical correlator is illustrated schematically in Figure 2.1. Here, each lens has focal length of L. The process starts with an image (e.g., the reference image, possibly retrieved from a holographic memory disc or with a Spatial Light Modulator (SLM) connected to a computer) in the input plane: P_1. Let us denote this image by the potentially complex electric field distribution $U_{A1}(x_1, y_1)$, in normalized units. The time dependence is not included in this function; all fields in this discussion are assumed to be at the same frequency.

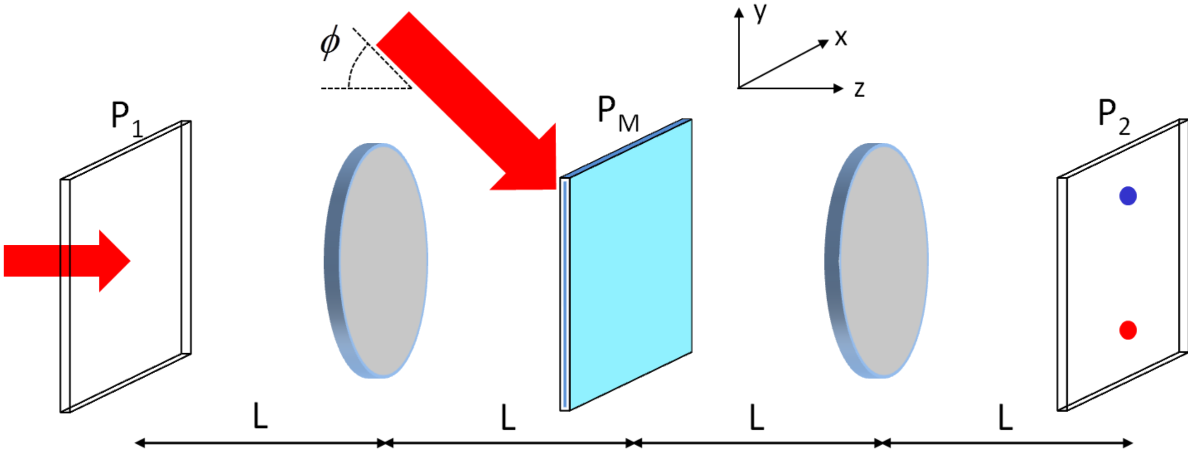


Figure 2.1: Schematic illustration of a conventional optical correlator

From the results shown in Appendix B, we can write the field in the middle plane, P_M, as

$$U_{AM}\left(x_M, y_M\right) = \frac{e^{j2kL}}{j\lambda L} \cdot \tilde{U}_{A1}\left(\frac{kx_M}{L}, \frac{ky_M}{L}\right) \tag{2.1}$$

where $\tilde{U}_{A1}$ is the Fourier Transform of U_{A1}. We now apply a plane wave, at an angle ϕ in the y-z plane, to interfere with the transformed image in the middle plane. This field can be written as:

$$U_W\left(x_M, y_M\right) = U_{W0} e^{-jk_\phi y_M}; \quad k_\phi = k\sin\phi \tag{2.2}$$

where k is the wave number of the laser field. The interference is recorded in a thin photographic plate, which produces a transmission function that is proportional to the interference pattern:

$$t\left(x_M, y_M\right) = \alpha I\left(x_M, y_M\right); \quad I\left(x_M, y_M\right) = \left|U_{A1}\left(x_M, y_M\right) + U_W\left(x_M, y_M\right)\right|^2 \tag{2.3}$$

For simplicity, we will assume that $\alpha = 1$, so that:

$$\begin{gathered} t\left(x_M, y_M\right) = t_1 + t_2 + t_3 + t_4 \\ t_1 = U_{W0}^2; \quad t_2 = \left|U_{AM}\right|^2; \quad t_3 = U_{W0} U_{AM}^* e^{-jk_\phi y_M}; \quad t_4 = t_3^* = U_{W0} U_{AM} e^{jk_\phi y_M} \end{gathered} \tag{2.4}$$

Once the recording is fixed, the second image (e.g., the query image, from a camera connected to an SLM) is presented in the input plane: P$_1$. Let us denote this image ,again, potentially complex by the electric field distribution $U_{B1}\left(x_1, y_1\right)$. In the middle plane, before the plate, the corresponding field is:

$$U_{BM,b}\left(x_M, y_M\right) = \frac{e^{j2kL}}{j\lambda L} \cdot \tilde{U}_{B1}\left(\frac{kx_M}{L}, \frac{ky_M}{L}\right) \tag{2.5}$$

where $\tilde{U}_{B1}$ is the Fourier Transform of U_{B1}. After passing through the photographic plate, the field can be written as:

$$U_{BM,a}\left(x_M, y_M\right) = tU_{BM,b}\left(x_M, y_M\right) = \left(t_1 + t_2 + t_3 + t_4\right)U_{BM,b}\left(x_M, y_M\right) \tag{2.6}$$

Only the term proportional to t_3 produces the cross-correlation signal. After passing to the second lens, the cross-correlation signal in the output plane, P_2, corresponding to this term becomes (as can be shown by carrying out the beam propagation analysis, using the Fourier Transforming property of the second lens):

$$S_{CC}\left(\vec{\rho}_2\right) = \beta \int d\vec{\rho}\,' U_{B1}\left(\vec{\rho}\,'\right) U_{A1}\left(\vec{\rho}\,' + \vec{\rho}_W - \vec{\rho}_2\right) \tag{2.7}$$

Here, we have used the compact notation of two-dimensional integration introduced in Appendix A (see eqn. A.5). Specifically, we have defined:

$$\vec{\rho}_2 \equiv \left(\hat{x}x_2 + \hat{y}y_2\right) \tag{2.8}$$

as the coordinate vector for the output plane. The integration is carried out over a dummy coordinate vector defined as:

$$\vec{\rho}\,' \equiv \left(\hat{x}x' + \hat{y}y'\right) \tag{2.9}$$

The factor in front, β, is complex, but does not depend on the coordinates in the output plane, and corresponds to a constant factor in the detected intensity corresponding to the field shown in eqn.(**2**.**7**) . Finally, the offset vector is eqn. (**2**.**7**) is related to the propagation vector of the writing beam shown in eqn. (**2**.**7**), and is given by:

$$\vec{\rho}_W \equiv \hat{y} L \sin\phi \tag{2.10}$$

Thus, the cross-correlation signal appears as an intensity localized at:

$$x_2 = 0, y_2 = -L\sin\phi \tag{2.11}$$

as indicated by the red dot in Figure 2.1. The amplitude of this cross-correlation signal is high when the reference image and the query images are matched. If the query image is shifted with respect to the input image in the x-y plane, the correlation spot will also appear shifted, without a loss in signal strength.

The other terms in eqn. (**2**.**6**) also produce signals (for example, the term proportional to t_4 gives the convolution between the two images, as shown by the blue dot in Figure 2.1), but they do not overlap with the cross-correlation signal if ϕ is chosen to be sufficiently large.

2.2 Limitations of the Conventional Holographic Correlator (CHC)

The above described correlator is so called Translation Invariant- Spatial Holographic Correlator (TI-SHC). It is also possible to realize the TI-SHC as a single-step process, where the write beam, the reference beam and the query beams are applied simultaneously to a dynamic photo-refractive medium, thus realizing what is known as the joint-transform correlator[1]. Using images pre-processed via so-called polar Mellin transforms, it is also possible to make the system also invariant with respect to scale and rotation[12,13 ,31].

Over recent years, our group have been investigating the feasibility of realizing a high-speed automatic target recognition system using the TI-SHC technique[,1,2,3,4,5,6,7]. One approach pursued by us employed a high-capacity holographic video disc (HVD) with a terabyte storage capacity, corresponding to 10 million images. An image from this disc can be retrieved very rapidly and applied to a joint-transform correlator (JTC), without requiring any electronic processing. The query image, captured by a camera, provides the second input to the same JTC, along with a reference beam. The JTC we used is constructed from an organic photorefractive film, developed jointly by Nitto-Denko, Inc., and University of Arizona. The process we employed is illustrated schematically Figure 2.2. Briefly, as shown in the top-left, the query image is acquired by a suitable input device, such as a video camera. The image is then

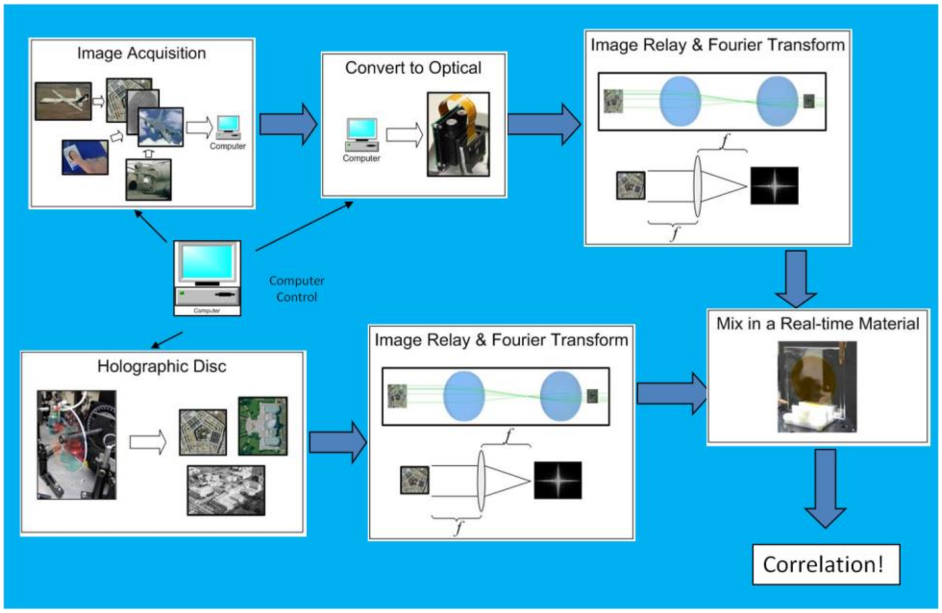


Figure 2.2: Schematic illustration of the process employed by the holographic correlator shown in Figure 2.3.

converted to an optical one by reflecting a laser beam off a Spatial Light Modulator (SLM), as shown in the top-middle. The output of the SLM is sent through a lens, which produced a spatial Fourier Transform (FT), as discussed above. The reference images are pre-recorded and stored in the holographic video disc (HVD). Due to a high degree of Bragg selectivity for a thick medium, many images can be stored at the same physical location, using angular multiplexing. For example, a typical disc that is 1 cm thick and has a diameter of about 15 cm, which is compartmented into nearly 2000 locations, can store up to 10 million images, each with a 512X512 pixels. During the correlation process, images from this disc are retrieved sequentially, using a laser beam transmitted through the disc, as shown in bottom-left. This retrieval also occurs very fast, corresponding to the time it takes for light to pass through the disc. Furthermore, by using high-speed acousto-optic modulators, the image from any location and at any angle can be accessed with a very short latency time. The low latency, high speed and high-capacity thus make such a memory ideal for target recognition. The reference image, retrieved from the disc, is passed through a lens, producing its FT, as shown in bottom-middle of Figure 2.2.

The FT's of both the query image and the reference image are then applied to a photorefractive polymer film, configured as a joint-transform correlator (JTC), as shown in bottom-right of Figure 2.2. A JTC can be realized in many different ways. In our system, we used a four-wave mixing geometry to realize a JTC. Briefly, three different optical beams are

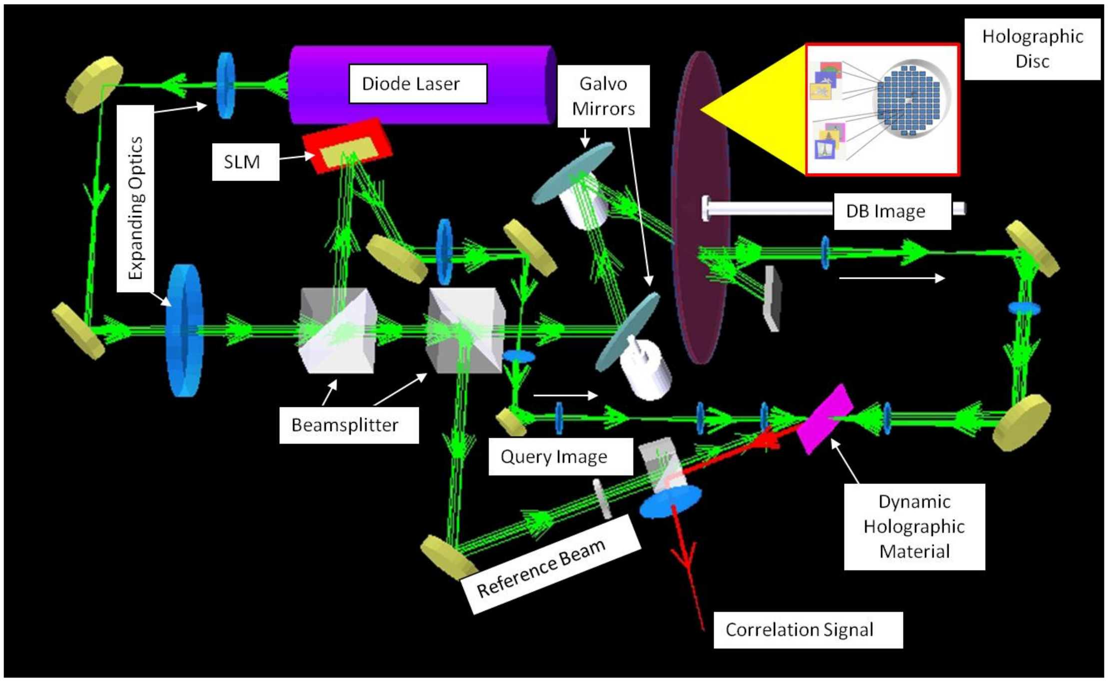


Figure 2.3: A holographic target recognition system demonstrated previously in our lab. See text for details.

grating in the film. The query image then diffracts off this interference grating. This process happens simultaneously, through four-wave mixing produced by the film[1,2]. In case of a match, a bright signal is produced, at an angle that depends on the relative translation between the query image and the reference image. If there is no match, the process is repeated with another reference image retrieved from the memory disc. Note that this is simply a variation of the two-step conventional optical correlator described in section 2.1. The only difference here is that the two steps (recording and correlation) occur simultaneously.

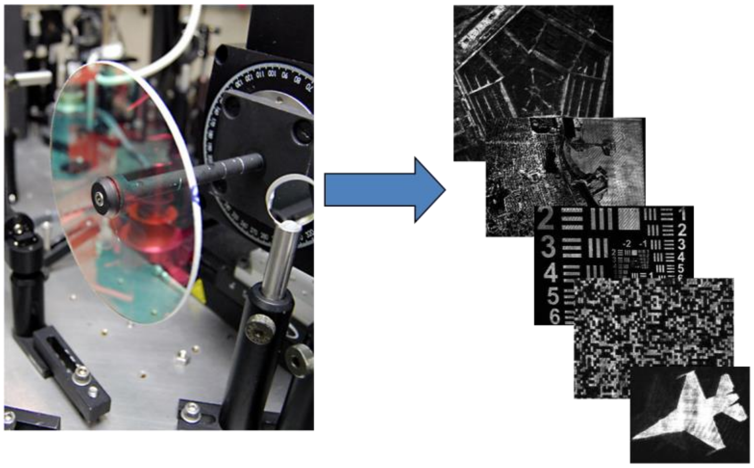


Figure 2.4:Samples of images retrieved from the holographic video disc (HVD).

Such a system was constructed and demonstrated[1,2] in our lab. The details of this system are shown in Figure 2.3. For this implementation, we used a holographic video disc made from a substrate produced by Aprilis, Inc. The photorefractive polymer film was obtained from Nitto-Denko, Inc. A 512X512 pixels SLM made by Displaytech, Inc. was used to convert camera images into optical ones. Various spatial locations on the memory disc were accessed by an automated rotation and translation stage. The angular multiplexing at a given location was achieved by using a pair of galvo-mounted mirrors. A typical set of images stored and retrieved from the holographic memory disc is shown in Figure 2.4. Figure 2.5 shows a set of correlation peaks produced by tracking the image of a moving plane.

The primary limitation of this system was the poor nature of the material used for making the JTC. First, it is very fragile, and gets destroyed rather easily for reasons not well understood. Second, the diffraction efficiency is rather small, and it produces a lot of scattering, leading to a very poor signal-to-noise (SNR) ratio. Third, after carrying out some correlations, the residual gratings generated in the medium have to be erased by applying a high voltage; this process takes quite a bit of time. To our knowledge, all dynamic holographic films suffer from similar limitations.

Figure 2.5: Example of a set of correlation seen by the system in Figure 2.3.

Our proposed Hybrid Optoelectronic Correlator (HOC) gets around this problem by making use of concepts developed recently in the context of digital holography[32,33]. Namely, the non-linearity provided by the JTC medium is replaced by the non-linearity of high-speed detectors (since detectors measure intensities, they are inherently non-linear). Of course, this requires some modification of the architecture, as well as post-processing of signals. In The next chapter, we specify an explicit and novel architecture that enables the process of correlating images using detectors only.

Chapter 3 Shift Invariant Hybrid Optoelectronic Correlator

3.1 Proposed Architecture of Hybrid Optoelectronic Correlator

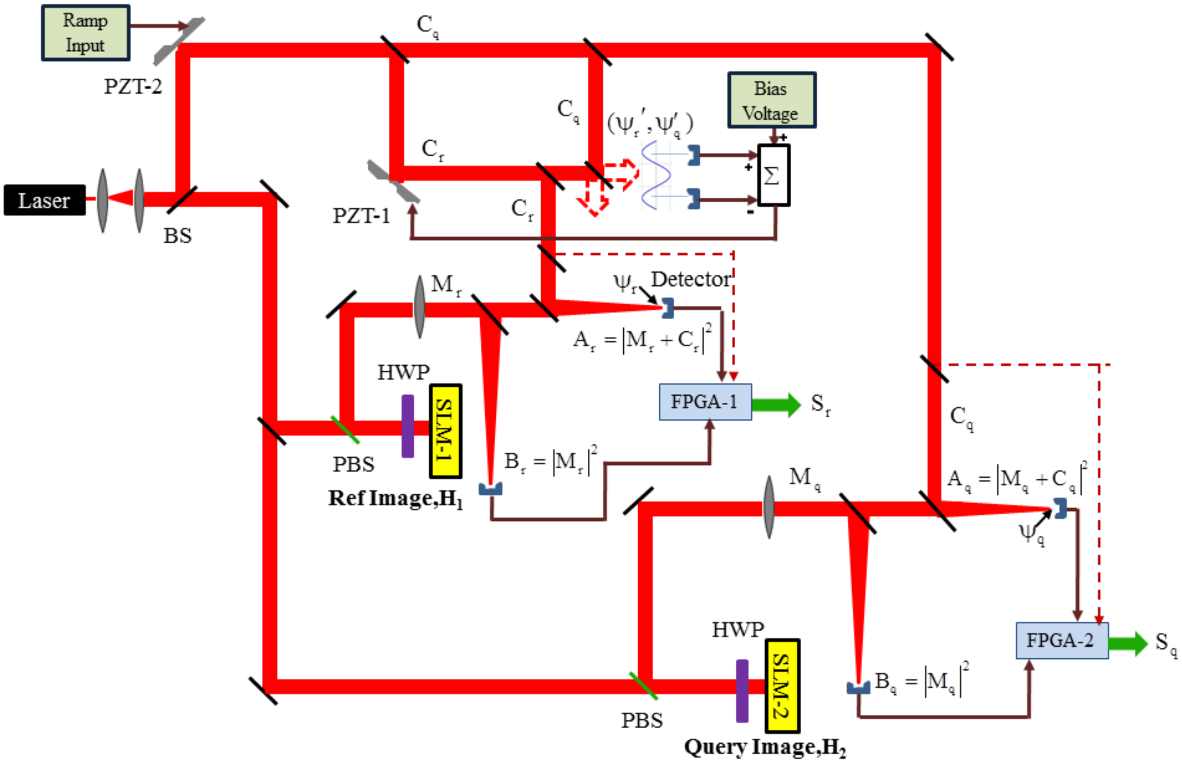


Figure 3.1: Proposed architecture of the hybrid optoelectronic correlator with phase stabilization and scanning circuit; BS = beam splitter; PBS = polarizing beam splitter; HWP = half wave plate; PZT = Piezo-electric Transducer]

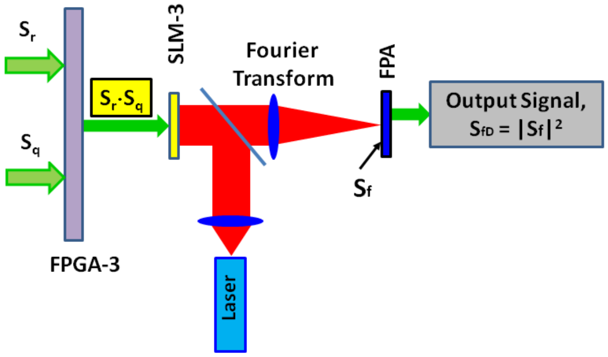


Figure 3.2: The final stage of the Hybrid Optoelectronic Correlator

The proposed architecture of the Hybrid Optoelectronic Correlator (HOC)[34] is summarized in Figure 3.1 and Figure 3.2. Briefly, the reference image, H_1 is retrieved from the database and transferred to an optical beam using a Spatial Light Modulator (SLM-1), and is Fourier transformed using a lens. The Fourier transformed image (M_r) is split into two identical ports. In one port, the image is detected by an array of detectors, which could be a high resolution focal plane array (FPA) or a digital CMOS camera. As an explicit example, we consider the USB2.0 CMOS Camera (DCC1545M), which has 1280H x 1024V pixels and sends 10 bit data of each pixel at a 48 MHZ clock rate, thus requiring about 27 ms to send an image. The signal array produced by the camera is denoted as B_r. The camera is interfaced with a Field Programmable Gate Array (FPGA) via a USB cable. B_r can be stored in the built in memory of the FPGA [FPGA-1]. In the other port, M_r is interfered with a plane wave C_r, and detected with another CMOS Camera, producing the digital signal array A_r and is stored in the memory of FPGA-1. A_r & B_r can be expressed as:

$$A_r = \left|M_r e^{j\phi_r} + C_r e^{j\psi_r}\right|^2 = \left|M_r\right|^2 + \left|C_r\right|^2 + \left|M_r\right|\left|C_r\right| e^{j(\phi_r-\psi_r)} + \left|M_r\right|\left|C_r\right| e^{-j(\phi_r-\psi_r)} \quad (3.1)$$

$$B_r = \left|M_r\right|^2 \quad (3.2)$$

In addition, the intensity profile of the plane wave ($|C_r|^2$) is measured, by blocking the image path momentarily, using a shutter (not shown), and the information is stored in the memory component of FPGA-1. FPGA-1 then computes and stores S_r which can be expressed as:

$$S_r = A_r - B_r - \left|C_r\right|^2 = M_r C_r^* + M_r^* C_r = \left|M_r\right|\left|C_r\right| e^{j(\phi_r-\psi_r)} + \left|M_r\right|\left|C_r\right| e^{-j(\phi_r-\psi_r)} \quad (3.3)$$

Here $\phi_r(x,y)$ is the phase of the Fourier transformed image, M_r, and Ψ_r is the phase of the plane wave, C. It should be noted that ϕ_r is a function of (x,y), assuming that the image is in the (x,y) plane. This subtraction process has to be done pixel by pixel using one or more subtractors available in the FPGA. Consider, for example, the Virtex-6 ML605 made by Xilinx as a candidate FPGA, which has an oscillator frequency of 200 MHz, so that each subtraction takes about 5 - 10 ns depending on the implementation of the adder circuit**.** Thus, the total subtraction process of an image of size 1280 x 1024 pixels would take about 6 ms when the subtraction is done with one subtractor. This process can be speeded up by using multiple subtractors that can operate in parallel in many FPGAs.

The captured query image, H_2 is transferred to an optical beam using another SLM (SLM-2), and split into two paths after being Fourier transformed with a lens. The resulting image in each path is designated as M_q. In a manner similar to what is described above for the query image, the signal $S_q = A_q - B_q - |C_q|^2$ is produced using two cameras and an FPGA (FPGA-2) and stored in FPGA-2 memory. Here, C_q is the amplitude of an interfering plane wave, and the other quantities are given as follows:

$$A_q = \left|M_q + C_q\right|^2 = \left|M_q\right|^2 + \left|C_q\right|^2 + \left|M_q\right|\left|C_q\right|e^{j(\phi_q-\psi_q)} + \left|M_q\right|\left|C_q\right|e^{-j(\phi_q-\psi_q)} \tag{3.4}$$

$$B_q = \left|M_q\right|^2 \tag{3.5}$$

$$S_q = M_q C_q^{\,*} + M_q^* C_q = \left|M_q\right|\left|C_q\right|e^{j(\phi_q-\psi_q)} + \left|M_q\right|\left|C_q\right|e^{-j(\phi_q-\psi_q)} \tag{3.6}$$

As before, $\phi_q(x,y)$ is the phase of the Fourier transformed image, M_q, and Ψ_q is the phase of the plan wave D.

In the final stage of the hybrid correlator (as shown in Figure 3.2), these two signals (S_r and S_q) described in equations (**3**.**3**) and (**3**.**6**) are multiplied together using the multiplier in FPGA-3. Four quadrant multiplication can easily be implemented using an FPGA. The resulting signal array, S is stored in FPGA-3 memory. This can be expressed as:

$$\begin{aligned} S = S_r \bullet S_q &= \left(M_r C_r^{\,*} + M_r^* C_r\right)\left(M_q C_q^{\,*} + M_q^* C_q\right) \\ &= \left[\alpha^* M_r M_q + \alpha M_r^* M_q^* + \beta^* M_r M_q^* + \beta M_r^* M_q\right] \end{aligned} \tag{3.7}$$

where, $\alpha \equiv \left|C_r\right|\left|C_q\right|e^{j(\psi_r+\psi_q)}; \quad \beta \equiv \left|C_r\right|\left|C_q\right|e^{j(\psi_r-\psi_q)}$.

This signal array, S, is now transferred to another SLM (SLM-3) from FPGA-3 through the digital visual interface (DVI) port. For the X-Y series SLM made by Boulder Nonlinear Systems (BNS), for example, this image update would take about 65 ms. Since S can be positive or negative, the SLM should be operated in a bipolar amplitude mode. The optical image produced by SLM-3 is Fourier transformed using a lens, and detected by an FPA. The output of the FPA will provide the main correlation signal. The final signal can thus be expressed as:

$$S_f = \left[\alpha^* F(M_r M_q) + \alpha F(M_r^* M_q^*) + \beta^* F(M_r M_q^*) + \beta F(M_r^* M_q)\right] \tag{3.8}$$

Here, F stands for the Fourier Transform and α, β are constants. Since M_r and M_q are the FT of the real images H_j, j=1,2 respectively, using the well-known relations between the FT of products of functions, and convolutions and cross-correlations, we can express the final signal as the sum of four terms:

$$S_f = \alpha^* T_1 + \alpha T_2 + \beta^* T_3 + \beta T_4 \tag{3.9}$$

$$\begin{aligned} T_1 &= H_1(x,y) \otimes H_2(x,y) \\ T_2 &= H_1(-x,-y) \otimes H_2(-x,-y) \\ T_3 &= H_1(x,y) \odot H_2(x,y) \\ T_4 &= H_2(x,y) \odot H_1(x,y) \end{aligned}$$

Where $\otimes$ indicates two-dimensional convolution, and $\odot$ indicates two-dimensional cross-correlation. We can now make the following observations:

- T_1 represents the two-dimensional convolution of the images, H_1 and H_2.
- T_2 represents the two-dimensional convolution of the images, H_1 and H_2, but with each conjugated and inverted along both axes.
- T_3 represents the two-dimensional cross-correlation of the images, H_1 and H_2.
- T_4 represents the two-dimensional cross-correlation of the images, H_2 and H_1. (Cross-correlation is non-commutative; hence, T_3 is not necessarily equal to T_4)

- If the images, H_1 and H_2, are not symmetric in both x and y directions, we have $T_1 \neq T_2 \neq T_3 \neq T_4$
- If the images, H_1 and H_2, are symmetric in both x and y directions, we have $T_1 = T_2 = T_3 = T_4$

The cross-correlation technique is usually used to find matches between two objects. In our final result we have convolution terms (T_1 & T_2) in addition to cross-correlation terms (T_3 & T_4). The convolution terms can be washed out by implementing a phase stabilization and scanning technique in the HOC architecture.

3.2 Phase Stabilization & Scanning Circuit

From eqn (**3.9**), it is obvious that the final signal S_f depends nontrivially on Ψ_1 & Ψ_2. To make this dependence more transparent, we can rewrite eqn (**3.7**) as follows:

$$S = 2|C_r||C_q||M_r||M_q|\left[\cos(\psi_r + \psi_q - \phi_M) + \cos(\psi_r - \psi_q - \tilde{\phi}_M)\right] \tag{3.10}$$

where, *and* $M_r M_q^* = |M_r||M_q|e^{j\tilde{\phi}_M(x,y)}$. Here the first term corresponds to the convolution and the second term corresponds to the cross-correlation. To eliminate the convolution term, we continuously scan ($\Psi_r + \Psi_q$) over a range of 2π at a certain frequency ω_s, while keeping (Ψ_r-Ψ_q) zero. The convolution term varies as we scan ($\Psi_r + \Psi_q$), whereas the cross-correlation term remains constant (since $\Psi_r - \Psi_q = 0$). While scanning is going on, we pass the signal S through a low-pass filter (LPF) with a bandwidth less than ω_s, so that the last term which corresponds to cross-correlation is passed. Such a filter can be easily implemented with the FPGA. The low-pass filtered version of S is Fourier transformed using a lens and detected by an FPA, producing only the cross-correlation signals.

Figure 3.1 shows the architecture for phase stabilization and scanning, where $(\Psi_r - \Psi_q)$ is kept to a constant value using a simple interferometer along with a feedback loop and $(\Psi_r + \Psi_q)$ is scanned over a range of 2π using a Piezoelectric Transducer (PZT). To start with, the output of the laser is split into two paths. One path is used for generating images via reflections from SLM (as already shown in Figure 3.1). The other path is used for generating both reference beams, C_r & C_q. Using a PZT (PZT-1) mounted on a mirror before the beam is split into C_r & C_q, and applying a saw-tooth type ramp voltage on it, result in a repeated linear scan of $(\Psi_r + \Psi_q)$ over a range of 2π. Pieces of C_r & C_q are now split off and interfered with each other. The resulting interference pattern along with a feedback signal applied to another PZT (PZT-2) mounted on a mirror in the path of C_r, can be used to control the value of $(\Psi_r - \Psi_q)$. To be specific, note that, $\Psi_r - \Psi_q = \xi + (\Psi_r' - \Psi_q')$ where, Ψ_r & Ψ_q are the phases at the detectors D_1 and D_2, respectively, Ψ_r' & Ψ_q' are the phases of C and D at the detectors of the interferometer (D_3 & D_4) and ξ is a constant, determined by path length difference. This condition of $(\Psi_r - \Psi_q) = 0$ can be achieved if $(\Psi_r' - \Psi_q') = -\xi$. We describe below how this can be achieved.

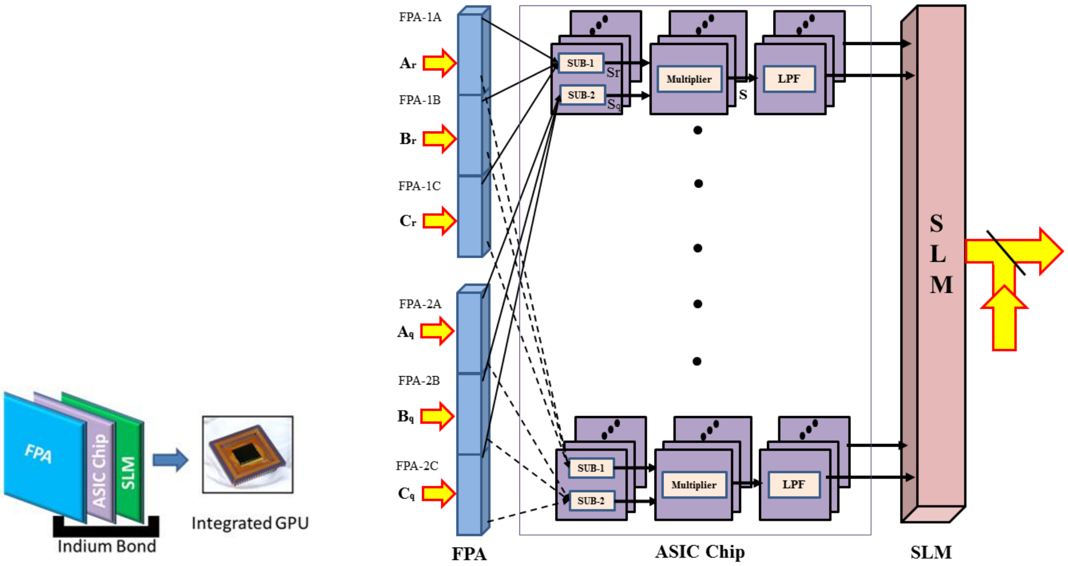
The interference signal is detected by a pair of matched detectors (D_3 & D_4). The voltages from these detectors are subtracted from each other using a subtractor circuit. The resultant voltage from the subtractor is added with a bias voltage using an adder circuit and the output of this adder is fed to PZT-2. The feedback loop can lock the interference pattern at any desired position. By performing several correlation operations of two identical images with this setup at different bias voltages, we find the bias voltage that gives the maximum peak value of the cross-correlation signal. At this position, $(\Psi_r' - \Psi_q') = -\xi$, which makes $(\Psi_r - \Psi_q) = 0$.

With the system locked at this position, $(\Psi_r + \Psi_q)$ is varied over a range of 2π by applying a ramp voltage to PZT-1, as mentioned above. The response of the feedback loop should be faster than ω_s, in order to ensure that the servo can hold $(\Psi_r' - \Psi_q')$ to a constant value. While the scanning is going on, the signal is passed through an LPF with a bandwidth less than ω_s. This low pass filtered version of S is then processed to yield cross-correlation signals only, as discussed in section 3.1.

While the image detection process is going on, the stability of the phases should be checked after some characteristic time, T_c. This characteristic time is defined as the time during which $(\Psi_r - \Psi_q)$ can drift within a certain allowable range, for example, a few milliradians. After time T_c, we have to adjust the bias voltage again and perform several correlation of two known images with the HOC to get the highest correlation peaks. The characteristic stability time would depend on the stability of the optical mounts, and can easily exceed 100's of seconds in a well designed system.

3.3 Possible Future Work to Speed up the Operation of the HOC

In describing the architecture of the HOC in section 3.1, we have considered the use of commercially available components such as cameras, FPGAs and SLMs. However, it is obvious from the analysis that the overall process is severely slowed down during the serial communication between these devices. In order for the HOC to achieve its ultimate operating speed, it is thus necessary to resort to novel components that operate in parallel.



(a) (b)

Figure 3.3: (a) Shows the block diagram of the integrated graphic processing unit (IGPU) (b) Shows the block diagram of the specialized IGPU in detail that combines, in parallel, the FPA, the ASIC (Application-specific Integrated Circuit) chip for signal processing, and the high-speed SLM. [SUB = Subtractor; FPA= Focal Plane Array; LPF = Low Pass Filter; GPU= Graphic processing Unit, SLM=Spatial Light Modulator]

Consider the set of steps (shown in Figure 3.2) whereby signals captured by the cameras are processed to produce the signal S, which then appears as an optical field at the output of the final

SLM (SLM-3). This process is serving as a conduit between signals that start in the optical domain (i.e. inputs for the cameras) and end in the optical domain (i.e. output of the SLM-3). With proper use of current technologies, it should be possible to combine these tasks in an integrated graphic processing system for high-speed operation, as shown in Figure 3.3(a). Figure 3.3(b) shows the block diagram of the specialized integrated graphic processing unit (IGPU) that combines, in parallel, the FPA, the ASIC (Application-specific Integrated Circuit) chip for signal processing, and the high-speed SLM.

Briefly, FPA-1A, FPA-1B and FPA-1C would capture, respectively, signals, $A_r(\equiv|M_r + C_r|^2), B_r(\equiv|M_r|^2)$ and $|C_r|^2$, corresponding to the reference channel[35]. Similarly, FPA-2A, FPA-2B and FPA-2C would capture, respectively, signals, $A_q(\equiv|M_q + C_q|^2), B_q(\equiv|M_q|^2)$ and $|C_q|^2$, corresponding to the query channel. The FPA would be connected to the ASIC chip through Indium bonding, for example. The ASIC chip would consist of 512 X 512 signal-processing units, corresponding to each pixel and geometrically matched to the FPA. Each unit of the ASIC chip would have two analog subtractor circuits, one analog multiplier circuit and an analog LPF. The signals from the matching pixels in the three FPAs (FPA-1A, FPA-1B and FPA-1C) will be processed by a single subtractor element (denoted as SUB-1) in the ASIC chip, producing the signal $S_r = A_r - B_r - |C_r|^2$. Similarly, FPA-2A, FPA-2B and FPA-2C, along with SUB-2, would generate the signal array S_q, corresponding to the query channel. These two signals, S_r and S_q, would be applied in parallel, to an analog multiplier, producing the signal array S. The signal S is passed through an LPF to get rid of the convolution terms, as discussed in section 3.1. On the back end of the ASIC chip array would be the SLM array,connected via indium bonding.

The speed of operation of such an IGPU would, of course, depend on the specific technology employed. In what follows, we describe a specific example of the technology that could be employed to realize each part of the IGPU, thus enabling us to reach a definitive estimate of the speed of operation.

Consider first the FPA. One possible choice for this would be an array of nano-injection detectors [36,37]. These detector elements operate at a low bias voltage (~ 1 V) at room temperature, and has a response time of 1 ns. Consider next the subtractor elements. Each of these could be implemented easily with a simple operational amplifier consisting of as few as six transistors[38]. The response time of such an operational amplifier is expected to be similar (~0.3 us) to that of a bulk operational amplifier chip, such as LM741C. The multiplier could be realized with the Gilbert cell, which requires only six transistors[36]. The properties of such a multiplier should be similar to that of a bulk multiplier chip, such as AD835 which is a complete four-quadrant, voltage output analog multiplier. It generates the linear product of its X and Y voltage inputs with a rise time of 2.5 ns. The LPF can be implemented with 0.1 nf capacitance and 10 ohm resistance would have a response time of 1ns. Finally, the SLM could be realized with an array of high speed stepped quantum wells[39,40,41,42,43]. An SLM based on these elements has a response time of ~16 ns, which is orders of magnitude faster than the conventional SLM discussed in section 3.1.

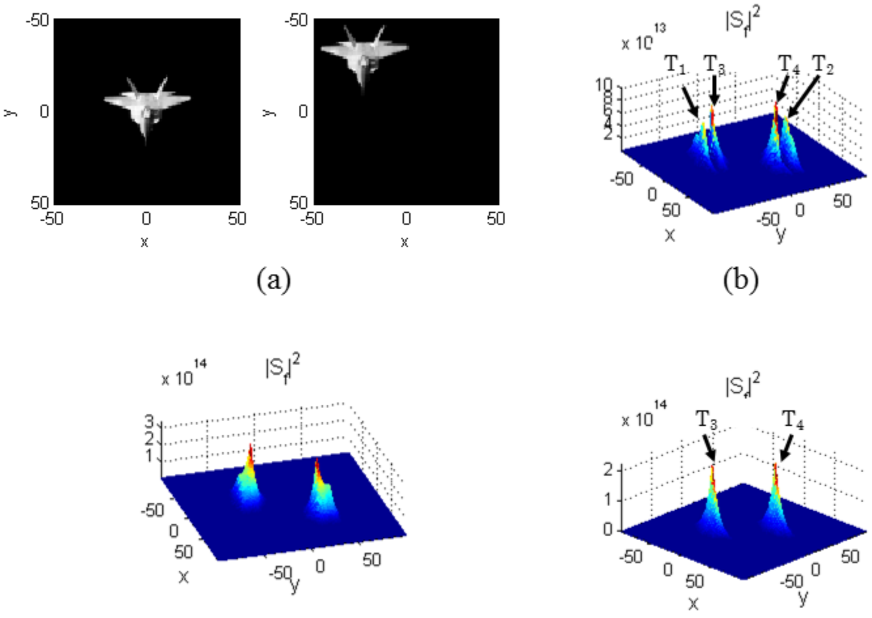
Therefore, the integrated graphic processing unit would take less than 0.4 μs to perform the whole process of capturing optical signals, computing the signal S and converting it back to optical domain. In contrast, for performing the same operation of 512X512 image, the technology discussed in section 3.1 takes about 22 ms since the data communication is done serially. Thus, the IGPU would lead to a speed-up of operation by a factor of nearly $5*10^4$. Of course, we have simply outlined a sketch of the type of the integrated graphic processing unit that is required to make the HOC achieve its ultimate speed. We are working with collaborators to design and implement this chip, and results from these efforts will be reported in the future.

3.4 Conclusion

We have presented theoretical details and the underlying architecture of the HOC that correlates images using SLMs, detector arrays and FPGAs. The HOC architecture bypasses the need for photorefractive polymer films by using detectors, and the phase information is yet conserved by the interference of plane waves with the images. The output signal of such an HOC has four terms: two convolution signals and two cross-correlation signals. The convolution terms can be eliminated by implementing a phase stabilization and scanning circuit, so that the behavior of an HOC becomes essentially identical to that of a conventional Holographic Correlator (CHC). To achieve the ultimate speed of such a correlator, we also propose an opto-electronic chip which would perform all the electrical processes in a parallel manner. The HOC architecture along with the phase stabilization technique would thus be as good as a CHC, capable of high speed image recognition in a shift invariant manner. In addition to shift invariant property of the HOC, rotation and scale invariant correlation can also be achieved by applying Polar Mellin Transform (PMT) to both query and reference images[31]. With the future implementation of an opto-electronic chip and PMT, the HOC architecture holds the promise of a practical, versatile and high speed image recognition system.

Chapter 4 Simulation and Experimental Results of Shift Invarinat Hybrid Optoelectronic Correlator (HOC)

4.1 Numerical Simulation Results



(a) (b)

(c) (d)

Figure 4.1: (a) Two identical but shifted images are inputs to the HOC. (b) All four terms are present in the output of the detector when $\Psi_r = \Psi_q = 0$ *(without phase stabilization circuit). (c) The output of the detector,* $|S_f|^2$ *when* $\Psi_r = \Psi_q = \pi/3$ *(This result is also without phase stabilization circuit) (d) Phase stabilization and scanning gives the optimum value of* $|S_f|^2$ *by eliminating the convolution terms (*T_1 *&* T_2*) and keeping only the cross-correlation terms (*T_3 *&* T_4*).*

As discussed in section 3.2, the convolution terms can be washed out by scanning continuously ($\Psi_r + \Psi_q$) over a range of 2π at a certain frequency ω_s, while keeping ($\Psi_r - \Psi_q$) zero. This can be verified through some simulation results shown in Figure 4.1. Two identical but shifted images (as shown in Figure 4.1(a)) are inputs to the HOC architecture without phase stabilization circuits. When phase stabilization and scanning technique is not implemented and Ψ_r & Ψ_q are both kept zero, FPA detects a signal $|S_f|^2$ which contains both convolution (T_1 & T_2) and cross-correlation (T_3 & T_4) terms, as shown in Figure 4.1(b). Another case without phase stabilization and scanning technique is shown in Figure 4.1(c) where Ψ_r & Ψ_q are both set to the value of $\pi/3$. In this case also, the convolution terms appear in $|S_f|^2$ along with the cross-correlation terms.

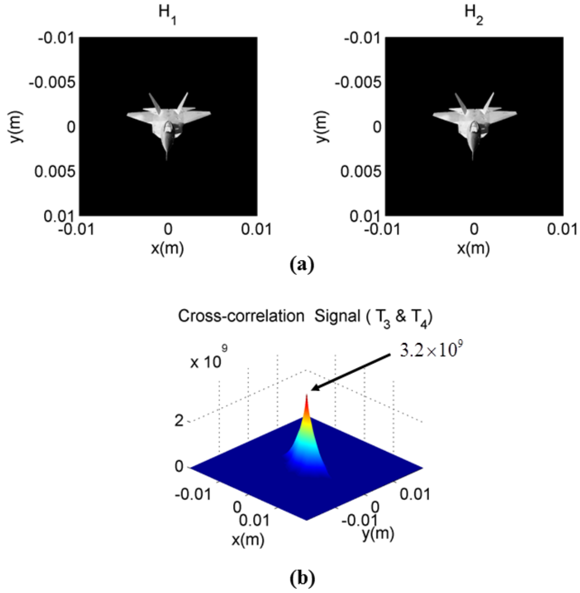


Figure 4.2: (a) Two similar images with same position, H_1 & H_2 are the inputs to the HOC (b) The cross-correlation signals (T_3 & T_4) give sharper peak due to matched images

Next we vary $(\Psi_r + \Psi_q)$ from 0 to 2π, in 20 intervals, and average the values of S for these 20 cases. This is, of course, equivalent to the process of low pass filtering with a bandwidth less than the inverse of the duration for a linear scan of $(\Psi_r + \Psi_q)$ over a range of 2π, as described in the section 3.2. As shown in Figure 4.1(d), this averaging eliminates the convolution terms, and $|S_f|^2$ contains only the cross-correlation terms. In what follows, we will assume that such an averaging/filtering is carried out, and plot only the cross-correlation terms.

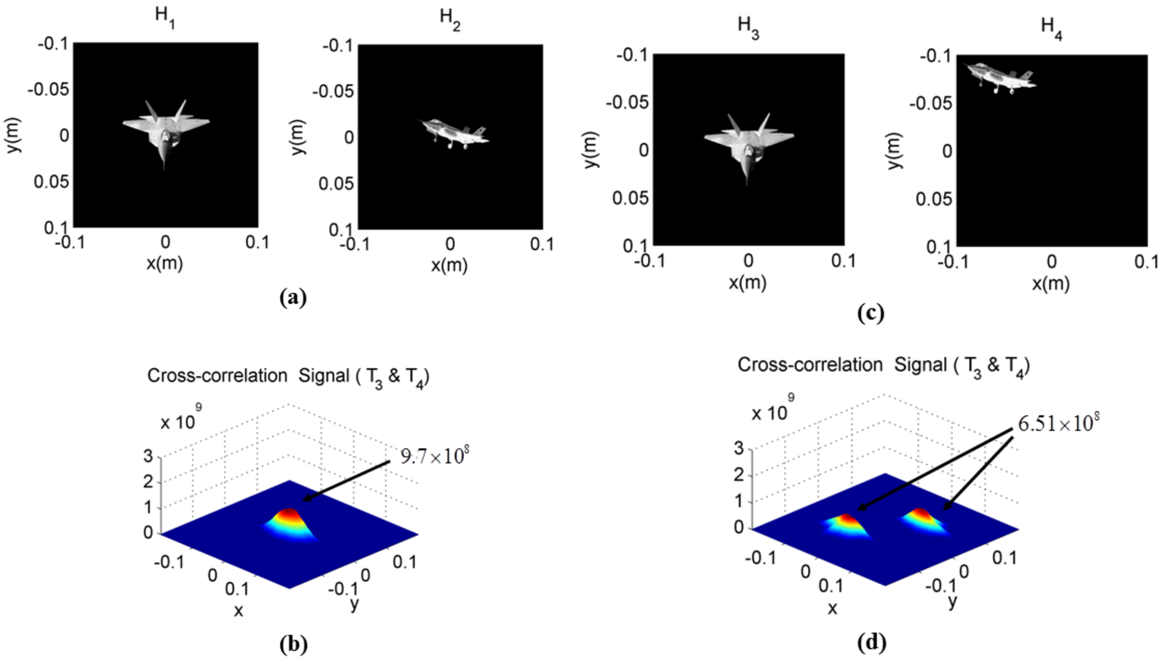


Figure 4.3: (a) Two completely different images H_1' and H_2' are inputs to the HOC (b) The cross-correlation terms, T_3 & T_4 do not reveal any sharper peak, since they are different. (c) Two different and shifted images, H_3' and H_4' are inputs to the HOC. (d) The cross-correlation terms also (T_3 & T_4) do not reveal sharper peak.

Next, we illustrate the behavior of the HOC architecture for various types of objects and reference images. In Figure 4.1(d) we have already shown the result for identical but shifted images. Figure 4.2(a) shows the case where the reference image, H_1 and the query image, H_2 are identical and not shifted. The cross correlation signals (T_3 & T_4) are shown in Figure 4.2(b). In this case, the cross correlation signals (T_3 & T_4) have a peak at the center of the detector plane[44]. Figure 4.3(a) shows another case, where two different images, H_1 and H_2, are the inputs to the HOC. In this case, as seen in Figure 4.3(b), the cross-correlation signals (T_3 & T_4) do not reveal any distinct peak. Figure 4.3(c) shows the case of two different images, H_3 and H_4, shifted with respect to each other. In this case, the cross-correlation terms (Figure 4.3(d)) also do not have sharp peaks, and the maximum value is less than that shown in Figure 4.3(b). Of course, it is also much smaller than the matched cases shown in Figure 4.1(d) and Figure 4.2(b).

It is important to note that if the query and reference images are not shifted with respect to one also another, then the cross-correlation peaks will be at the center of the detection plane, overlapping each other. If the images are shifted with respect to one another, the cross-correlations will be shifted symmetrically around the center by a distance corresponding to the shift.

4.2 Resolving Power of the HOC

The cross-correlation technique is usually used to find matches between two objects. In our architecture, we have convolution terms (T_1 and T_2) in addition to cross-correlation terms (T_3

and T_4). The convolution terms can be washed out by implementing a phase stabilization and scanning technique in the HOC architecture which has been discussed in details in section 3.2.

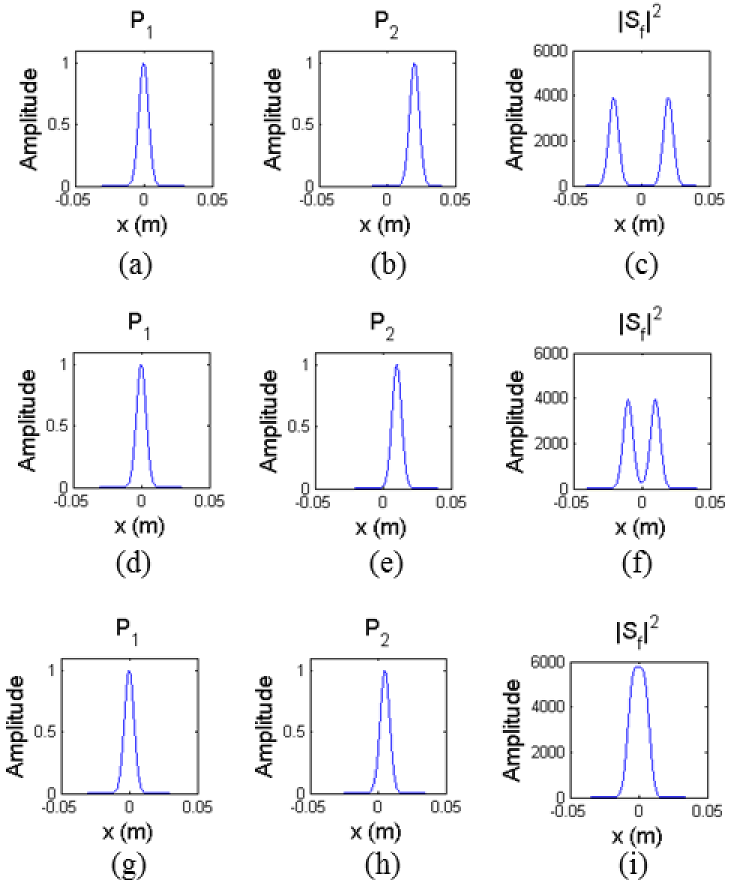


Figure 4.4: Illustration of the resolving power of the HOC architecture for one dimensional identical images

The conventional holographic correlator (CHC)[1,2,8] has one convolution term and one cross-correlation term, along with some dc and low spatial frequency outputs. In contrast, the HOC architecture, after incorporating the phase stabilization and scanning, has only two outputs, a cross correlation (T_3) and an anti cross-correlation term (T_4). The strengths of T_3 and T_4 are the same, and they appear as symmetric signals in the presence of a shift. This difference has to be kept in mind in using the HOC architecture.

The signal observed by the final FPA, is, of course, given by $S_{fD} \equiv \alpha\,|S_f|^2$, where α is a proportionality constant, which we set to be unity for simplicity of discussion. Thus, assuming that the contributions from T_1 and T_2 are eliminated via the combination of phase scanning and low pass filtering[44], the final FPA signal can be expressed as:

$$S_{fD} = |\beta^* T_3 + \beta T_4|^2 = \beta^* FT(M_r M_q^*) + \beta FT(M_r^* M_q) \tag{4.1}$$

It is useful to consider two different scenarios in order to interpret the information one can glean from this signal.

Scenario 1: Perfectly matched images with no relative shift:

In this case, $T_3 = T_4$, so that eqn (**10.6**) can be expressed as: $S_{fD} = |\beta^* + \beta|^2 |T_3|^2 = 4|\beta|^2 \cos^2(\Psi_r - \Psi_q)|T_3|^2$. This signal is maximum when $(\Psi_r - \Psi_q) = 0$. As discussed in ref. 8, this situation (i.e. use of identical, unshifted images) thus can be used to keep the servo locked to the position where $(\Psi_r - \Psi_q) = 0$. Thus, we will assume from now on that $(\Psi_r - \Psi_q) = 0$, so that $\beta = \beta^* = |\beta|$.

Scenario 2: Perfectly matched images with a relative shift in position:

If the query image is shifted by a vector, $\overrightarrow{\rho_0}$, then, $M_q = M_r \exp(j2\pi\vec{f}.\overrightarrow{\rho_0})$. Hence, we can write: $M_r M_q^* = |M_r|^2 \exp(-j2\pi\vec{f}.\overrightarrow{\rho_0})$; $M_r^* M_q = |M_r|^2 \exp(j2\pi\vec{f}.\overrightarrow{\rho_0})$. Now we define: $F(|M_r|^2) \equiv G_0(\vec{\rho})$, where F stands for FT. It then follows that, $T_3 = F(M_r M_q^*) = F(|M_r|^2 \exp(-j2\pi\vec{f}.\overrightarrow{\rho_0})) = G_0(\vec{\rho} - \overrightarrow{\rho_0})$. Similarly, $T_4 = F(M_r^* M_q) =$

$F(|M_r|^2 \exp(j2\pi\vec{f}.\overrightarrow{\rho_0})) = G_0(\vec{\rho} + \overrightarrow{\rho_0})$. The spatial extent of $G_0(\vec{\rho})$ is determined by the size of the image. Let us quantify this by defining a radial extent, $|\overrightarrow{\rho_m}|$ such that $G_0(\vec{\rho}) = 0$ for $|\vec{\rho}| \geq |\overrightarrow{\rho_m}|$. The behavior of the final signal depends on the value of the parameter, $\eta \equiv |\overrightarrow{\rho_0}|/|\overrightarrow{\rho_m}|$.

Case I: Consider first the situation where, $\eta \geq 1$. In this case, there is no overlap between $G_0(\vec{\rho} - \overrightarrow{\rho_0})$ and $G_0(\vec{\rho} + \overrightarrow{\rho_0})$ (i.e. between T_3 and T_4). Thus, the final detector signal of eqn (**10.6**) can be expressed as: $S_{fD} = |\beta|^2(|T_3|^2 + |T_4|^2)$. In this case, we will see two distinct peaks, corresponding to the cross-correlation (T_3) and anti cross correlation (T_4). We would like to point out that, for the sake of simplicity, this condition was implicitly assumed to hold in the discussions presented in ref. 34.

Case II: Consider next the situation where, $\eta < 1$. In this case, the final detector signal can be expressed as: $S_{fD} = |\beta|^2|G_0(\vec{\rho} - \overrightarrow{\rho_0}) + G_0(\vec{\rho} + \overrightarrow{\rho_0})|^2 = |\beta|^2|T_3 + T_4|^2$. The shape of this signal depends on the details of the images, and thereby on the details of T_3 and T_4. In what follows, we illustrate the shape of the signal S_{fD} for both case I and II, with a few examples.

For clarity, we consider first examples of one-dimensional images. Figure 4.4(a) and Figure 4.4(b) shows two images, with gaussian profiles, shifted from each other by $|\overrightarrow{\rho_0}| = 4\sigma$, where $\sigma = 5$ mm is the half- width of each image. Here, a reasonable estimate for $|\overrightarrow{\rho_m}|$ is 3σ, so that $\eta > 1$ is satisfied, corresponding to case I. As can be seen in Figure 4.4 (c), the signal now has two distinct peaks. Next, we consider again the same images, but with a smaller shift: $|\overrightarrow{\rho_0}| = 2\sigma$, shown in Figure 4.4(d) and Figure 4.4(e). In this case η =0.67, corresponding to case II, so that there will be overlaps between T_3 and T_4. However, the two peaks can still be discerned in the final signal, shown in Figure 4.5(f). Finally we consider a case where $|\overrightarrow{\rho_0}| = \sigma$, as shown in Figure 4.4(g) and Figure 4.4(h). In this case, $\eta = 0.25$ and

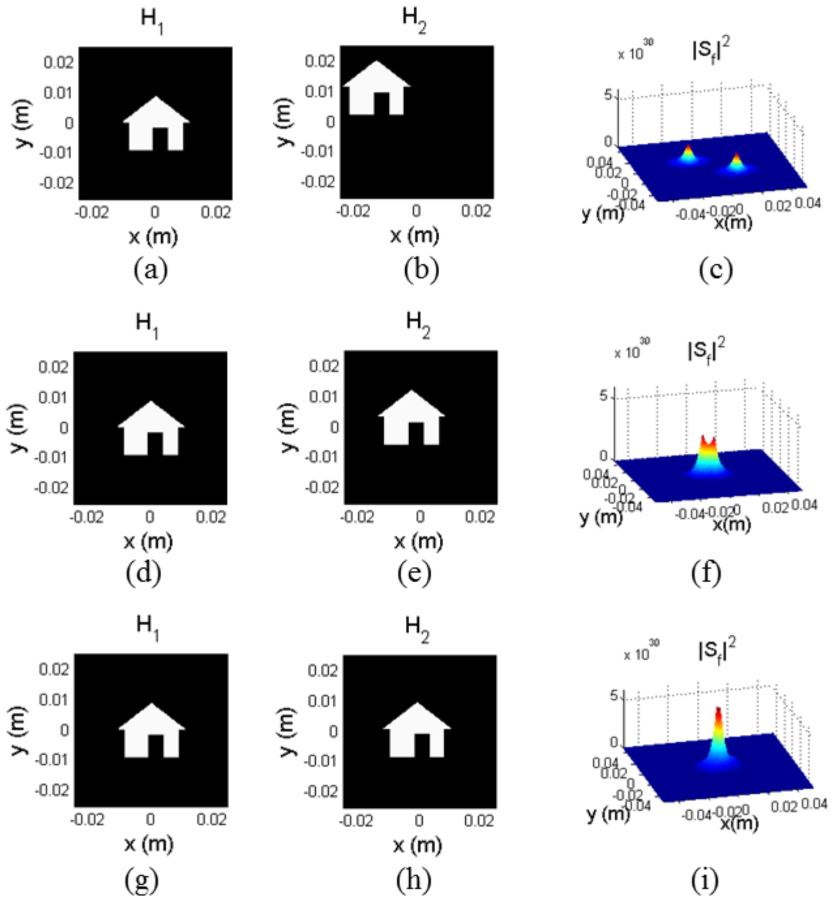


Figure 4.5: Illustration of the resolving power of the HOC architecture for two-dimensional identical images

we can see that the peaks are no longer distinguishable, and have merged into each other, as shown in Figure 4.4(i).

Next, we consider some examples in two dimensions. Figure 4.5(a) and Figure 4.5(b) show two identical images, shifted from each other by $|\overrightarrow{\rho_0}| \approx 0.02$ m. For this image $|\overrightarrow{\rho_m}| \approx 0.01$ m, so that

$\eta > 1$ is satisfied, corresponding to case I. As can be seen in Figure 4.5(c), the signal now has two distinct sharp peaks.

Next, we consider again the same images, but with a smaller shift: $|\overrightarrow{\rho_0}| \approx 0.004$ m, shown in Figure 4.5(d) and Figure 4.5(e). In this case $\eta \approx 0.4$, corresponding to case II, so that there will be overlaps between T_3 and T_4. However, the two peaks can still be discerned in the final signal, shown in Figure 4.5(f). Finally we consider a case where, $|\overrightarrow{\rho_0}| = 0.001$ m, as shown in Figure 4.5(g) and Figure 4.5(h). In this case, $\eta \approx 0.09$ and we can see that the peaks are no longer distinguishable, and have merged into each other, as shown in Figure 4.5(i).

If the cross-correlation peaks are clearly resolved, then we can infer the distance between the two matched images, given by half the separation between the peaks. However, this information cannot be retrieved when the peaks are not resolved. On the other, the ability of the HOC architecture to determine whether a match is found is not adversely affected by the potential overlap between T_3 and T_4. Additional implications of this potential overlap between T_3 and T_4 are addressed in other sections of this thesis. Note that in rest of the thesis, we will assume that the convolution terms (T_1 and T_2) have been eliminated by the phase stabilization and scanning circuit.

4.3 Experimental Results and Discussion

To overcome the material constraints of the Conventional Holographic Correlator (CHC), we proposed a Hybrid Opto-electronic Correlator (HOC) using only photo-detectors, Spatial Light Modulators (SLM), phase stabilization circuits, and VLSI chips[34]. In the HOC, the amplitude and phase information of the query as well as the reference image are recorded with photo-detectors via interference with plane waves. In this section, we show experimentally the basic working principle of the HOC architecture using current available technologies. We demonstrate experimentally, for a matched reference and query image, the output signal shows a sharp peak indicating a match is found. Whereas for unmatched case a lower peak value is observed in the output signal indicating no match. We also show the dependence of output signal on the phases of the interfering plane waves and describe a technique for optimizing the output signal. This result validates the scheme proposed in ref [34] for optimizing the correlation signal by averaging out the effect of the absolute phase. As such, the work reported here paves the way for further development of the HOC for practical applications.

4.3.1 Experimental Setup

The overall architecture of the proposed HOC is discussed in details in chapter 3. Primarily in describing the architecture of the HOC, we have considered the use of commercially available components such as cameras, FPGAs and SLMs. However, the overall process is severely slowed down during the serial communication between these devices. In order for the HOC to achieve its ultimate operating speed, we proposed a novel component called Integrated Graphic Processing Unit (IGPU)[34] to operate in parallel, which can perform the a single correlation operation in microseconds. However, the IGPU is not currently available in market and it is a long-term expensive project. Hence, for now we show the working principle of the HOC using a computer instead of an IGPU. Efforts are also underway to develop an IGPU in order to realize a high-speed version of the HOC.

To demonstrate the basic working principle of the HOC experimentally, we modify the architecture, as shown in Figure 4.6. A continuous wave diode pumped solid-state laser (Verdi V2) with 532 nm wavelength is used for this experiment. The laser beam has a diameter of 1 mm, which can be expanded to 1 inch using a spatial filter, and a lens. The expanded beam is split into two paths using a 50-50 beam splitter (BS). One path is used to retrieve the images and the other path is used to create the Phase Stabilization Circuit (PSC), which is described in details later in section 4.3.2. Instead of using the SLM, as proposed in the original architecture, we use a portion of the Air Force Resolution Chart (ARC) to

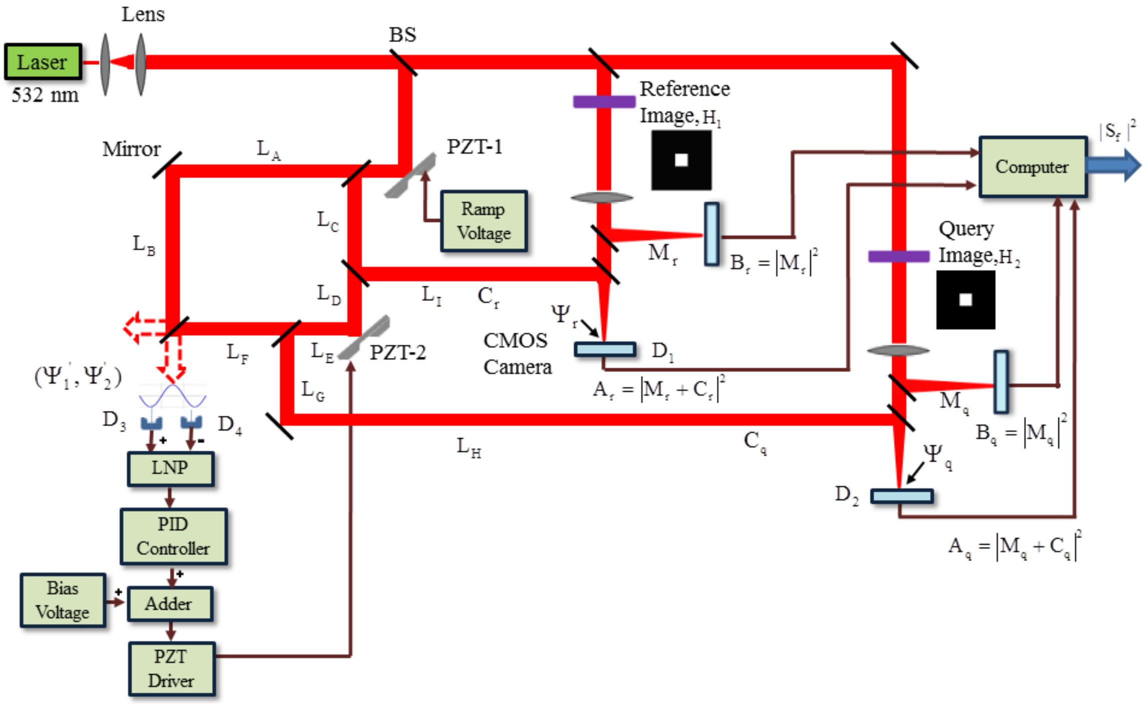


Figure 4.6: Simplified architecture of the HOC for demonstrating basic working principle.[BS = Beam Splitter]

produce the reference and query images. The reference image, H_1 is Fourier transformed (FT) using a lens. The Fourier transformed image (M_r) is split into two identical ports. In one port, a digital CMOS camera detects the image. In one port, a digital CMOS camera detects the signal M_r. We have used the USB2.0 CMOS Camera (DCC1545M), which has 1280H x 1024V pixels. The camera is interfaced with a computer via an USB cable. The signal array produced by the camera is denoted as B_r and stored in the computer memory. In the other port, M_r is interfered with a plane wave C_r, and detected with another CMOS Camera, producing the digital signal array A_r . This signal array is also stored in the memory of the computer. The equation of A_r and B_r are defined previously in eqn (**3.1**) and (**3.2**). Here $\phi_r(x,y)$ is the phase of the Fourier transformed image, Mr, and Ψ_r is the phase of the plane wave, C_r at the detector D_1. In addition, the intensity profile of the plane wave ($|C_r|^2$) is measured, by blocking the image path momentarily, using a shutter (not shown). A MATLAB script then uses the stored data to compute the signal, Sr according to eqn (**3.3**). In a similar manner described above, the query image, H_q is transferred to an optical beam using another ARC, and split into two paths after being Fourier transformed with a lens. The resulting image in each path is designated as M_q. In one port, a digital CMOS camera detects the FT image, where the detected signal is denoted as B_q. In the other port, M_q is interfered with a plane wave C_q, and detected with another CMOS Camera, producing the digital signal array A_q. The same MATLAB script mentioned above calculates the signal Sq according to eqn (**3.6**). As before, $\phi_q(x,y)$ is the phase of the Fourier transformed image, M_q, and Ψ_q is the phase of the plan wave C_q at the detector D_2. In the final stage of the HOC, these two signals (S_r and S_q) are multiplied together. The resulting signal array, S can be expressed as eqn (**3.7**). According to the original architecture of the HOC, the signal array, S, is transferred to the optical domain using an SLM (not shown in Figure 4.6). The optical image produced by the SLM is Fourier transformed using a lens, and detected by a Focal Plane Array (FPA). The output of the FPA provides the main correlation signal. The final signal can thus be expressed as:

$$|\mathrm{S_f}|^2 = |\alpha^*\mathrm{F}(\mathrm{M_r M_q}) + \alpha\mathrm{F}(\mathrm{M_r^* M_q^*}) + \beta^*\mathrm{F}(\mathrm{M_r M_q^*}) + \beta\mathrm{F}(\mathrm{M_r^* M_q})|^2 \tag{4.2}$$

Where, $\alpha \equiv |C_r||C_q|e^{j(\psi_r+\psi_q)}$; $\beta \equiv |C_r||C_q|e^{j(\psi_r-\psi_q)}$ and F stands for the Fourier Transform. Here, we calculate the final output signal according to eqn 6 using MATLAB in computer for simplicity. We can see that the final signal is a sum of four terms; the first two terms corresponds to convolution of the reference and the query images, H_1 and H_2, respectively. The last two terms corresponds to the cross-correlation and anticross-correlation of H_1 and H_2. The cross-correlation signal is usually used to find matches between two objects. If a match is found between the reference and the query images, the output signal gives a high amplitude sharp peak whereas, for unmatched images sharp peak is not observed. The convolution terms can be washed out by implementing the phase stabilization and scanning technique, which is described in details in section 3.2. Since we are limiting our discussion for centered images, the cross-correlation and the convolution terms overlap one another and it is out of scope of this thesis to eliminate the convolution terms.

Since the output signal depends non-trivially on the phases of plane waves (eqn (**10.6**)), it is important to stabilize and adjust the phase to get the optimum output signal. First, we show experimentally that the Phase stabilization circuit is capable of stabilizing the phases of the plane waves (C_r and C_q). Then we show the simulation and experimental results that the output signal depends non-trivially on both the *relative* and the *absolute* phases of the two plane waves at the planes of interference. Lastly, we illustrate the behavior of the HOC for identical and different reference and query images under optimized phase conditions.

It is very important to encapsulate the whole setup and tightly secure the optical mounts to the table, as the whole setup is extremely sensitive to environmental vibrations and air currents. Hence, a floating optical table is used to isolate the system from any kind of floor vibration. An optical enclosure made of plexi glass is used to ensure that the air current does not change the optical path lengths.

4.3.2. Demonstration of the Phase stabilization circuit

As we have mentioned in section 4.3.1, the output of the laser is split into two paths. One path is used for generating images and the other path is used for generating the Mach-Zehnder type interferometer for the PSC. The interference pattern along with a servo and PZT (PZT-2) is used to stabilize the phases, Ψ_r and Ψ_q .Since we cannot directly adjust the phase of the plane waves, Ψ_r and Ψ_q at the detector plane D_1 and D_2, we take an indirect measure to adjust the phases assuming all other optical path lengths are constant. The mirror is on the path of C_q is used to control the value $(\Psi_r - \Psi_q)$ while keeping the Ψ_r constant assuming the optical path length is fixed. Now, Ψ_r and Ψ_q can be written as $\Psi_r = k(L_C + L_I)$ and $\Psi_q = k(L_C + L_D + L_E + L_G + L_H)$. Here k is the wavenumber and L_j is the optical path length starting from the Beam splitter (BS) (which creates the two beams for the interferometer). Now the phases of the interfering plane waves Ψ'_1 and Ψ'_2 at the detectors, D$_3$ and D$_4$ can be written as $\Psi'_1 = k(L_A + L_B)$ and $\Psi'_2 = k(L_C + L_D + L_E + L_F)$. Thus, we can write $(\Psi_r - \Psi_q)/k = (\Psi'_1 - \Psi'_2)/k + \xi$ where $\xi = (L_C + L_F + L_I) - (L_A + L_B + L_G + L_H)$ is the path length difference and it is assumed to be constant. Hence, if we make an adjustment in the path L$_E$ keeping the other path lengths constant (i.e. Ψ_r) , we will be able to adjust the phase Ψ_q. A feedback loop along with a servo can fixed the value $(\Psi'_1 - \Psi'_2)$ which is also adjusting the value $(\Psi_r - \Psi_q)$. Below we describe how this can be achieved experimentally.

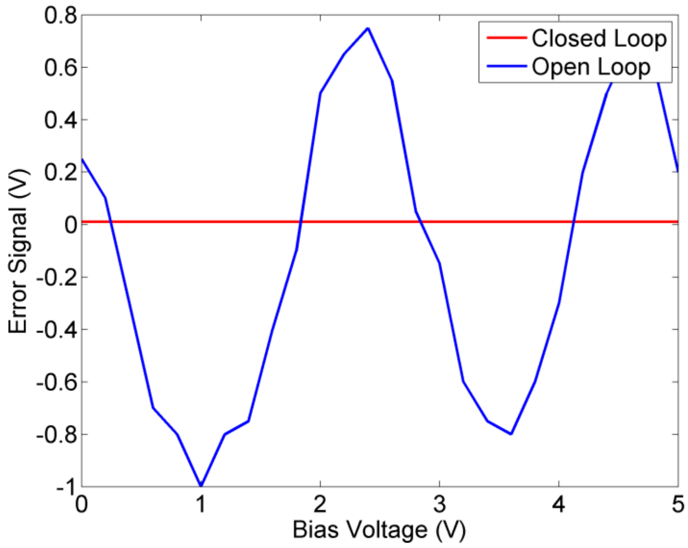


Figure 4.7: Experimental demonstration of the feedback loop to stabilize the phases of the plane waves. See text for details.

A pair of matched detectors (D$_3$ & D$_4$) detects the interference signal due to interferometer. The voltages from these detectors are subtracted from each other using a Low Noise Pre-amplifier (LNP). The Model number is SR- 560 and it has the capability of amplifying the subtracted signal. The resultant voltage from the LNP is fed to the analog PID controller (SIM 960). The output from the PID controller is added with a bias voltage using an adder circuit and the output of this adder is fed to PZT-2 via a PZT Controller (MDT 694A). The parameters of the PID controller have to be tuned correctly such that the system is locked at the desired position despite the environmental noise. The PID controller also ensures that the system does not become unstable while collecting data. The feedback loop can lock the interference pattern at any desired position.

Figure 4.7 shows that the feedback loop can fix the interference fringe to a desired position by keeping ($\Psi_r - \Psi_q$) to a fixed position. As mentioned before, we cannot measure the value of Ψ_r and Ψ_q, but implementing the feedback loop we can fix $(\Psi_1' - \Psi_2')$ i.e. ($\Psi_r - \Psi_q$) to a certain value . The detectors D_3 and D_4 detect the intensity of two portions of the interference fringe and the voltage from the subtractor is denoted as the error signal. If a ramp bias voltage is applied, it would make the error signal vary sinusoidally in an open loop system as no correction is applied to the PZT-2 to nullify the changes. However, in a closed loop system, the change due to bias voltage can be nullified using the PID controller based feedback loop. Hence, the error signal and the interference fringe is fixed to zero or to any desired value as shown in Figure 4.7. Thus, we can claim that the phases Ψ_r and Ψ_q are constant in a closed loop system. From Figure 4.7 we can also show that with an increasing bias voltage the error signal varies sinusoidally in a open loop system while the error signal is constant in a closed loop system.

4.3.3 Phase dependence of the Output signal

From eqn (**10.6**), it is obvious that the final signal S_f depends nontrivially on Ψ_r & Ψ_q. Here the first term corresponds to the convolution and the second term corresponds to the cross-correlation. In this section, we show that the final output signal $|S_f|^2$ is a function of the phases Ψ_r and Ψ_q, which is verified using the simulation and experimental results.

Figure 4.8 shows the simulation of the output signal that depends on the relative and the absolute phase of the phase Ψ_r & Ψ_q which can be shown by varying the phase Ψ_q keeping the

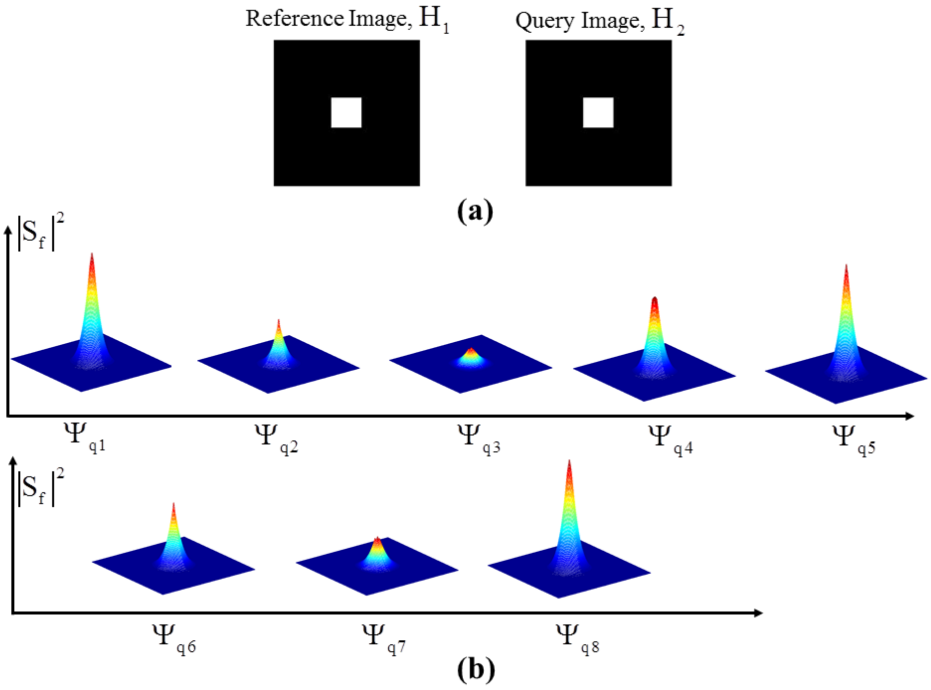


Figure 4.8 : (a) The reference and the query images (b) Simulation result of the output signal varying with phase ψ_q, *considering the phase* ψ_r *is constant.*

phase Ψ_r. Here, the phase Ψ_q is varied between 0 to 2π i.e. over one cycle while keeping Ψ_r constant. For this particular example, the reference and the query image are identical images as shown in Figure 4.8(a). For this particular case, since the reference and the query image are similar and positioned at the center, the four terms in eqn (**10.6**) overlap with one another at the center. Figure 4.8(b) shows the results that the final output signal $|S_f|^2$ varies as Ψ_q is varied between 0 to 2π keeping Ψ_r constant.

Figure 4.9 shows the experimental result demonstrating the output signal $|S_f|^2$ is varying

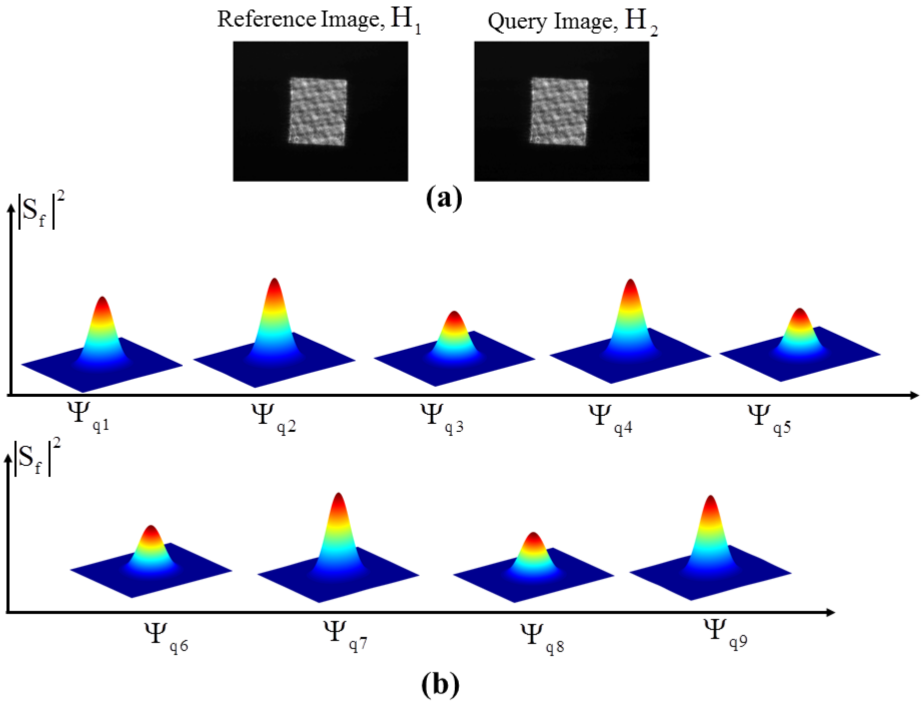


Figure 4.9: (a) The reference and the query images retrieved from a portion of the ARC (b) Experimental result of the output signal varying with phase as we change the bias voltage.

with the phase ψ_q . Figure 4.9(a) shows the reference and the query image both retrieved from two different ARC in the reference arm and the query arm. To make sure that both images are at the center of the camera plane, they have to be adjusted in x and y directions while the lens is out of the path. As discussed before, the terms A_r, B_r, Cr, A_q, B_q and C_q are captured via a CMOS camera (DCC1545M) and transferred to the computer via an USB cable. The final output signal $|S_f|^2$ is calculated using a MATLAB script in the local computer.

As mentioned above, any adjustments in the position of the PZT-2 changes the phase Ψ_q , as

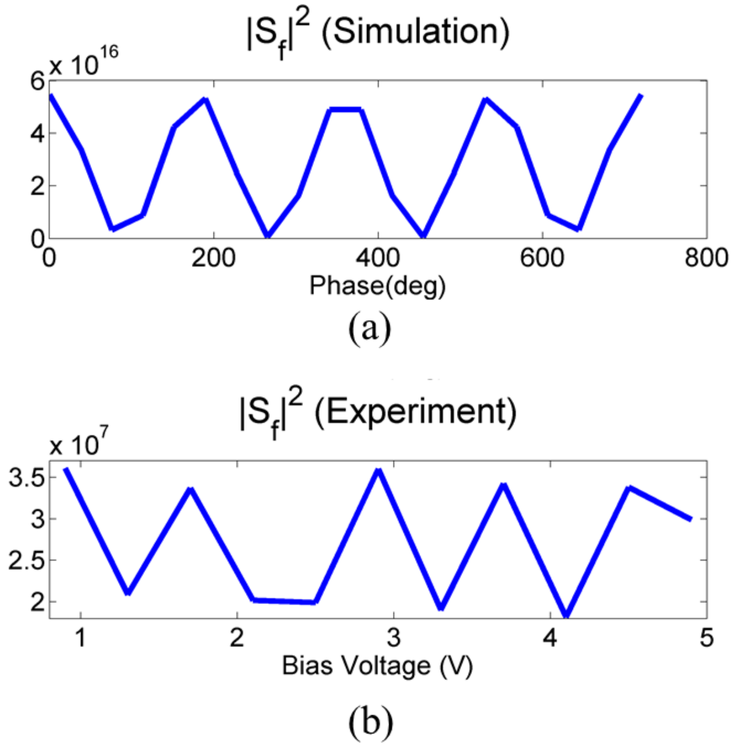


Figure 4.10: The amplitude of the output signal varies sinosoidally with the applied bias voltage or phase as shown (a) in simulation results and (b) experimental results

the PZT -2 is in the path of the plane wave $C_{q.}$ Now, PZT-2 can be deflected by a fixed amount by changing the bias voltage. Here we have assumed the path of the plane wave C_r is fixed hence the phase Ψ_r is constant. Now we increase the bias voltage linearly, lock the system, and capture the output signal $|S_f|^2$ for each bias voltage while the system is locked. In Figure 4.10, we shows the comparison between the simulation and experimental results and shows how the amplitude of the output signal vary as a function of bias voltage i.e. the phase of the plane wave C_q.

4.3.4 Object Detection using HOC

Figure 4.11 illustrates the experimental results of the behavior the HOC for a set of identical reference and query images and for a set of unmatched reference and query images. Figure 4.11(a) shows the case where the reference image H_1 and query image H_2 where they are a square and are in the center of the image plane. For this particular case, the convolution and cross-correlation terms of eqn (**10.6**) overlap with one another as can be seen in the simulation result in Figure 4.8. Figure 4.11(b) shows the experimental result of the output signal $|S_f|^2$ has a sharp peak at the center which indicates that a match is found. Now Figure 4.11(c) shows another case, where two different images, H_1' and H_2' , are inputs to the HOC. Since they are not identical, we expect to get a lower peak of the output signal compared to the matched case. Figure 4.11(d) shows that the peak value of output signal for unmatched images are around 14 times lower than the matched case. Hence, we can set a threshold to make a decision if match is found, or not.

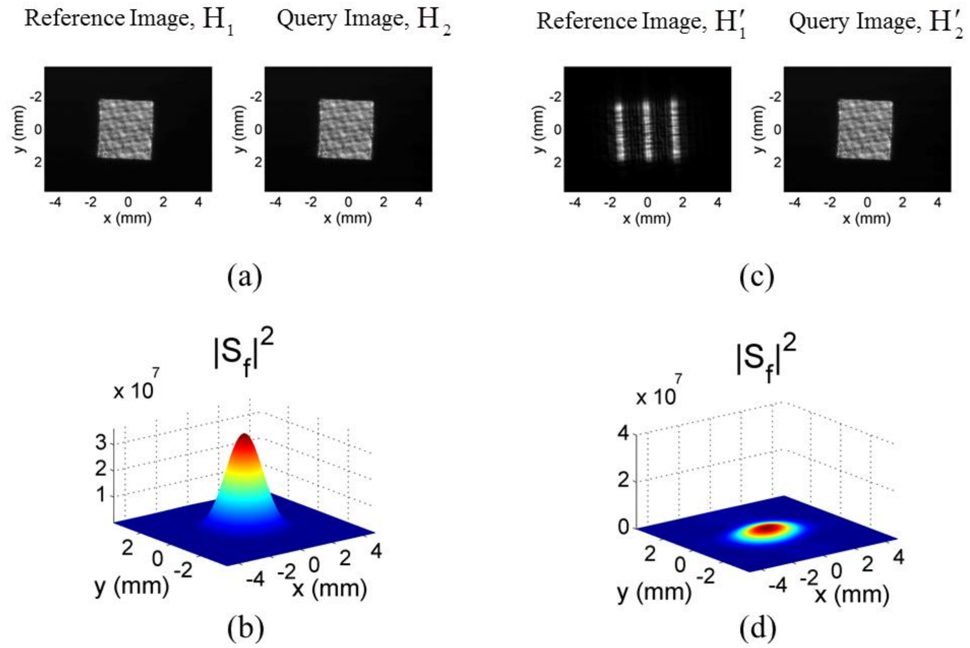


Figure 4.11: Experimental results of the HOC for identical images and unmatched images.

It is important to note that if the query and reference images are not shifted with respect to one also another, then the cross-correlation and the convolution peaks will be at the center of the detection plane, overlapping each other. If the images are shifted with respect to one another, the cross-correlations will be shifted symmetrically around the center by a distance corresponding to the shift.

4.4 Conclusion

Optical target recognition using correlators is an important technique for fast verification and identification of images. The Hybrid Optoelectronic Correlator (HOC) recently proposed by us bypasses the need for nonlinear materials such as photorefractive polymer films by using detectors instead, and the phase information is yet conserved by the interference of plane waves with the images. Here, we show through simulation and experimental results, that the correlation signal depends non-trivially on both the *relative* and the *absolute* phases of these two plane waves at the planes of interference. We then describe the technique for controlling and optimizing these phase parameters by using the interferometer with a servo. We show experimentally the behavior of the HOC for identical and different reference and query images under optimized phase conditions and verify that the HOC is capable of detecting objects using currently available technologies.

Chapter 5 Introduction to Polar Mellin Transform

5.1 Scale and Rotation Invariant employing Polar Mellin Transform

A position, rotation and scale invariant (PRSI) correlator can be realized using the same optical architecture as shown in Figure 3.1. To understand this process, it is useful first to summarize **five important properties** of 2D Fourier Transform, as follows:

- ***Property 1:*** *2D FT can be implemented using a lens (Appendix A and B)*
- ***Property 2:*** *The magnitude of 2D FT remains unchanged if the input images is shifted in position (Appendix A and B). The shift information appears as phase-shift in the 2D FT.*
- ***Property 3:*** *If an image is scaled so that its area is amplified by a factor* σ^2*, then its 2D Fourier Transform is amplified in area by a factor of* $(1/\sigma^2)$*.*
- ***Property 4:*** *If an input image is rotated, the 2D FT is rotated by the same angle*
- ***Property 5:*** *If an input image is expressed in terms of coordinates that are natural logarithms of the original coordinates, then the magnitude of its FT is insensitive to any change in scale of the original image (since it is Mellin Transform in the original coordinates, as shown in Appendix D). The scale change appears as a phase shift in the 2D FT. This property holds in both one and two dimensions.*

Using these properties of 2D FT, it is possible to realize an optical correlator that is position, rotation and scale invariant (PRSI), by employing the following set of steps:

- ***Step 1:*** *The 2D FT of a given image (e.g., the query image as well as the reference image), and its magnitude is determined. This magnitude will be independent of the position of the input image.*
- ***Step 2:*** *The FT magnitude is converted in polar coordinates. Qualitatively, this step can be thought of as decomposing the original image into a sum of polar-coordinate harmonics (which are simply the well known spherical harmonics, with the azimuthal angle set to zero). If the input image is rotated, the radial part of each harmonic remains unchanged, and the angular part is shifted by the angle matching the rotation.*
- ***Step 3:*** *Change the radial part of the polar coordinate to its natural logarithm, after normalizing it to a convenient length scale.*

5.2 Steps to Convert Images to Log Polar Domain

These steps are illustrated schematically in the flow diagram shown in Figure 6.1 in Chapter 6. Note that we transform the amplitude of the FT, $\left|V\left(x,y\right)\right|$, into the function $F\left(r,\theta\right)$. For $F\left(r,\theta\right)$, the coordinates r and θ are rectilinear (as opposed to curvilinear). To be specific, this means that for a two-dimensional integration over this (r,θ) space, the area element is given by $ds = dr \cdot d\theta$, and ***not*** by $ds = dr \cdot rd\theta$. More explicitly, the conversion is carried out as follows:

For a given combination of coordinates in this space, say $\{r = r_1; \theta = \theta_1\}$, we determine the corresponding values of *x* and *y* by using the relations $x_1 = r_1 \cos\theta_1$ and $y_1 = r_1 \sin\theta_1$. The value of the function F is then given by $F(r_1, \theta_1) = \left|V(x_1, y_1)\right|$. To plot the function *F*, this value will correspond to the point that is a distance r_1 away from the origin along the horizontal axis, and a distance θ_1 away from the origin along the vertical axis.

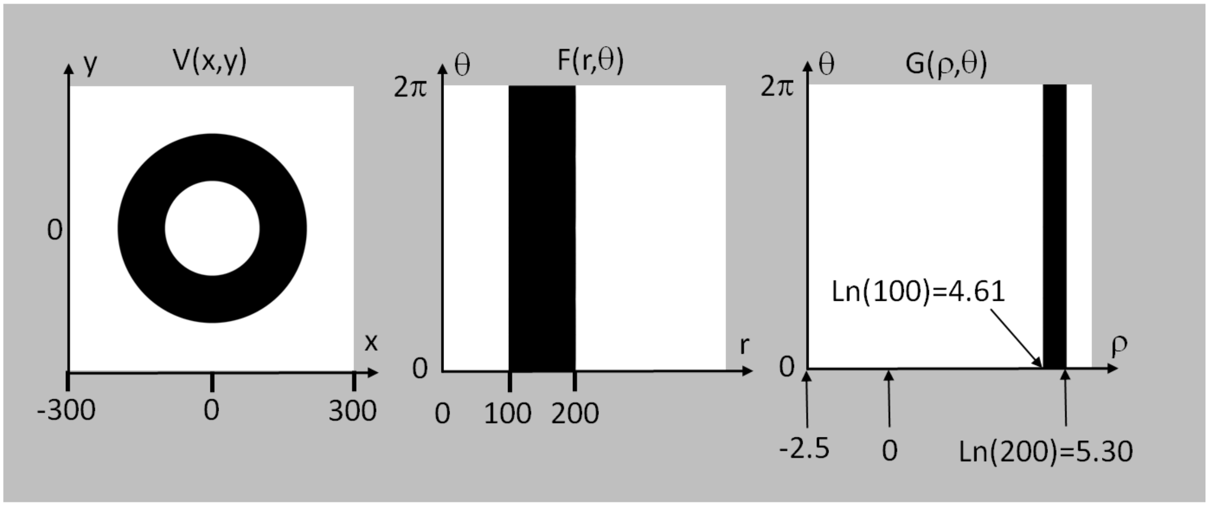


Figure 5.1: Illustration of the steps to the Polar Mellin Transform for a simple case where the amplitude of the FT of an input image is a circle of uniform intensity, with a hole at the center.

To generate the function $G(\rho, \theta)$, we again note that the coordinates ρ and θ are rectilinear. For a given combination of coordinates in this space, say $\{\rho = \rho_1; \theta = \theta_1\}$, we determine the corresponding values of *r* and *θ* by using the relations $\rho_1 = \ln(r_1)$ and $\theta_1 = \theta_1$. The value of the function G is then given by $G(\rho_1, \theta_1) = F(r_1, \theta_1)$. To plot the function *G*, this value will correspond to the point that is a distance ρ_1 away from the origin along the horizontal axis, and a distance θ_1 away from the origin along the vertical axis.

5.3 Examples of the PMT process

A very simple example of this process is illustrated in Figure 5.1. Here, we assume a very artificial case where the *amplitude of the Fourier Transform of an image* [i.e., V(x,y)] happens to be a nice function: a circle with uniform intensity, with a hole in the center, as shown on the left of Figure 5.1. Because of the inherent symmetry in this case, the computation of the corresponding polar function, $F(r,\theta)$ is rather trivial, and the result is shown in the middle. The corresponding polar-logarithmic function, $G(\rho,\theta)$ is shown on the right. A somewhat more

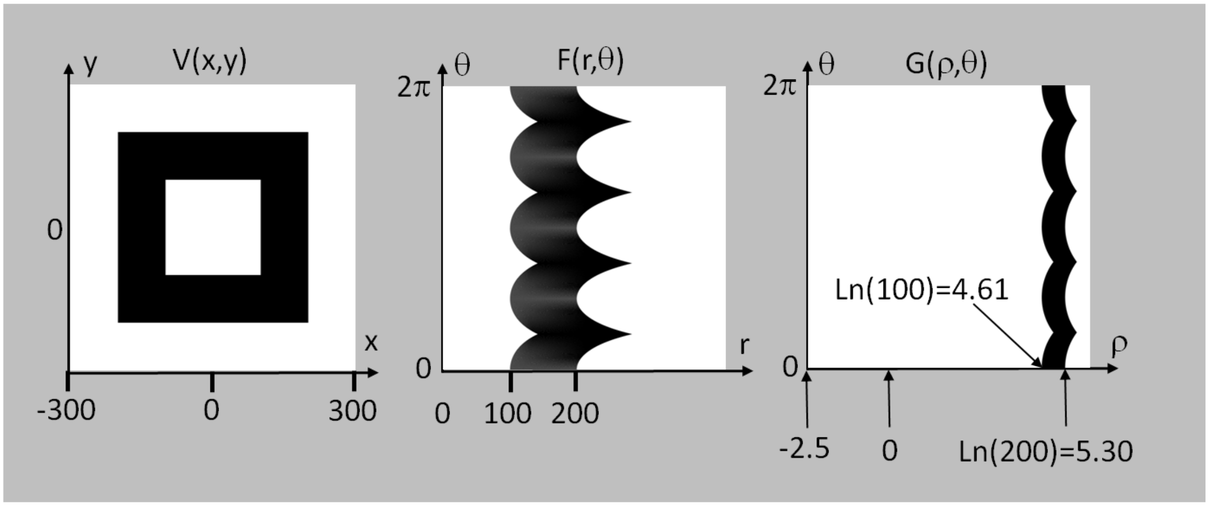


Figure 5.2: Illustration of the steps to convert to the log-polar domain for a non-trivial case where the amplitude of the FT of an input image is a square of uniform intensity, with a square shaped hole at the center.

complicated case is shown in Figure 5.2. for which the *amplitude of the Fourier Transform of an image* [i.e., V(x,y)] is assumed to be uniform square, with a square shape hole in it.

It is instructive to consider the effect of scale change on this conversion process. Let us assume, for example, that the physical image corresponding to the case shown in Figure 5.2 is reduced by a factor of 4 in each direction (representing a factor of 16 reduction in area). According to the one of the properties of FT's summarized earlier, this would imply that the area of the corresponding V(x,y) would be reduced by a factor of 16 (representing a linear reduction by a factor of 4 in each direction). Figure 5.3 shows such a case, where A represents

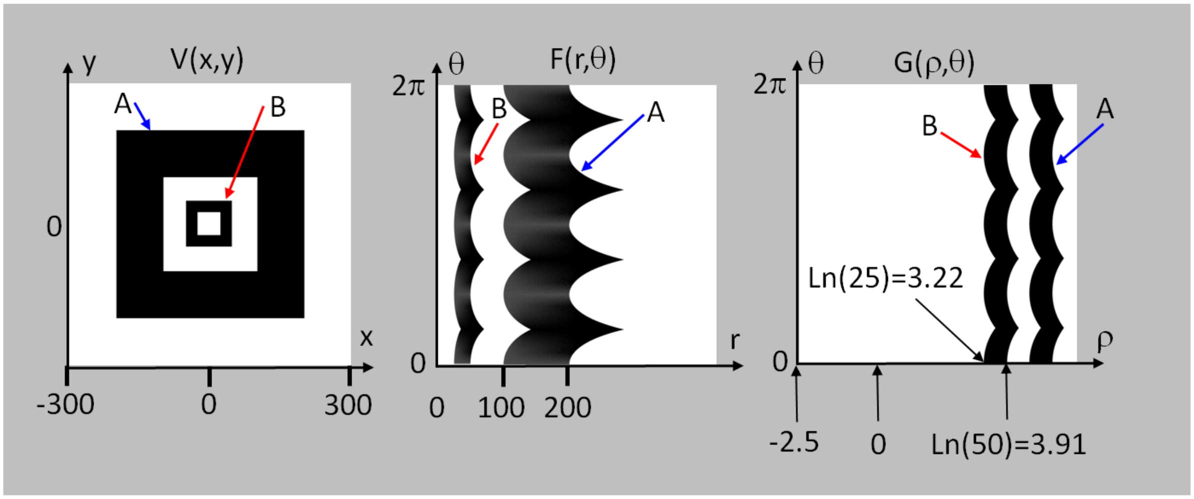


Figure 5.3: Illustration of the effect of scaling on the image conversion process. See text for details.

the same case as shown in Figure 5.2., and B represents the case with the changed scale. Thus, $V_B(x,y)$ is smaller in area than $V_A(x,y)$ by a factor of 16. Note first that the corresponding polar distributions, $F_B(r,\theta)$ and $F_A(r,\theta)$, are the same in the θ-direction, but differ by the linear scaling factor (4 in this case) in the *r*-direction. Note next that the polar-logarithmic distributions, $G(\rho,\theta)$, are ***identical in shapes***, except for a shift in the ρ-direction [equaling the natural logarithm of the scale factor: ln(4)≅1.39]. This illustrates the essence of how scale invariance is achieved via these transformations. The Fourier Transforms of $G_A(\rho,\theta)$ and $G_B(\rho,\theta)$ would be identical in amplitude, thus leading to strong cross-correlation when applied to the FT-based correlator described in Chapter 2 and 3.

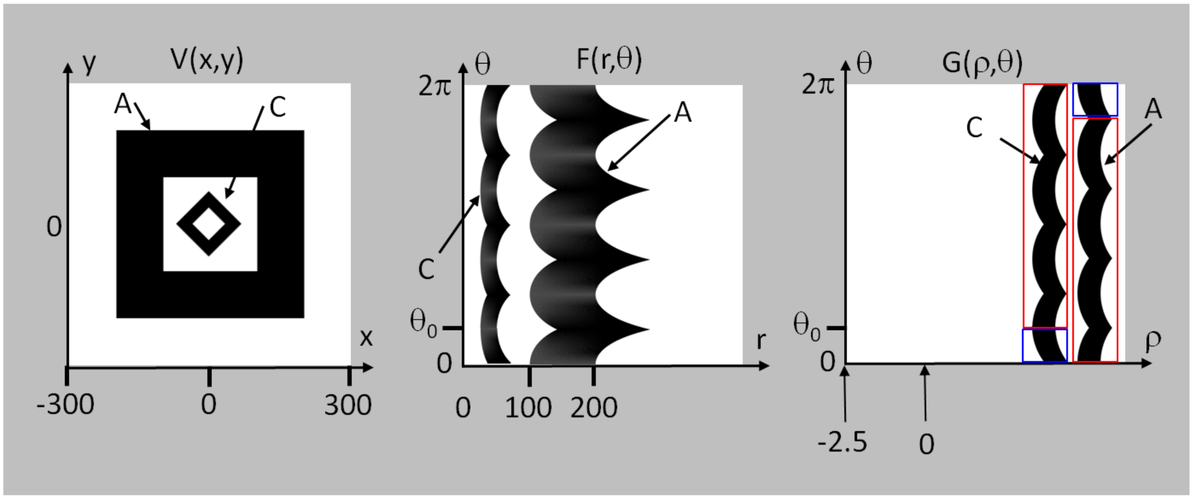


Figure 5.4: Illustration of the effect of scaling and rotation on the image conversion process. See text for details.

This feature is rather easy to understand by considering the following. Suppose that a given line at a fixed value of θ has end points r_{A1} and r_{A2} for $F_A(r,\theta)$, and has end points r_{B1} and r_{B2} for $F_B(r,\theta)$. If we assume a linear scale factor of σ, then $r_{B1}=\sigma\, r_{A1}$, and $r_{B2}=\sigma\, r_{A2}$. Now, we see that $\{\ln(r_{B2})- \ln(r_{B1})\}= \{\ln(r_{A2})- \ln(r_{A1})\}$, which shows the invariance of shapes in the polar-logarithmic distribution. The fact that $\ln(r_{B1}) = \ln(\sigma)+\ln(r_{A1})$, for example, shows why the pattern are shifted by $\ln(\sigma)$.

Consider next the effect of rotation in addition to the scale change. Again, we consider the same scale change as considered in Figure 5.3. In addition, the changed scaled image is assumed to be rotated by an angle θ_0. By virtue of the fourth of the properties of 2D FT's summarized above, the corresponding FT's would be rotated by the same angle. Figure 5.4 illustrates such a case, where A represents the same case as shown in Figure 5.2, and represents the case where the image is scaled up by a factor of 4 in each direction, and rotated by $\theta_0=\pi/4$. Thus, $V_C(x,y)$ is smaller in area by a factor of 16 than $V_A(x,y)$, and is rotated by $\pi/4$. Note first that the corresponding polar distributions, $F_C(r,\theta)$ *and* $F_A(r,\theta)$, now differ in both the r-direction and the θ-direction. In the r-direction, the change is the same as that shown in Figure 5.3. In the θ-direction, the pattern is shifted by $\theta_0=\pi/4$. Note next that the polar distributions, $G(\rho,\theta)$, are now still identical in shapes, except that $G_C(\rho,\theta)$ is now shifted from $G_A(\rho,\theta)$ in both the ρ-direction and the θ-direction. It is particularly important to consider the shift in the θ-direction very carefully, since this coordinate is limited to a range of 0 to 2π. Specifically, the distribution along the θ-direction in $G_A(\rho,\theta)$ can be broken up in two parts: the part enclosed in the red box and the part enclosed in the blue box. In $G_C(\rho,\theta)$ the part in red box is shifted by a angle θ_0, while the part in the blue box is shifted by $-(2\pi-\theta_0)$. Thus, when $G_A(\rho,\theta)$ and $G_C(\rho,\theta)$ are applied (e.g., as the query image and the reference image, respectively) to the conventional FT-based correlator, the cross correlation will have two peaks corresponding to these two parts. Each peak will be shifted in the ρ-direction by an amount of ln(σ) [where σ=4 is the linear scaling factor in this case]. In the θ-direction, the spot corresponding to the red box will be shifted in one direction by an angle of $\theta_0=\pi/4$, while the spot corresponding to the blue box will be shifted in the opposite direction by an angle of $(2\pi-\theta_0)=7\pi/4$. In general, the brightness of the spots will be unequal, since the amounts of signals in the two parts are unequal. The sum of the signals from the two spots is a constant, matching the value of the signal for the single spot that would result for no rotation.

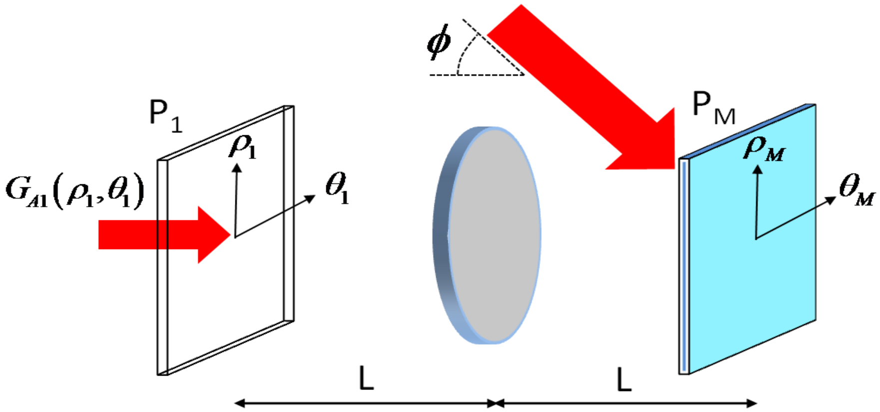


Figure 5.5: Recording the correlation filter for the PRSI correlator, with the polar-logarithmic version of the reference image at the input plane of a conventional optical correlator.

The next step in the PRSI process is to apply the polar-logarithmic version of the reference image[$U_{A1}(x,y)$], denoted now as $G_{A1}(\rho_1,\theta_1)$, at the input of a conventional optical correlator set-up, as shown in Figure 5.5. A plane wave beam is now applied to the middle plane at an angle ϕ, just as in the case of the conventional correlator as described in Chapter 2. Once the interference is recorded in the thin photographic plate, we now apply to the input plane the polar-logarithmic version of the query image [$U_{B1}(x,y)$], denoted now as $G_{B1}(\rho_1,\theta_1)$, as shown in Figure 5.6. If there is no relative rotation or scale change between the matching query and reference images, we will see a single cross-correlation spot, just as in the case of the conventional correlator, located in the output plane at:

$$x_2 = 0; y_2 = -L\sin\phi \quad \Leftrightarrow \quad \theta_2 = 0; \rho_2 = -L\sin\phi$$

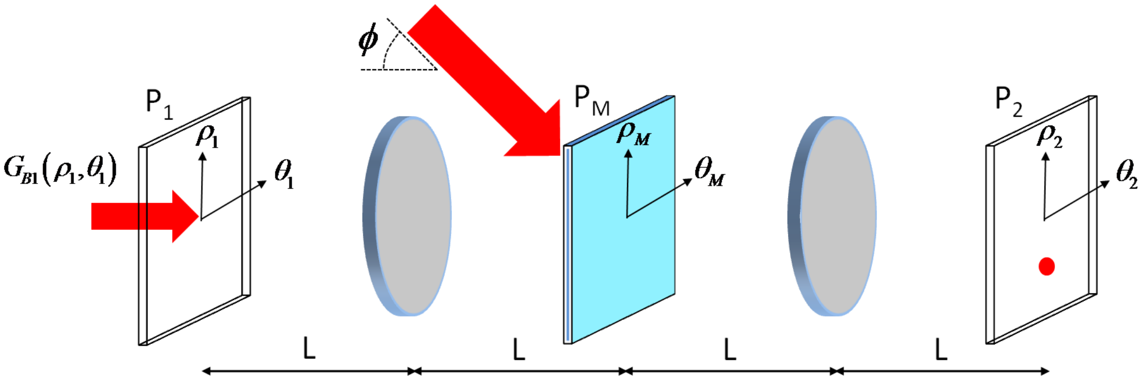


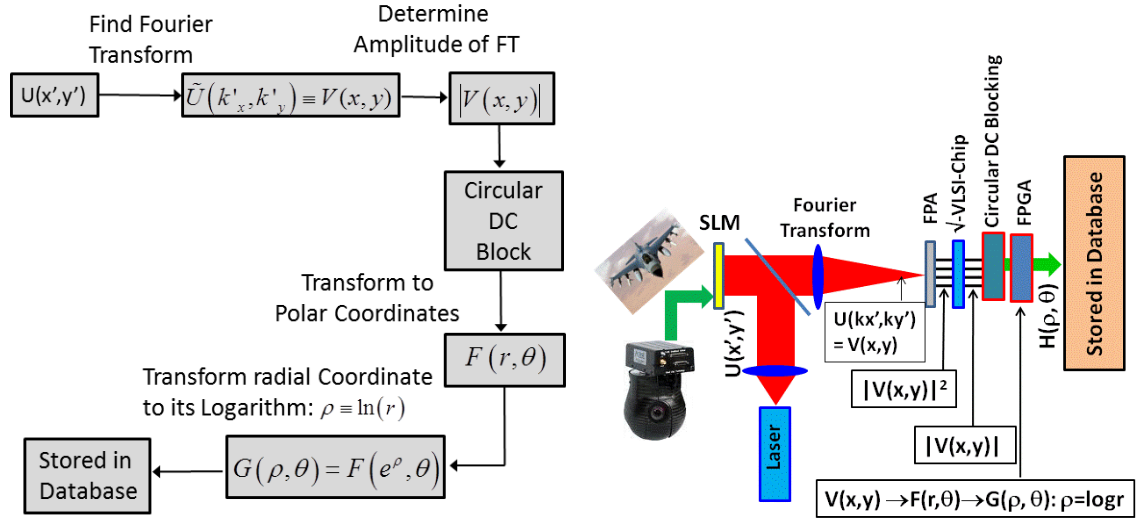
Figure 5.6: Illustration of the correlation stage of the PRSI correlator.

Here, we have assumed, for convenience that, that the arbitrary choice for the origin of the r-coordinate is chosen to be so that the function $F(r,\theta)$ corresponding to this image is centered at $r = 1 \Leftrightarrow \rho = 0$. If the matched query image is a scaled version of the reference image, by a linear scale factor σ, then the correlation spot will be further shifted radially, by an amount that corresponds to the log of the scale factor: $\ln(\sigma)$.

If the matched query image is a rotated version of the reference image, by an angle θ_0, then the correlation spot will be split in two parts. One spot will be shifted by an angle equaling θ_0, while another spot will be shifted by an angle $-\left(2\pi - \theta_0\right)$. These two spots will in general be of unequal intensities, with the sum of the intensities matching the intensity of the spot for the unrotated case, as discussed above. These effects of scaling and rotation on the appearance of the correlation spots are discussed in details in next chapter.

Chapter 6 Incorporation of Polar Mellin Transform in the HOC for Scale and Rotation Invariant Target Recognition

6.1. Incorporating Polar Mellin Transform in the HOC architecture



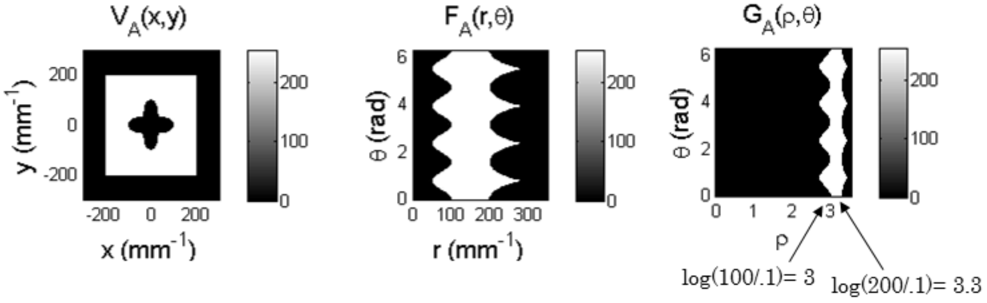
(a) (b)

Figure 6.1: (a) Flow diagram for transforming of query/ reference image to the log-polar domain. (b) Schematic illustration of the architecture for implementing PMT.

If we start with regular reference and query images for the HOC architecture described in chapter 3, it can detect images in shift invariant manner only. The architecture can be extended to achieve scale and rotation invariance, along with shift invariance, by transforming the reference and query images to log-polar domain. The flow diagram of performing this transformation, which is generally called the polar mellin transform (PMT) is described in Figure 6.1(a) and the detailed architecture is described in Figure 6.1(b). We start with an image U(x',y'), (query or reference image and converted to optical domain using a Spatial Light Modulator (SLM) as shown in Figure 6.1(b)). The co-ordinates have dimensions of distance, e.g. meter. The next step is to find the FT of the image, $\widetilde{U}\left(k'_x, k'_y\right)$ where the co-ordinates have dimensions of per meter. For notational convenience, we redefine $k'_x \rightarrow x$ and $k'_y \rightarrow y$, and denote as V(x,y) to be the same as $\widetilde{U}\left(k'_x, k'_y\right)$. In practice, the original image, U(x',y') would be represented in the optical domain by using an SLM linked to a camera or a computer data base. A lens would be used to find the fourier transform: $\widetilde{U}\left(k'_x, k'_y\right) \equiv V(x, y)$. An FPA detects the intensity of the Fourier transformed image i.e. $|V(x, y)|^2$. The FPA is interfaced with the FPGA which determines the value of $|V(x, y)|$, thus eliminating the phase information. The magnitude of the FT of an object or function is invariant to a shift in the function $|FT\{f(x\,, y)| = |FT\{f(x - x_0)|$, but not to a scale change in the input. A circular hole of small radius (e.g. radius of 5 unit) on $|V(x, y)|$, is created using the FPGA. The necessity of creating this hole is discussed in detail in section 6.2, where we also point out that in general this hole does not affect significantly the performance of the correlator. To achieve scale and rotation invariance, the amplitude of $V(x, y)$ is transformed to the polar co-ordinate function, $F(r, \theta)$, using the FPGA. Then it is converted to the log-polar co-ordinate function $G(\rho, \theta)$ using the same FPGA. The reference images are polar mellin transformed and stored in the database. The database can be a computer or a holographic memory disk if fast retrieval of the reference image is required. The captured query image also goes through the same procedure of PMT and supplied to the correlator's input port. In the next section, we describe the PMT process in detail, including illustrative examples.

6.2 Pre Preprocessing the image using polar mellin transform

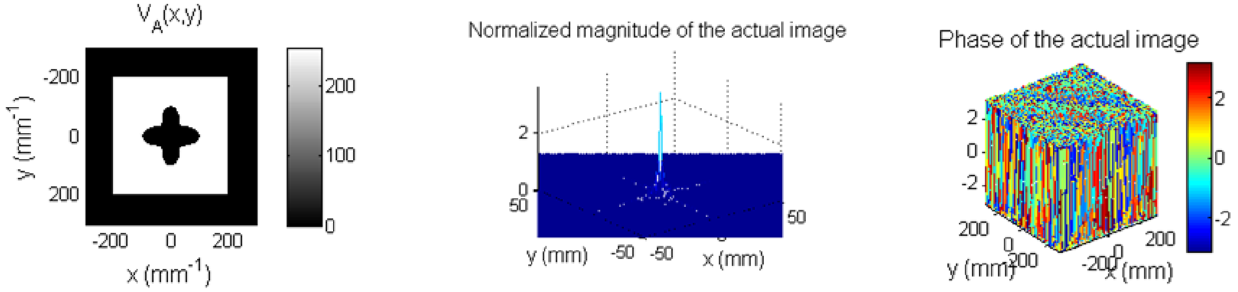
A simple example of the PMT process is illustrated in Figure 6.2. Here, we assume an artificial case



(a) (b) (c)

Figure 6.2:(a) V(x,y) is the amplitude of the Fourier Transform of an image with a flower shape hole in it (b) The corresponding polar function, F(r,θ) is shown (c) polar-logarithmic function, G(ρ,θ) is shown [here, $r_0 = 0.1$*].*

where the amplitude of the FT [i.e., $V(x, y)$] of an image is chosen to be an uniform square, with a flower shape hole in it, as shown in Figure 6.2(a). The corresponding polar function, $F(r, \theta)$is shown in Figure 6.2(b) and the corresponding polar-logarithmic function, $G(\rho, \theta)$ is shown in Figure 6.2(c). Note that for $F(r, \theta)$the coordinates r and θ are rectilinear (as opposed to curvilinear). For a given combination of coordinates in polar space, say $\{r = r_1;\theta = \theta_1\}$, we determine the corresponding values of x and y by using the relations $x_1 = r_1\cos\theta_1$ and $y_1 = r_1\sin\theta_1$. The value of the function F is then given by $F(r_1, \theta_1) = |V(x_1, y_1)|$. To plot the function F, we put the corresponding value of F at the co-ordinate (x_1, y_1) to a point that is a distance r_1 away from the origin along the horizontal axis (which has the dimension of inverse length, mm^{-1}) and a distance $\theta_1$away from the origin along the vertical axis (which is in the dimension of radian, and span from 0 to 2π) .



(a) (b) (c)

Figure 6.3:(a) Shows $V_A(x,y)$ which is the FT of an image. (b) Shows the magnitude of an image whose FT is $V_A(x,y)$ (c) Shows the phase of the actual image whose FT is $V_A(x,y)$.

To generate the log-polar co-ordinate function $G(\rho,\theta)$, we proceed as follows, For a given combination of coordinates in this space, say $\{\rho = \rho_1; \theta = \theta_1\}$, we determine the corresponding values of r and θ (i.e. r_1 and θ_1) by using the relations $\rho_1 = \log(r_1/r_0)$ and $\theta_1 = \theta_1$. Here the choice of the scaling distance, r_0 , is arbitrary. Note that the value of $\log(r/r_0)$ approaches -∞ as r approaches zero, for any finite value of r_0. Obviously, this is an impractical situation. To circumvent this problem, we choose to ignore the information contained in a small circle of radius r_0 (in the $V(x,y)$ plane), centered around r = 0, thus restricting the lower range of $\rho = \log(r/r_0)$ to 0, corresponding to $r = r_0$. The magnitude of r_0 should be chosen judiciously so as not to exclude any critical feature that may be present within the exclusion zone of $0 \le r \le r_0$. Of course, for the particular case shown in Figure 6.2, already a dark part in the center of $V_A(x,y)$. Thus, there is no loss of information if the circle of radius r_0 is fully contained in the small dark part. Figure 6.3(b) and Figure 6.3(c) show the amplitude and the phase, respectively, of the image whose magnitude of FT is $V_A(x, y)$ as shown in Figure 6.3(a). From Figure 6.3(c) it is clear that such an image whose magnitude of the FT is $V_A(x, y)$, is unrealistic because the phase of the actual image is spanning between $-\pi$ to π. While in this section, we restrict ourselves to unrealistic images where FT's have holes at the center, in section 6.3 and 6.4 we will consider realistic images for which it would be essential to exclude a small circle in FT plane.

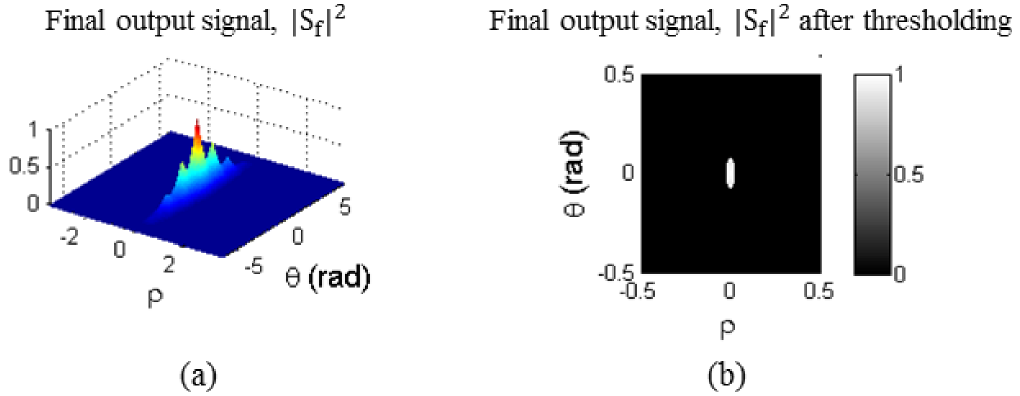


Figure 6.4: (a) The final output signal when two identical images are inputs to the HOC. The PMT version is shown in Figure 6.2(c). (b) The output signal after thresholding shows that there is a peak in the center when match between two objects are found.

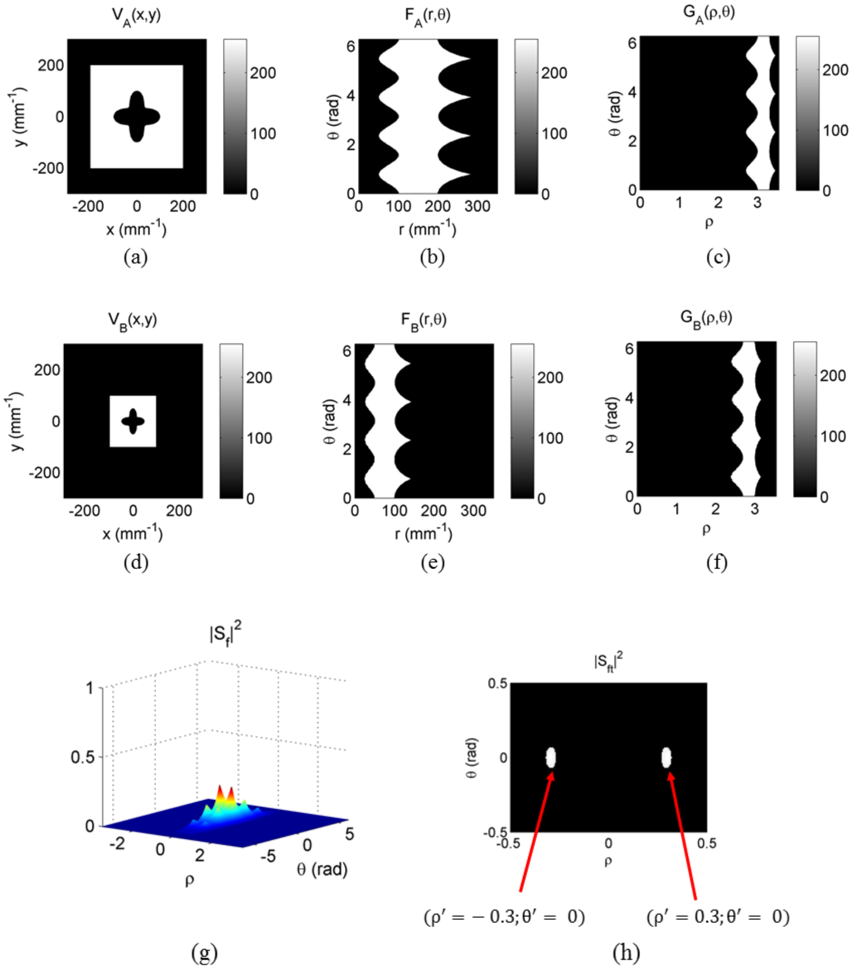
Before proceeding further, we note that the necessity of excluding a small circle in the FT plane was not addressed in the previous original investigations[13,14,45] pertaining to the use of PMT's for rotation and scale invariant correlations. This is due in part to the fact that these papers considered artificial images where FT's already contained holes in the center. Any realistic image, on the other hand, is bound to have a non-zero value at the center of the FT, corresponding to the average value of the image amplitude. However, in later work involving this approach, this issue was addressed[18,19].

Consider now a situation where two identical PMT images, each corresponding to Figure 6.2(c), are applied as inputs to the HOC. The resulting final output signals, $|S_f|^2$ of the HOC is illustrated in Figure 6.4(a). Because of the perfect match, the output has a sharp peak at the center, corresponding to the sum of the terms T_3 and T_4 (which are identical in this case) of eqn. (**4.1**). The value (~ 1) at the peak, of course, is arbitrary, depending on the magnitude of the images. In Figure 6.4(b), we show the peak clearly by applying a threshold value of 0.9.

6.3 Scale Invariant Image Detection using an Idealized Image

We consider now the effect of scale change on this PMT conversion process. Figure 6.5 shows such a case, where $V_A(x, y)$ represents the amplitude of the FT of the same image as considered in Figure 6.2, and $V_B(x, y)$ represents the amplitude of the FT of the same image, but scaled *up* by a linear factor of $\sigma = 2$ (i.e., x 4 larger in area). Thus $V_B(x, y)$ is scaled *down* by a linear factor of 2 (x4 smaller in area) compared to $V_A(x, y)$. Note first that the corresponding polar distributions (as shown in Figure 6.5(b) and Figure 6.5(e)), $F_B(r, \theta)$ and $F_A(r, \theta)$ are the same in the θ-direction, but differ by the linear scaling factor (2 in this case) in the r-direction. Note next that the corresponding polar-logarithmic distributions, $G_A(\rho, \theta)$ and $G_B(\rho, \theta)$ as shown in Figure 6.5(c) and Figure 6.5(f) respectively, are identical in shape, except for a shift in the ρ-direction, equaling the logarithm of the scale factor: $\log(2) \cong 0.3$. This illustrates the essence of how scale invariance is achieved via these transformations. The fourier transforms of $G_A(\rho, \theta)$ and

$G_B(\rho,\theta)$ would be identical in amplitudes, thus leading to a strong cross-correlation when applied to the HOC, which eliminates the effect of relative shift**.**



(a) (b) (c)

(d) (e) (f)

(g) (h)

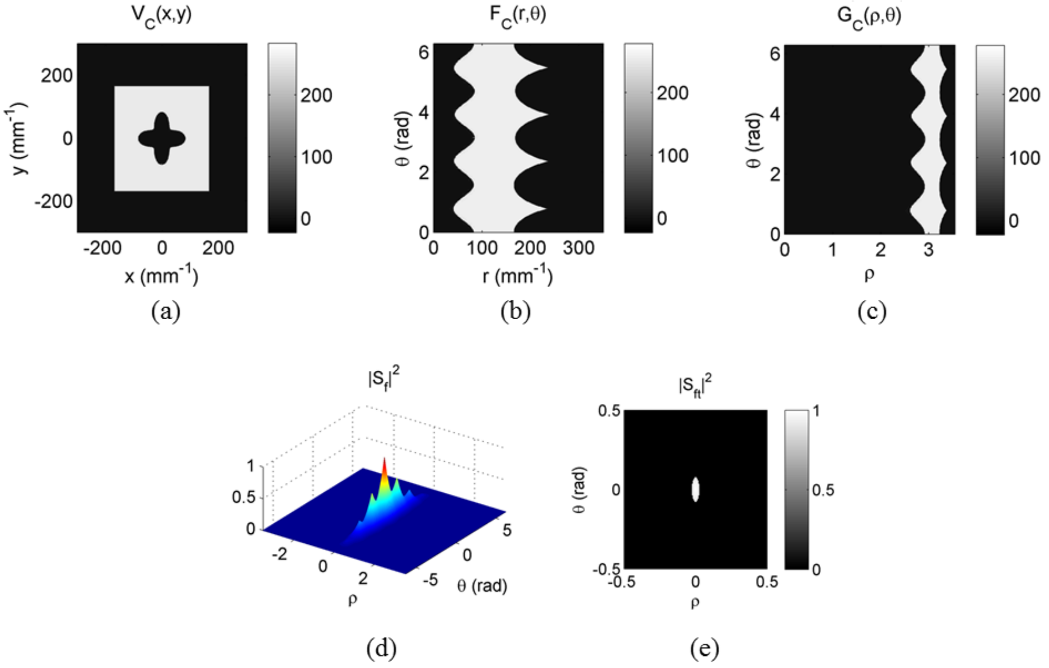
Figure 6.5:(a) We consider an artificial case where,$V_A(x,y)$ is the FT of an image (b) the corresponding polar distribution $F_A(r,\theta)$ (c) the corresponding log-polar distribution $G_A(\rho,\theta)$ (d) $V_B(x,y)$ is smaller in area than $V_A(x,y)$ by a factor of 4. (e) the corresponding polar distributions, $F_B(r,\theta)$. (f) the corresponding polar-logarithmic distribution, $G_B(\rho,\theta)$. (g) The final output signal $|S_f|^2$ when $G_A(\rho,\theta)$ and $G_B(\rho,\theta)$ are applied to the correlator. (h) The final signal after thresholding for $\sigma = 2$.

Figure 6.5(g) and Figure 6.5(h) illustrate the output signals of the HOC, if these PMT images [(G_A and G_B)] are applied as inputs. The output signal of the correlator, $|S_f|^2$ contains the two cross-correlation terms, T_3 and T_4. From eqn (**4.1**), T_3 represents the cross-correlation of the PMT images, $G_A(\rho,\theta) \odot G_B(\rho,\theta)$ and T_4 represents the anti cross-correlation of the same PMT images, but in reverse order : $G_B(\rho,\theta) \odot G_A(\rho,\theta)$. Thus T_3 and T_4 signals are shifted by equal amount but in opposite direction according to the shift in position between $G_A(\rho,\theta)$ and $G_B(\rho,\theta)$. As shown in Figure 6.5(g), the output now contains two peaks: a) the cross-correlation, $G_A(\rho,\theta) \odot G_B(\rho,\theta)$ at $\rho' = \log(2) = 0.3$ and $\theta' = 0$ and b) the anti cross-correlation $G_B(\rho,\theta) \odot G_A(\rho,\theta)$ at $\rho' = -\log(2) = -0.3$ and $\theta' = 0$, where (ρ',θ') are the co-ordinates in the correlation plane. Here, as expected, the magnitude of each peak is ~0.27, which is approximately one fourth of the peak value shown in Figure 6.4. In Figure 6.5(h), we apply a threshold of 0.25 to illustrate the peaks clearly**.**

It is important to consider the limit imposed on this process by the fact that the final signal contains both T_3 and T_4 terms. As we discussed in detail in section 4.2, there are essentially two distinct scenarios, characterized by the parameter η. For the scale invariant recognition, the value of this parameter is proportional to the scaling factor. The case shown in Figure 6.5 corresponds to $\eta > 1$, producing two peaks that are clearly resolved. We next consider a case where the scaling factor is small, corresponding to $\eta < 1$. In this case, $V_C(x,y)$ is scaled *down* by a small factor of $\sigma = 1.2$ compared to $V_A(x,y)$ as shown in Figure 6.6(a). The corresponding polar distribution and log-polar distributions

$F_C(r,\theta)$ and $G_C(\rho,\theta)$, are shown in Figure 6.6(b) and Figure 6.6(c), respectively. For this case, the polar-logarithmic distributions, $G_A(\rho,\theta)$, and $G_C(\rho,\theta)$ are shifted from one another by an amount log (1.2) ≈ 0.08 which corresponds to $\eta < 1$.

Figure 6.6(d) shows the final signal while $G_A(\rho,\theta)$ and $G_C(\rho,\theta)$ are applied as input to the HOC



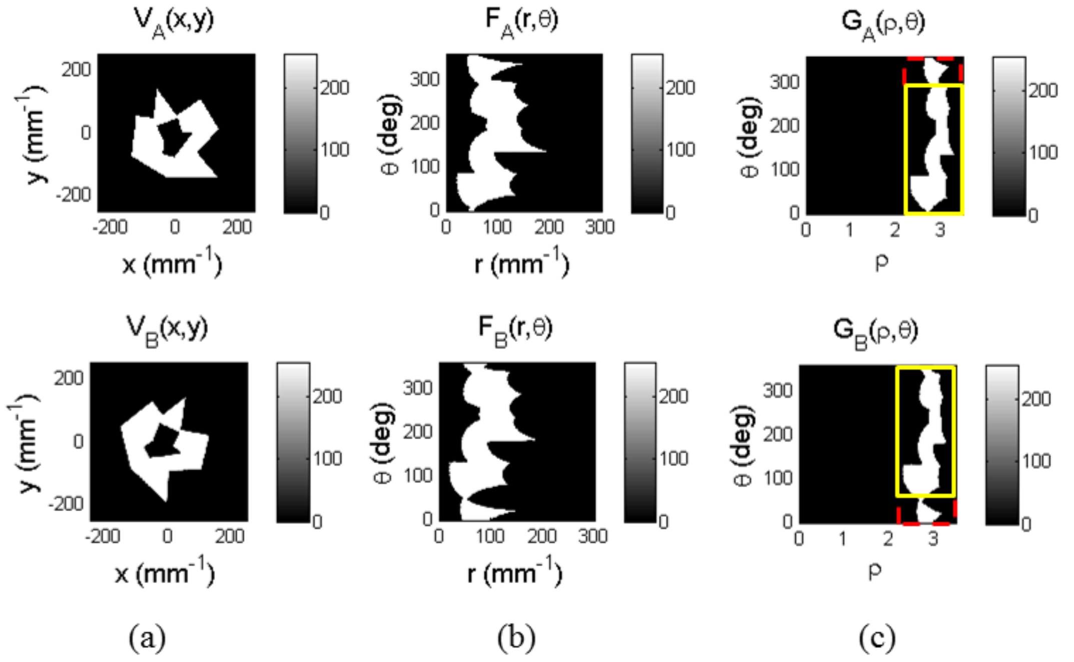
(a) (b) (c)

(d) (e)

Figure 6.6:(a) $V_C(x,y)$ is smaller in area than $V_A(x,y)$ by a factor of 1.44. (b) the corresponding polar distributions, $F_C(r,\theta)$ (c) the corresponding polar-logarithmic distribution is $G_C(\rho,\theta)$. (d) the final output signal $|S_f|^2$ when $G_A(\rho,\theta)$ and $G_C(\rho,\theta)$ are applied to HOC ; (d) The final signal after thresholding for $\sigma = 1.2$

architecture. In this case, the corresponding shift between $G_A(\rho,\theta)$ and $G_C(\rho,\theta)$ is log(1.2) ≈ 0.08. Hence, the two peaks have merged into a single peak of magnitude ~ 0.7. Note that in this case, while we are still able to determine the fact that the images are matched, the information about the relative scale between the images is lost. Figure 6.6(e) shows the final output signal after thresholding.

6.4. Rotation Invariant Image Detection using Idealized Image



(a) (b) (c)

Figure 6.7:(a) $V_A(x,y)$ is the amplitude of the FT of an arbitrary image and $V_B(x,y)$ represents the amplitude of the FT of the same image, but rotated by an angle of $\theta_0 = 45^0$. (b) the corresponding polar distributions, $F_B(r,\theta)$ and $F_A(r,\theta)$, now differ in the θ-direction (c) The dotted part of $G_B(\rho,\theta)$ is now shifted from the solid part of $G_A(\rho,\theta)$ in the θ-direction by an amount of $\theta_0 = 45^0$; the dotted part of $G_B(\rho,\theta)$ is now shifted from the solid part of $G_A(\rho,\theta)$ in the θ-direction by an amount of $-(2\pi - \theta_0) = -315^0$.

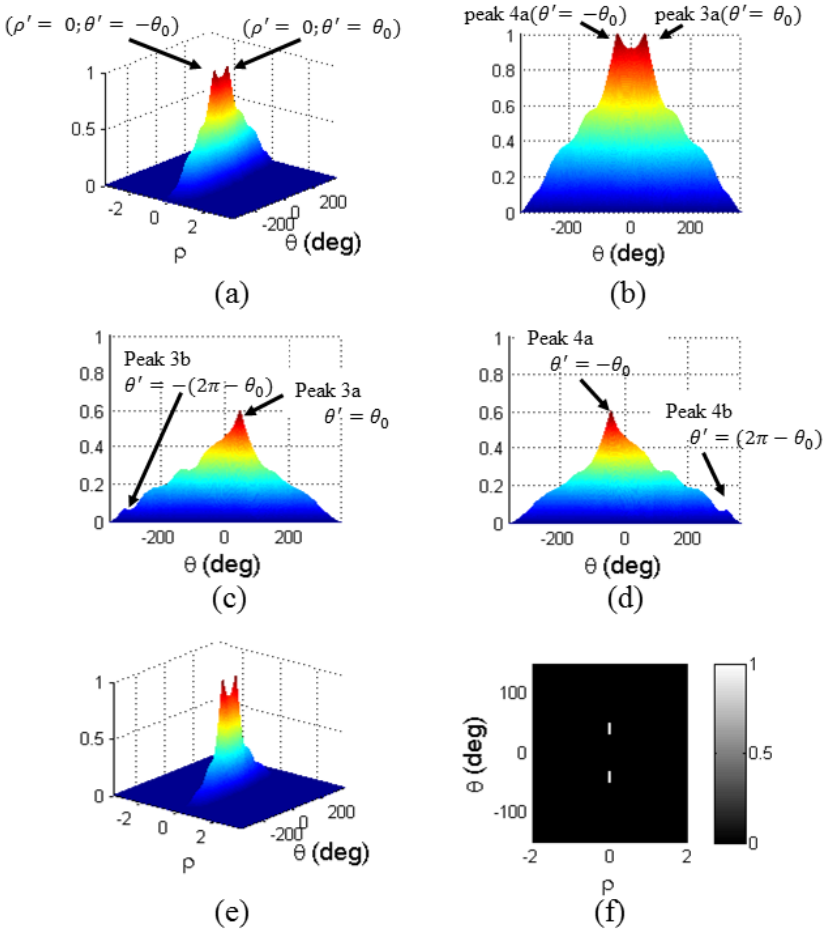


Figure 6.8: (a) The cross-correlation and anti cross correlation (T_3 + T_4) signal when $G_A(\rho,\theta)$ and $G_C(\rho,\theta)$ are applied to the HOC; (b) The cross-sectional view of the T_3 + T_4 signal showing two peaks shifted in the θ-direction by an amount of θ_0 and $-\theta_0$, which correspond to the cross-correlation terms $G_{A1}(\rho,\theta) \odot G_{B1}(\rho,\theta)$ and $G_{B1}(\rho,\theta) \odot G_{A1}(\rho,\theta)$, respectively (c) The cross-correlation signal T_3 which shows the two peaks corresponding to the cross-correlation $G_{A1}(\rho,\theta) \odot G_{B1}(\rho,\theta)$ and $G_{A2}(\rho,\theta) \odot G_{B2}(\rho,\theta)$. (d) *The anti cross-correlation signal T_4 which shows the two peaks corresponding to the cross-correlations $G_{B1}(\rho,\theta) \odot G_{A1}(\rho,\theta)$* and $G_{B2}(\rho,\theta) \odot G_{A2}(\rho,\theta)$. (e) The final output signal $|S_f|^2$. (f) The final output signal after thresholding.

Now, we consider the effect of rotation on the PMT images. Figure 6.7 shows such a case, where $V_A(x,y)$ is the amplitude of the FT of another arbitrary image and $V_B(x,y)$ represents the amplitude of the FT of the same image, but rotated by an angle of $\theta_0 = 45^0$. Thus, $V_B(x,y)$ is rotated by an angle of $\theta_0 = 45^0$ compared to $V_A(x,y)$. Here we have made use of the well-known fact that the process of FT preserves the angular information. Note first that the corresponding polar distributions (as shown in Figure 6.7(b)), $F_B(r,\theta)$, and $F_A(r,\theta)$ now differ in the θ-direction only and the pattern is shifted by $\theta_0 = 45^0$. Also note that the log-polar distributions are now still identical in shapes, except that $G_B(\rho,\theta)$ is now shifted from $G_A(\rho,\theta)$ in the θ-direction. It is particularly important to consider the shift in the θ-direction very carefully, since this coordinate is limited to a range of 0 to 2π (360^0). Specifically, the distribution along the θ-direction in G_A (G_B) can be broken into two parts: the part enclosed in the solid box in the upper (lower) part of Figure 6.7(c) can be denoted as $G_{A1}(G_{B1})$ and the part enclosed in the dotted box is denoted as G_{A2} (G_{B2}) . In $G_B(\rho,\theta)$, the G_{B1} part is shifted by an angle θ_0, while the GB_2 part is shifted by $-(2\pi-\theta_0)$, compared to G_{A1} & G_{A2}, respectively. Figure 6.8(a) shows the cross-correlation (T_3) and anti cross-correlation signal (T_4), when $G_A(\rho,\theta)$ & $G_B(\rho,\theta)$ are applied to the HOC. Figure 6.8(b) shows the

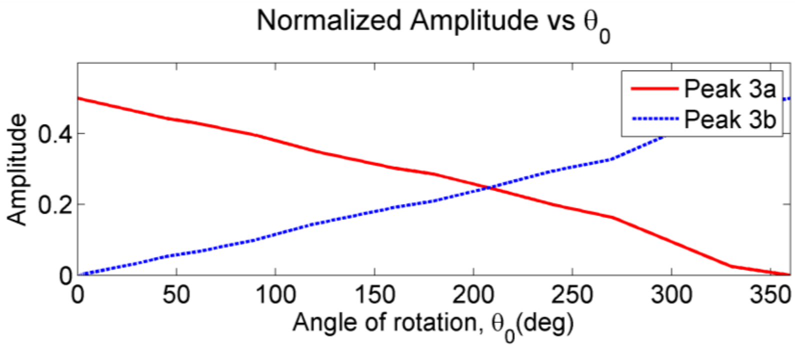
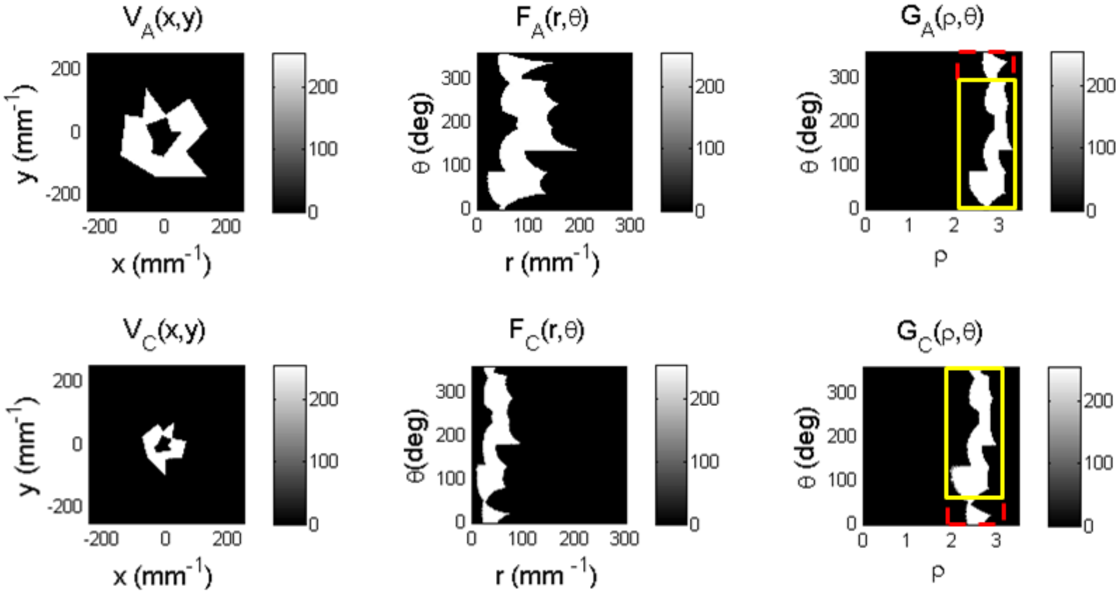


Figure 6.9: Normalized amplitude of peak 3a and 3b for different angle of rotation. See text for details.

cross-sectional view (as a function of θ for ρ = 0) of the cross-correlation (T_3) and anti cross- correlation signal (T_4). Consider first the T_3 term, which corresponds to $G_A(\rho,\theta) \odot G_B(\rho,\theta)$. Since there is no match between $G_{A1}(\rho,\theta)$ and $G_{B2}(\rho,\theta)$, and between $G_{A2}(\rho,\theta)$ and $G_{B1}(\rho,\theta)$ we get only two peaks: peak 3a corresponding to $G_{A1}(\rho,\theta) \odot G_{B1}(\rho,\theta)$ at $\rho' = 0$ and $\theta' = \theta_0 (= 45^0)$, and peak 3b corresponding to $G_{A2}(\rho,\theta) \odot G_{B2}(\rho,\theta)$ at $\rho' = 0$ and $\theta' = -(2\pi - \theta_0) = -315^0$. The signal corresponding to T_3, for $\rho = 0$, is plotted as a function of θ in Figure 6.8(c). As can be seen, peak 3a is prominent, while peak 3b is barely visible. This is due to the fact that θ_0 is small compared to $(2\pi - \theta_0)$ so that the energy contained in G_{A2} (G_{B2}) is smaller than that contained in G_{A1} (G_{B1}). Similarly, from the T_4 term, we get two other peaks: peak 4a corresponding to $G_{B1}(\rho,\theta) \odot G_{A1}(\rho,\theta)$ at $\rho' = 0$ and $\theta' = -\theta_0$; and peak 4b corresponding to $G_{B2}(\rho,\theta) \odot G_{A2}(\rho,\theta)$ at $\rho' = 0$ and $\theta' = (2\pi - \theta_0)$. These two peaks are shown in Figure 6.8(d), as a function of θ, for $\rho' = 0$. Again, we see that peak 4b is much smaller than peak 4a. The final output signal $|S_f|^2$ is shown in Figure 6.8(e). Here, we see peaks 3a and 4a clearly, while peaks 3b and 4b barely visible. However, the detection of just two peaks is enough to discern the relative angle of rotation. Figure 6.8(f) shows the final output signal $|S_f|^2$ after thresholding from where the locations of the peaks can is clearly visible. In Figure 6.9, we show how the relative amplitudes of peak 3a and 3b vary as a function of the rotation angle, θ_0. Similar behavior occurs (not shown) for peak 4a and peak 4b as well. The actual ratios of these peaks would, of course, depend on the details of the angular properties of the image.

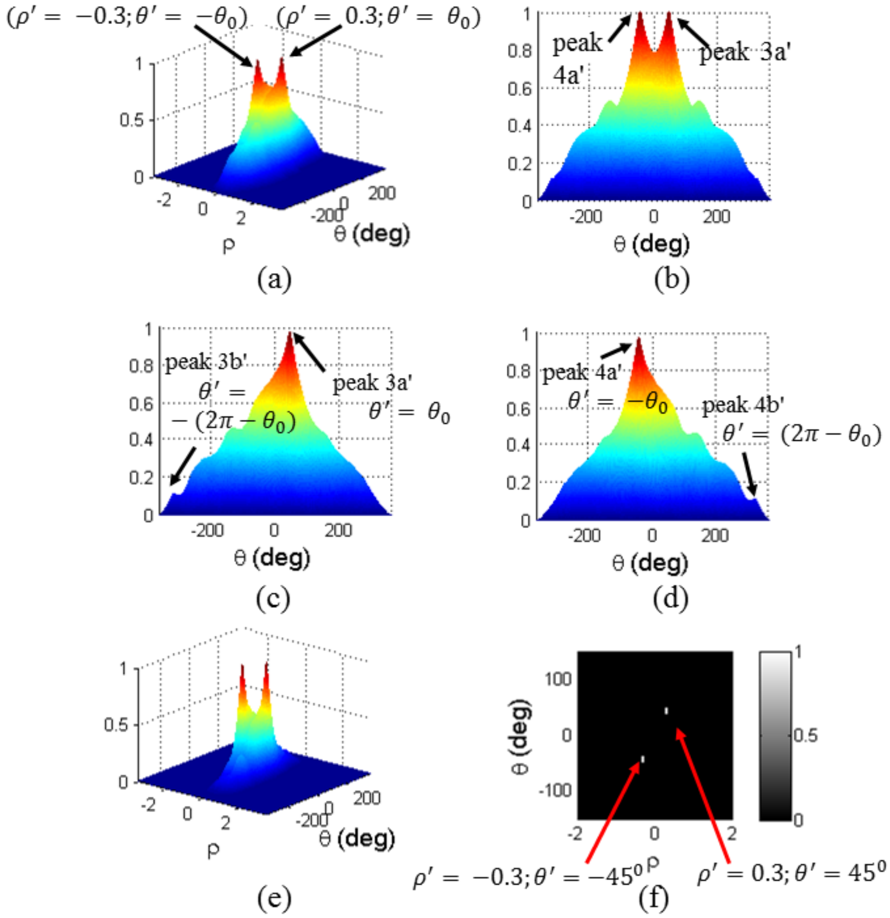
6.5. Scale & Rotation Invariant Image Detection using Idealized Image

Next, we consider the effect of rotation and scale change simultaneously. Figure 6.10 shows such a case, where $V_A(x,y)$ is the amplitude of the FT of an arbitrary image and $V_C(x,y)$ represents the amplitude of the FT of the same image, but scaled *up* by a linear factor of 2 and rotated by an angle of $\theta_0 = 45^0$. Note first that the corresponding polar distributions (as shown in Figure 6.10(b)),



(a) (b) (c)

Figure 6.10: (a) $V_C(x,y)$ is smaller in area by a factor of 4 than $V_A(x,y)$, and is rotated by 45^0. (b) the corresponding polar distributions, $F_C(r,\theta)$ and $F_A(r,\theta)$, now differ in both the r-direction and the θ-direction (c) $G_C(\rho,\theta)$ is now shifted from $G_A(\rho,\theta)$ in both the ρ-direction by an amount log(2) = 0.3 and the θ-direction by an amount of $\theta_0 = 45^0$.



(a) (b) (c) (d) (e) (f)

Figure 6.11: (a) The cross-correlation and anti cross-correlation (T_3+T_4) signal when $G_A(\rho,\theta)$ *and* $G_C(\rho,\theta)$ *are applied to the HOC, showing two peaks at* $(\rho' = 0.3; \theta' = 45^0)$ *and* $(\rho' = -0.3; \theta' = -45^0)$. *These correspond to the cross-correlation terms* $G_{A1}(\rho,\theta) \odot G_{C1}(\rho,\theta)$ *and* $G_{C1}(\rho,\theta) \odot G_{A1}(\rho,\theta)$, respectively . *(b) The cross-sectional view shows the peaks 3a' and 4a' (c) The cross-correlation signal T_3 which shows the two peaks corresponding to the cross-correlation* $G_{A1}(\rho,\theta) \odot G_{C1}(\rho,\theta)$ and $G_{A2}(\rho,\theta) \odot G_{C2}(\rho,\theta)$. (d) *The anti cross-correlation signal T_4 which shows the two peaks corresponding to the cross-correlations* $G_{C1}(\rho,\theta) \odot G_{A1}(\rho,\theta)$ and $G_{C2}(\rho,\theta) \odot G_{A2}(\rho,\theta)$. (e) The final output signal $|S_f|^2$. (f) The final output signal after thresholding.

$F_C(r,\theta)$, and $F_A(r,\theta)$ now differ in both the r-direction and the θ-direction. Note next that the log-polar distributions are now still identical in shapes, except that $G_C(\rho,\theta)$ is now shifted from $G_A(\rho,\theta)$ in the ρ-direction by an amount $\rho_0 = \log(2) = 0.3$ and the θ-direction by an amount $\theta_0 = 45^0$. Similar to the case of rotation change described above, the distribution along the θ-direction in G_A (G_C) can be broken into two parts: the part enclosed in the solid box in the upper (lower) part of Figure 6.10(c) can be denoted as $G_{A1}(G_{C1})$ and the part enclosed. in the dotted box is denoted as G_{A2} (G_{C2}) . In $G_C(\rho,\theta)$ the G_{C1} part is shifted by an angle θ_0, while the G_{C2} part is shifted by $-(2\pi-\theta_0)$, compared to G_{A1} and G_{A2} , respectively.

Figure 6.11(a) shows the cross-correlation and anti cross-correlation (T_3+T_4) signal, when $G_A(\rho,\theta)$ and $G_C(\rho,\theta)$ are applied to the HOC. The output contains several peaks, corresponding to the cross-correlation and anti cross correlation terms, T_3 and T_4, of eqn.(**3.9**). Consider first the T_3 term, which corresponds to $G_A(\rho,\theta) \odot G_C(\rho,\theta)$. Since there is no match between $G_{A1}(\rho,\theta)$ and $G_{C2}(\rho,\theta)$, and between $G_{A2}(\rho,\theta)$ and $G_{C1}(\rho,\theta)$, we get only two strong peaks: peak 3a' corresponding to $G_{A1}(\rho,\theta) \odot G_{C1}(\rho,\theta)$ at $\rho' = \log(2) = 0.3$ and $\theta' = \theta_0$ ($= 45^0$); and peak 3b' corresponding to $G_{A2}(\rho,\theta) \odot G_{C2}(\rho,\theta)$ at $\rho' = 0.3$ and $\theta' = -(2\pi - \theta_0)$ (shown in Figure 6.11(c)). As can be seen, peak 3a' is prominent, while peak 3b' is barely visible. Similarly, from the T_4 term, we get two other peaks: peak 4a' corresponds to $G_{C1}(\rho,\theta) \odot G_{A1}(\rho,\theta)$ at $\rho' = -0.3$ and $\theta' = -\theta_0$; and peak 4b' corresponds to $G_{C2}(\rho,\theta) \odot G_{A2}(\rho,\theta)$ at $\rho' = -0.3$ and $\theta' = (2\pi - \theta_0)$ (shown in Figure 6.11(d)). Figure 6.11(e) shows the final output signal $|S_f|^2$ and Figure 6.11(f) shows the output signal after applying a threshold of 0.9. From Figure 6.11(f), it is obvious that the positions of the peak of the cross-correlation signal (T_3) corresponds to rotation and scale change between two images.

Chapter 7 Simulation Results of Scale and Rotation Invariant HOC

7.1 Introduction

In chapter 6, we have shown that by incorporating PMT, the proposed HOC architecture can achieve scale and rotation invariance, in addition to the shift invariance feature. So far, we have considered artificial cases where the FT of the object or reference image has a distinct hole in the center. As discussed earlier this corresponds to an unrealistic situation where the image must have both positive and negative amplitudes. In this section, we consider real world scenarios where the image has non-negative values everywhere. Since such an image has a non-zero average value, its FT cannot have a hole in the center. By cutting a hole of a suitable radius in the center of the FT of the image, we produce an effective image, which is no longer positive definite, and is thus compatible with the PMT process.

7.2 Simulation Results and Discussion

In Figure 7.1(a), we consider two images, $U_1(x', y')$ and $U_2(x', y')$, where $U_2(x', y')$ is scaled down by a linear factor of 2 with respect to $U_1(x', y')$. The corresponding magnitudes of FTs, $V_1(x, y)$ and $V_2(x, y)$ are shown in fig 16(b), where $V_2(x, y)$ is scaled up by a factor of 2 compared to $V_1(x, y)$ in each dimension. Here, we take the magnitude of the FTs to get rid of any shift information. In fig 16(c), a hole of radius r_0 is created in the center of $V_1(x, y)$ and $V_2(x, y)$,

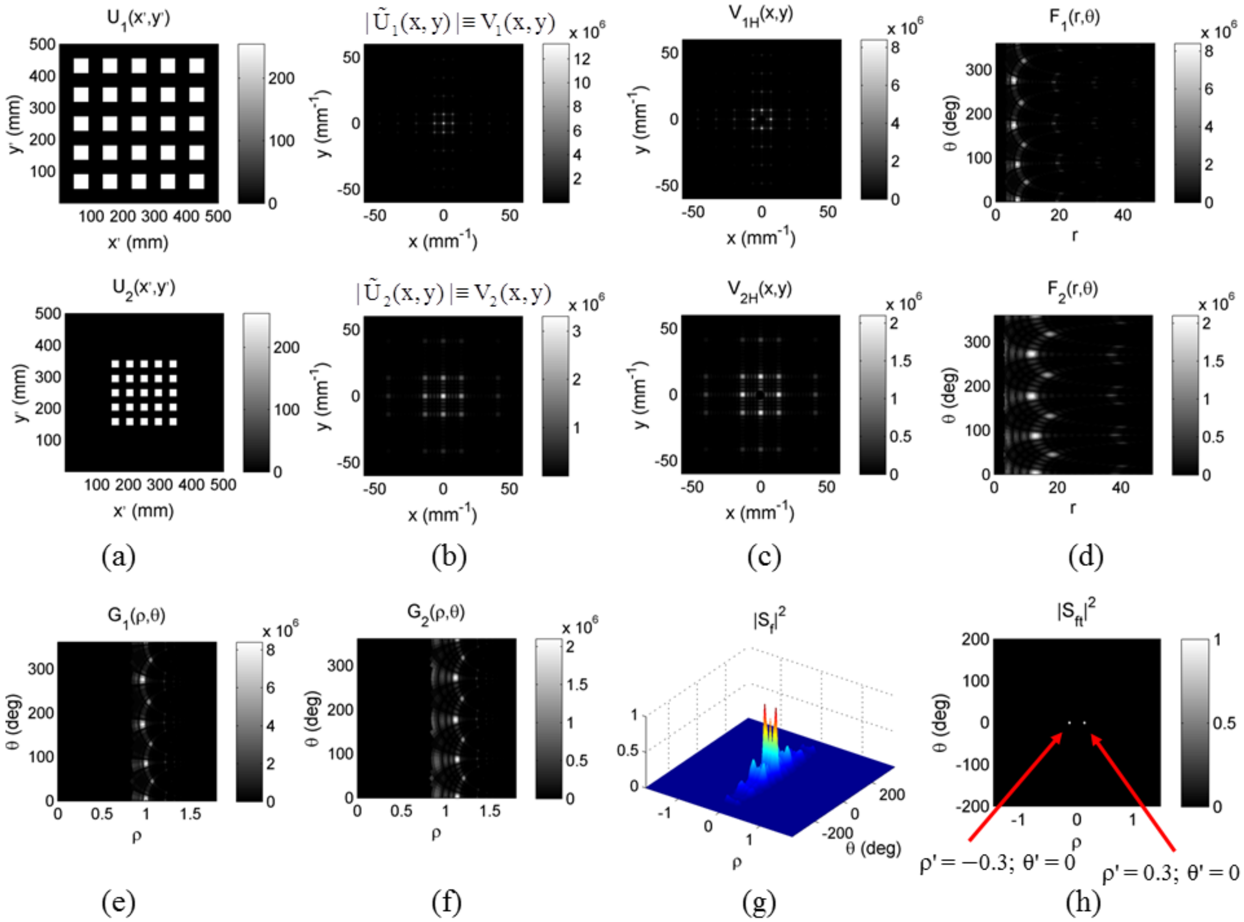


Figure 7.1: (a) we consider two images $U_1(x',y')$ and $U_2(x',y')$, where $U_2(x',y')$ is smaller in area than $U_1(x',y')$ by a factor of 4; (b) The corresponding magnitudes of FTs are denoted as $V_1(x,y)$ and $V_2(x,y)$, (c) A hole of radius r_0 is created in the center of $V_1(x,y)$ and $V_2(x,y)$ and the resulting function are denoted as $V_{1H}(x,y)$ and $V_{2H}(x,y)$, respectively; (d) The corresponding polar distributions, $F_1(r,\theta)$ and $F_2(r,\theta)$ (e) The polar-logarithmic distributions, $G_1(\rho,\theta)$ and $G_2(\rho,\theta)$, are identical in shapes, except for a shift in the ρ-direction [equaling the logarithm of the scale factor: $\log(2) \simeq 0.3$] (g) The final output signal, $|S_f|^2$ of the HOC when $U_1(x',y')$ and $U_2(x',y')$ are converted to the log-polar domain and then act as inputs to the HOC.(g) After thresholding we get two peaks corresponding to the two cross-correlation terms (T_3 and T_4).

which are denoted as $V_{1H}(x,y)$ and $V_{2H}(x,y)$, respectively. Fig 16(d) shows the corresponding polar distributions, $F_1(r,\theta)$ and $F_2(r,\theta)$.

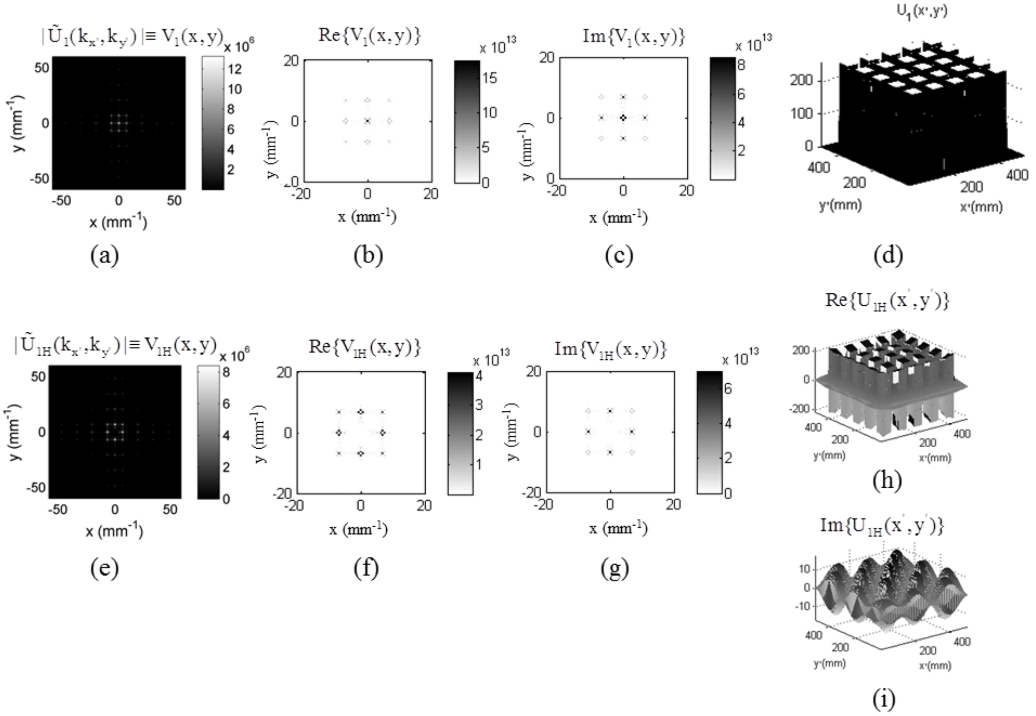


Figure 7.2: Illustration of the fact that cutting a hole of certain radius in the center of the FT of the image does not change the image significantly. See text for details. [Note that, in fig 16(b), (c), (f) and (g), the color has been inverted for clear visualization]

As shown in Figure 7.1(e), the polar-logarithmic distributions are nearly identical in shape, except that $G_2(\rho,\theta)$ is now shifted from $G_1(\rho,\theta)$ in the ρ-direction by an amount of log(2) = 0.3. The PMT processed images $G_1(\rho,\theta)$ and $G_2(\rho,\theta)$, are now inputs to the HOC architecture. Figure 7.1(f) shows the final output signals $|S_f|^2$ of the HOC where the peaks of the cross-correlation signals, T_3 and T_4 are shifted from the center in the ρ direction by an amount equaling log (σ) and −log (σ), respectively. Figure 7.1(g) shows that thresholding gives a clear view of the location of these peaks, from which we can determine the relative scaling factor. When a hole is cut in the FT of an image, the process is equivalent to the use of a modified image. We use one of the images considered above $U_1(x',y')$, above to illustrate what this modified image looks like. Before preceding it is instructive to document clearly the notations we have employed as shown in table 1.

Table 7.1: Summary of definitions of various transforms

Symbol	Meaning
$U_1(x',y')$	Original image; (x',y') are spatial coordinates, with units of millimeter (mm)
$\tilde{U}_1(k_{x'},k_{y'})$	FT of the original image; $(k_{x'},k_{y'})$ are wave number coordinates, with units of mm^{-1}
$\tilde{U}_1(x,y)$	Same as $\tilde{U}_1(k_{x'},k_{y'})$, except that we have defined $x \equiv k_{x'}$, $y \equiv k_{y'}$; Thus, (x,y) are wave number coordinates, with units of mm^{-1}. This redefinition is for convenience only.
$V_1(x,y)$	This is the magnitude of the FT of original image.
$U_{1H}(x',y')$	Image resulting from cutting a hole in the FT.
$\tilde{U}_{1H}(k_{x'},k_{y'})$	FT of $U_{1H}(x',y')$; Again, (x',y') are spatial coordinates.
$\tilde{U}_{1H}(x,y)$	Same as $\tilde{U}_{1H}(k_{x'},k_{y'})$ with the definition of $x \equiv k_{x'}$, $y \equiv k_{v'}$.
$V_{1H}(x,y)$	This is the magnitude of the FT of the modified image.

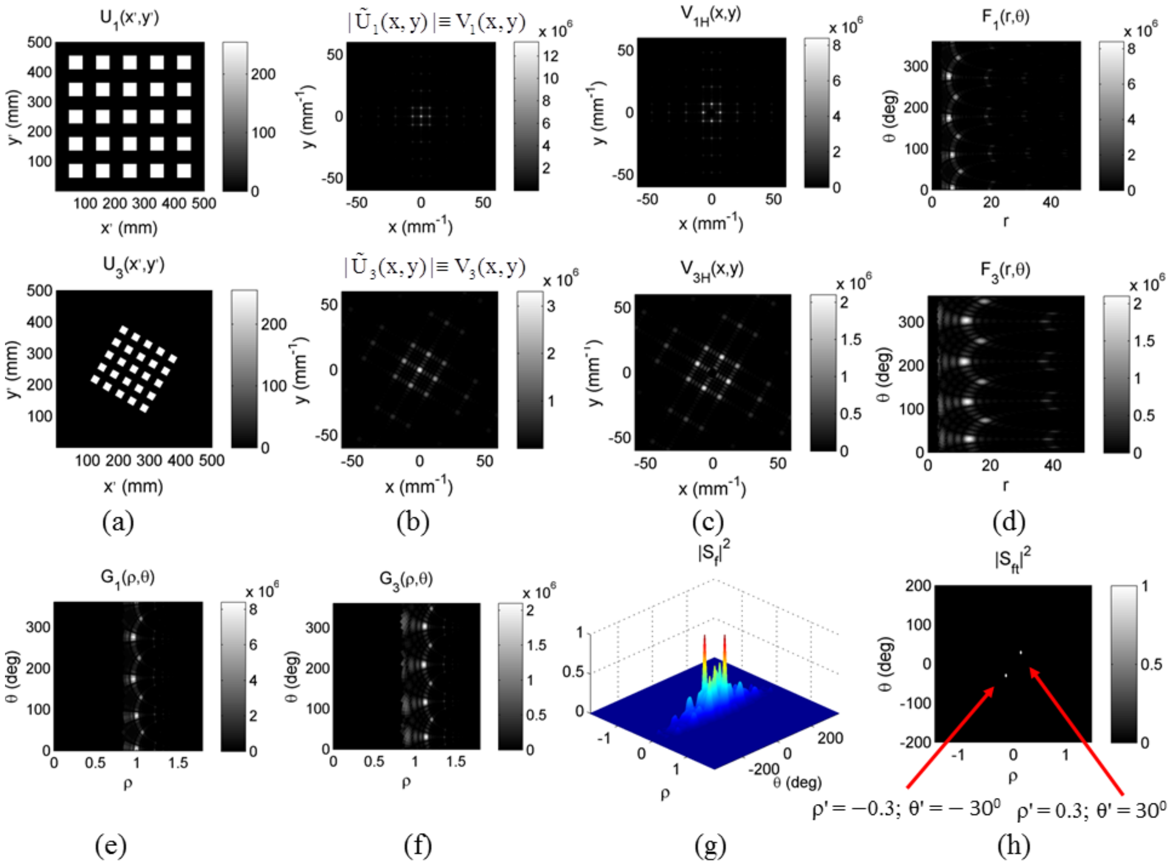


Figure 7.3: Simulation result showing the scale and rotation invariant object detection using the HOC

In the correlation process described in Figure 7.1, we made use of $V_{1H}(x,y)$ and $V_{2H}(x,y)$. The corresponding modified images are $U_{1H}(x',y')$ and $U_{2H}(x',y')$. As an example, we show below, in steps, how to determine $U_{1H}(x',y')$, and explain its shape. Figure 7.2(a) shows $V_1(x,y)$, the FT of the original image, while Figure 7.2(e) shows $V_{1H}(x,y)$, the FT of the modified image. However, in order to reconstruct the corresponding images, we require the complex FT's. These are illustrated next. Figure 7.2(b) and (c) show the real and imaginary parts respectively of $\tilde{U}_1(k_{x'},k_{y'}) = \tilde{U}_1(x,y)$, the FT of the original image. Inverse FT of $\tilde{U}_1(x,y)$, yields the original image, $U_1(x',y')$, which is shown in axonometric view in Figure 7.2(d). Note that the original image is real only. Similarly Figure 7.2(f) and (g) show the real and imaginary parts, respectively, of $\tilde{U}_{1H}(k_{x'},k_{y'}) = \tilde{U}_{1H}(x,y)$, the FT of the modified image. Inverse FT of $\tilde{U}_{1H}(x,y)$, yields the modified image, $U_{1H}(x',y')$. Note that $U_{1H}(x',y')$ is complex, as a result of the hole-cutting process. Real and imaginary parts of $U_{1H}(x',y')$ are shown in Figure 7.2(h) and (i) respectively. In Figure 7.3(a), we consider two images, $U_1(x',y')$ and $U_3(x',y')$, where $U_3(x',y')$ is scaled down by a linear factor of 2 and also rotated with respect to $U_1(x',y')$ by an angle of $\theta_0 = 30^0$. The magnitude of the FT of $U_3(x',y')$, denoted as $V_3(x,y)$, is also rotated by an angle of $\theta_0 = 30^0$ and the area is enlarged by a factor of 4, as shown in Figure 7.3(b). In Figure 7.3(c), a hole of radius r_0 is created in the center of each of $V_1(x,y)$ and $V_3(x,y)$, producing functions denoted as $V_{1H}(x,y)$ and $V_{3H}(x,y)$, respectively. The corresponding polar distribution, $F_1(r,\theta)$ and $F_3(r,\theta)$, are shown in Figure 7.3(d). As shown in Figure 7.3(e), the polar-logarithmic distributions are still identical in shape, except that $G_3(\rho,\theta)$, is shifted from $G_1(\rho,\theta)$, in both θ-direction and ρ-direction. In the process described above, two original images, $U_1(x',y')$ and $U_3(x',y')$, are converted to PMT images, $G_1(\rho,\theta)$, and $G_3(\rho,\theta)$, which act as inputs to the HOC architecture. Figure 7.3(f) shows the final results of the HOC architecture where the cross correlation signal T_3 has two peaks at positions ($\rho' = \log(\sigma)$, $\theta' = \theta_0$) and ($\rho' = \log(\sigma)$, $\theta' = -(2\pi - \theta_0)$) and the anti cross-correlation signal, T_4 also has two peaks at positions ($\rho' = -\log(\sigma)$, $\theta' = -\theta_0$) and ($\rho' = -\log(\sigma)$, $\theta' = (2\pi - \theta_0)$). As mentioned in section 3, the peaks at positions ($\rho' = \log(\sigma)$, $\theta' = -(2\pi - \theta_0)$) and ($\rho' = -\log(\sigma)$, θ'

$(2\pi - \theta_0)$) are very small compared to other peaks so that they are barely visible in the final output signal. Figure 7.3(g) shows that after thresholding the peaks of the cross-correlation signals at positions ($\rho' = \log(\sigma), \theta' = \theta_0$) and ($\rho' = -\log(\sigma)$, $\theta' = -\theta_0$) are clearly visible. From the location of the peaks, we can infer that the objects are rotated with respect to each other by an angle of 30^0 and also scaled down by a factor of 2.

Consider now a general case where the two images are shifted, rotated and scaled with respect to each other. In that case, the shift information can be retrieved by a two step process. First, the PMT versions of the images are used to determine the degree of rotation and scaling. This information is now used to scale and rotate the query image so that it is now not rotated, nor scaled with respect to the reference image. We now repeat the correlation process using the actual images rather than PMT images. This correlation will reveal the amount of relative shift (magnitude and direction) between the two images.

7.3 Conclusion

The proposed architecture of the HOC is very different, requiring many intermediate steps, servos and post processing. Furthermore, the output signals are also different (e.g., it contains a cross-correlation term and an anti cross-correlation term, but no convolution nor dc term).Thus, it may not be a priori obvious whether the HOC can also perform scale and rotation invariant correlation PMT. Here, we show how, indeed, it is possible to use PMT to achieve scale and rotation invariant image recognition with the HOC architecture. Furthermore, we identify some limitations of using PMT in the CHC architecture proposed previously, and show how to overcome these constraints by proper pre-processing of images for both CHC and HOC.

Chapter 8 Multiple Object Detection Using th HOC Arhitecture

8.1 Introduction

In reference 44, we showed how to recognize a single object using the HOC architecture in a shift invariant manner. In this thesis so far, we have shown how to recognize a single object in a shift, scale and rotation invariant manner. However, there are potential scenarios where the query field may contain multiple matches to the reference object. In this section, we describe how to achieve distinct detection of these multiple matches. We consider two different scenarios.

First, we consider the case where the multiple images in the query field are only shifted, without any scale change and rotation angle. In this case, the architecture needed does not employ the PMT process. However, it is somewhat different from the approach used for detecting multiple (unscaled and unrotated) matches using a conventional holographic correlator, because of the fact that the HOC produces both cross-correlation and anti cross-correlation signal. Second, we consider the case where the multiple images in the query field have potentially distinct values of shift, scale factor and rotation angle. In this case, the PMT process has to be employed. However, since the PMT process eliminates the shift information, multiple object recognition in this case requires a substantially different architecture.

8.2 Multiple objects detection for shifted images without rotation and scale change:

Detection of multiple objects using the HOC architecture without rotation and scale change is similar to single object detection under the same scenario (i.e., without rotation and scale change), in that it does not require using the PMT process. However, a potential complication arises due to the presence of both T_3 (cross-correlation) and T_4 (anti cross-correlation) terms. Specifically, in the presence of multiple matches, it becomes difficult to determine whether a peak corresponds to T_3 of a given matched image, or T_4 of another matched image, for example. This can be circumvented by applying the following technique. Assume first that the reference image is represented by a grid of N X N points. In contrast, we confine the query image, potentially containing multiple objects, to a grid of only N/2 x N points. We now map this query image to a grid of NXN points, thus producing a final query image where half the image is blank. Consider a situation where the blank half of the query image is on the left side. It is easy to see that, after carrying out the correlation process in the manner described in chapter 3, the peaks representing all the T_3's (corresponding to multiple matches) will appear on the right side of the final signal plane, while all the T_4's will appear on the left side of the final signal plane, thus avoiding the potential ambiguities between the T_3's and the T_4's mentioned above.

Figure 8.1 illustrates the results of a simulation for multiple object detection with the HOC using this approach. Figure 8.1(a) shows the reference image, Im_A, plotted on an N X N grid. Figure 8.1(b) show the query image, Im_B, which contains three images that match the reference (denoted as B,C and D), and one that does not (denoted as A). Note that these four images are confined to the right half

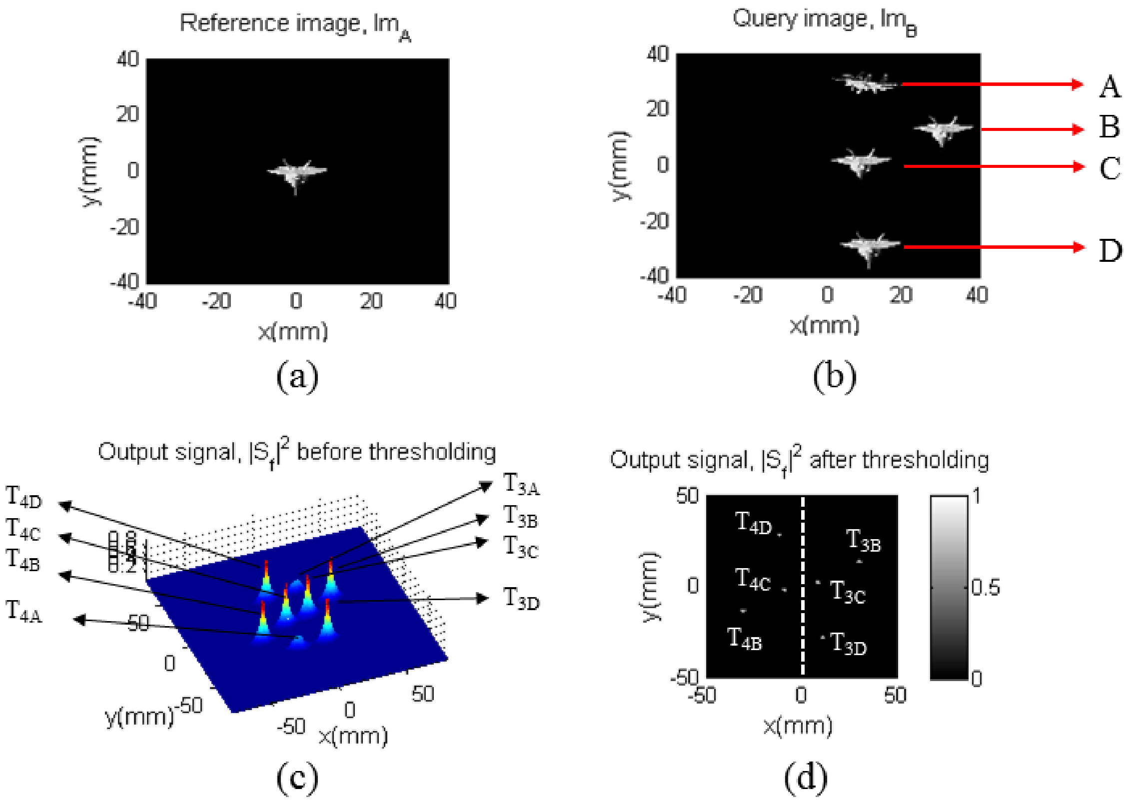


Figure 8.1: Illustration of multiple object detection without rotation and scale change using the HOC architecture. See text for details.

plane only, leaving the left half blank. Figure 8.1(c) shows the output signal, $|S_f|^2$ where we can see six sharp peaks corresponding to cross-correlation signals (T_{3B}, T_{3C} and T_{3D}) and anti-cross correlation signals (T_{4B}, T_{4C} and T_{4D}) of the three matched images. The other two signals in the output plane with lower peaks correspond to the cross-correlation (T_{3A}) and anti cross-correlation (T_{4A}) signals for the unmatched case. After thresholding (as shown in Figure 8.1(d)), the final signals are clearly visible from which we can infer that three matches are found. In addition, the distances of the peaks from the center reveal the locations of the matched images.

This approach of leaving a blank space also has an added advantage, in that it ensures that the overlap parameter, η (defined in section 4.2) is never less than unity. This can be seen clearly by considering image C, which is located at the left edge of the right half plane. Since it is contained fully in the right half plane, the distance between this and the reference images is $|\overrightarrow{\rho_m}|$, which corresponds to $\eta = 1$. All other images that are further away from the boundary between the left and right half planes would thus have a value of $\eta > 1$. Therefore, the cross-correlation and anti-cross correlation signals will be clearly resolved for all images.

Finally, we note that the use of the rectangular field (N/2 X N) in confining the query image may be inconvenient in some situations, especially if the camera used in acquiring the query image has an image field that is a square. This can be circumvented by confining the query image to a square field with N/2 x N/2 points, leaving the other three quadrants blank.

8.3 Multiple objects detection in shift, scale and rotation invariant manner using the HOC architecture:

Next, we consider a situation where the query field contains multiple replicas of the reference image, but each with potentially a different position, a different scale factor, and a different angular orientation. Obviously, this case would require the use of the PMT process. However, the PMT process loses the information about the relative position of any image once the phase information in the FT is eliminated by measuring the magnitude of the FT. Thus, the magnitude of the FT's of each of the matched images in the query field will overlap with one another, making it impossible to find any matches.

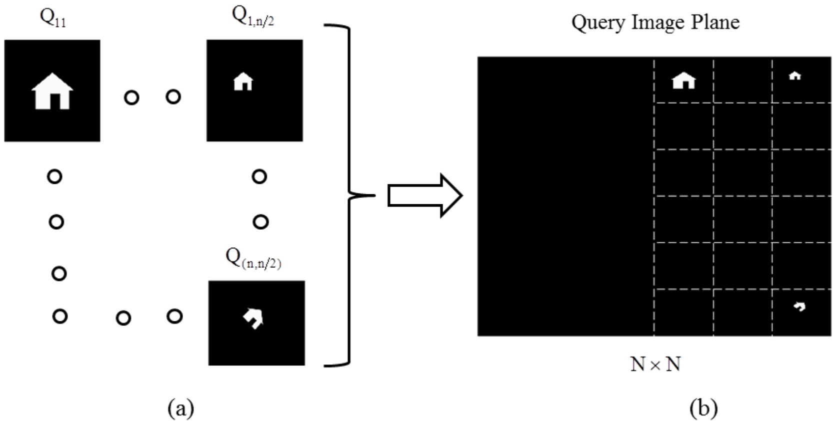


Figure 8.2: Illustration of the process of mapping multiple query objects each with potentially a different position, a different scale factor, and a different angular orientation, to the right half of the Query image plane which has a grid size of N x N. See text for details.

To overcome this problem, we propose the approach illustrated schematically in Figure 8.2 and Figure 8.3. The situation of interest here is as follows. We assume that we have one reference image, and L ≡ n/2 x n captured query images (n is an integer), each of which is a potentially shifted, scaled and rotated replica of the reference image. In principle, we can employ the PMT enhanced HOC process L times. The goal here is to carry out these L correlations simultaneously.

To start, we fit each captured image into a grid of (N/n) x (N/n), where N x N is the grid size for the reference image. We then fit these images into the right half plane of an N x N grid. This is illustrated in Figure 8.2. Next, we use an SLM to convert the query image to the optical domain. However, instead of sending the whole query image at once, we send to the SLM only one of the small grids (of size N/n x N/n) of the right half plane with all the other small grids being dark. This is illustrated schematically shown on the left edge of the Figure 8.3, for the first row and n/2-th column of the right half plane, for example. The image in this grid is denoted as $Q_{1,n/2}$. The lens, the FPA and the FPGA, as shown in the rest of Figure 8.3, are used to produce the corresponding PMT image, denoted as $G_{1,n/2}(\rho,\theta)$. This process is repeated $n^2/2$ times, which is the number of small grids containing images, without changing the positions of the SLM, the lens and the FPA. The $n^2/2$ numbers of PMT images produced and stored in the FPGA are then mapped to a corresponding set of small grids, which in turn is sent to the final SLM (SLM-3) for detecting cross-correlation and anti cross-correlation signals.

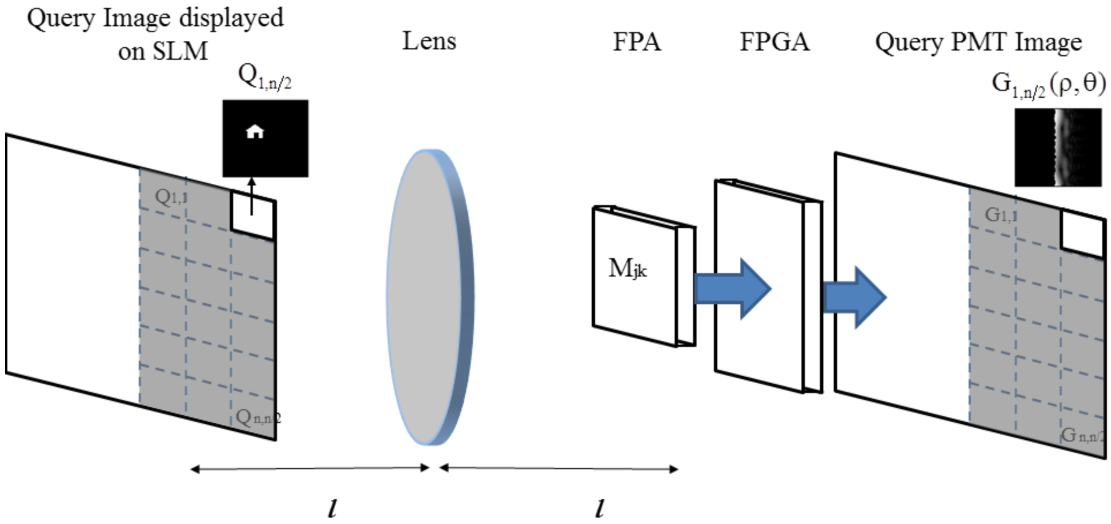


Figure 8.3: The process of Fourier transforming a query image plane with multiple images using an SLM and a lens. See text for details.

In Figure 8.4, we show results of numerical simulations used to illustrate the process described above. For simplicity, we have used an artificial reference image that has a clear hole in its FT's, similar to that shown earlier in Figure 6.7. The FT of the reference image is shown in Figure 8.4(a), denoted as $F_{00,}$ and the FTs of multiple query objects are shown in Figure 8.4(b), which are denoted as F_{ij} (i = row number; j = column number). For example, F_{42} is the image shown on the bottom right corner. The left half of the query plane is kept blank for the reason described above. As can be seen in Figure 8.4(b), F_{21} and F_{42} are similar to F_{00}; F_{22} and F_{31} are rotated from F_{00} by an angle of $\theta_0 = 30^0$; F_{11} and F_{32} are scaled down from F_{00} by a factor of $\sigma = 2$; F_{12} and F_{41} are scaled down by a factor of $\sigma = 2$ and also rotated by an angle of $\theta_0 = 30^0$ from F_{00}.The corresponding PMT images, $G_1(\rho,\theta)$ and $G_2(\rho,\theta)$ are shown in Figure 8.4(c) and Figure 8.4(d), respectively[46]. Figure 8.4(e) shows the final signal $|S_f|^2$ after thresholding. The right half side shows the cross-correlation signals and the left side shows the anti-cross- correlation signals. The red dots in Figure 8.4(e) correspond to the auto-correlation of the reference PMT image and the white dots on the right half plane represent the corresponding cross-correlation signals. From

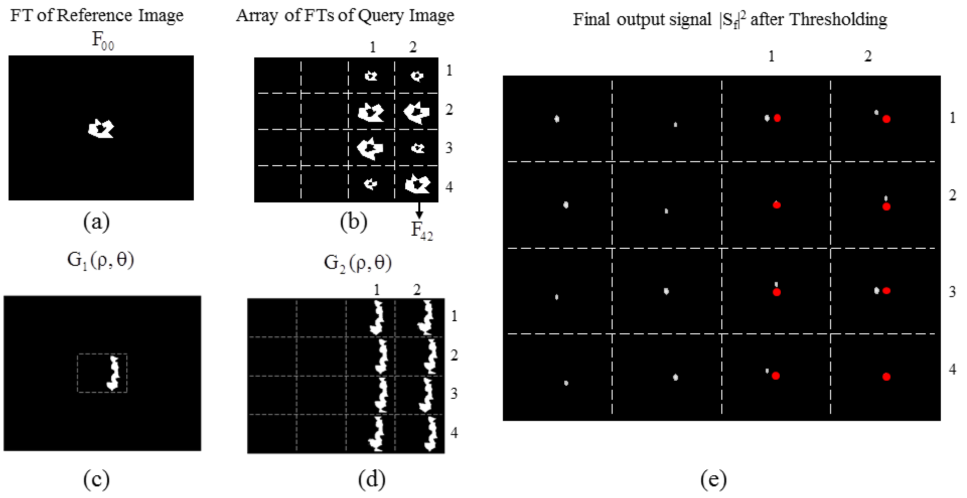


Figure 8.4: (a) FT of the reference image (b) The array of FTs of multiple objects where one side is intentionally left blank. (c) PMT of the reference image (d) PMT of the array of query images (e) Final results from the HOC. We have to consider only the right half of this plane. See text for details.

the distance between the red dot and the white dot in a given box, we can infer the scale and rotation change between the reference image and the query image.

8.4 Conclusion

We have shown that our proposed HOC architecture, which correlates images using SLMs, detectors and FPGAs, is capable of detecting objects in a scale and rotation invariant manner, along with the shift invariance feature, by incorporating PMT. For realistic images, we cut out a small circle at the center of the fourier transform domain, as required for PMT, and illustrate how this process corresponds to correlating images with real and imaginary parts. Furthermore, we showed how to carry out shift, rotation and scale invariant detection of multiple matching objects simultaneously, a process previously thought to be incompatible with PMT based correlators. We presented results of numerical simulations to validate the concepts.

Chapter 9 Translation Invariant Temporal Correlator

9.1 Introduction

The Temporal Correlator (TC), shift invariant in time, can recognize rapidly a query pulse that may be present in a reference pulse train, and determine the temporal location of the query pulse. To illustrate the Translation Invariant Temporal Correlator (TI-TC), consider a medium that is inhomogeneously broadened, meaning that the atoms inside the medium have a range of resonant frequencies. In such a medium, one can record a temporal data sequence by using a uniform recording pulse separated in time. In general, modeling the TI-TC requires determining the non-linear response of a resonant medium, which can be found by solving numerically the quantum mechanical density matrix equations, or simply the Schrodinger Equation (SE) if dissipative processes can be ignored. However, this approach is time consuming and cumbersome. Here, along with modeling the TI-TC using the SE for the temporal evolution of atoms excited by optical fields, we show that the response of the TI-TC can be determined by modeling the response of the interaction medium as a simple transfer function in the Spatio-Temporal Fourier Transform (STFT) domain. We explain the physical origin of analytical model, and then establish the validity of this model by comparing its prediction with that determined via the quantum mechanical dynamics. We show that the analytical model is gives idetical results to the numerical model but at a speed that is many orders of magnitude faster than solving the SE.

9.2 Architecture of the Translation Invariant Temporal Correlator (TI-TC)

To illustrate the Translation Invariant Temporal Correlator (TI-TC), consider a medium that is inhomogeneously broadened, meaning that the atoms inside the medium have a range of resonant frequencies. In such a medium, one can record a temporal data sequence by using a uniform recording pulse separated in time[25,26 ,47,48].We start by applying a write pulse. This is followed by the query data stream, with a certain time lag. The spectral-domain interference between the

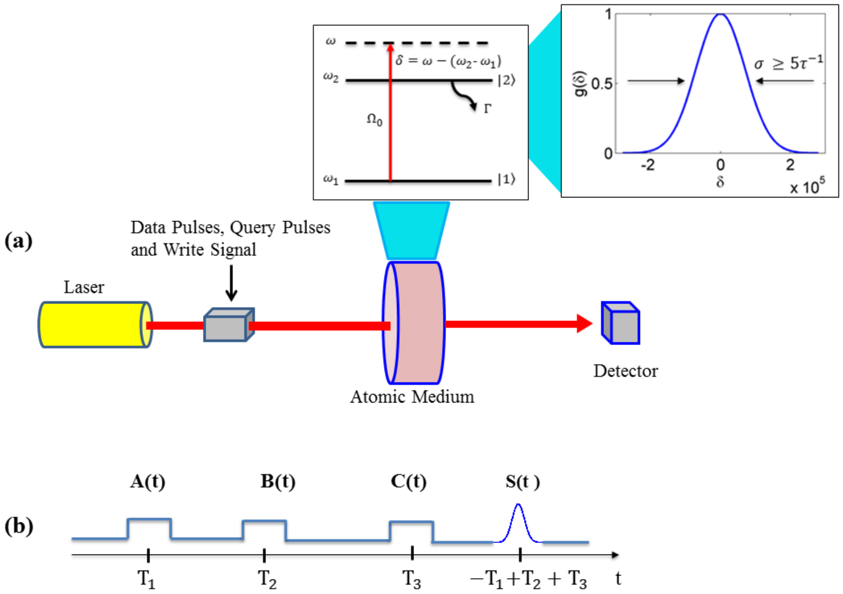


Figure 9.1: Schematic illustration of a temporal correlator using inhomogeneously broadened atomic media. After applying a short, uniform pulse of the writing beam, a(t) followed by the query, b(t) and reference data, c(t) stream, a correlation peak, s(t) is observed in a temporal shift invariant manner (A.U. = Arbitrary Unit)

writing beam and the query data stream (which can be viewed also as a manifestation of the Alford and Gold Effect[23,28 ,49,50]) is encoded in the coherence produced in the atomic medium. When the reference data stream is applied to this system, a correlation peak is observed in a temporally shift invariant manner. The temporal correlation process is illustrated schematically in Figure 9.1. Figure 9.1(a) shows the simple architecture of the temporal correlator and Figure 9.1(b) shows the sequence of pulses applied in an inhomogeneously broadened atomic medium. The numerical and analytical modeling of such a system is discussed later in this chapter. Here, we have used an idealized, decay-free two level system of atoms with an inhomogeneous broadening that is larger than the inverse of the temporal resolution of the data stream. It can be shown that an off-resonant excitation in a three-level system is equivalent to this model[51,52] which has been discussed in details in Appendix E.

There is a close analogy between this process and the holographic spatial correlator. The initial data stream corresponds to the reference image, and the recording pulse corresponds to the write beam, with the time separation being analogous to the angle between the reference image and the write beam. The matching pulse corresponds to the query image, and the output pulse corresponds to the correlation. Just as in the case of spatial holography, the process is translation invariant, meaning that if the matching pulse is shifted in time, the correlation peak appears at a different time, but with the same strength. This is again due to the fact that the FTs of two data streams that are identical but shifted in time have the same intensities, and differ only in phase.

9.3 Numerical Model of the Temporal Correlator using Two level Atoms

Consider an ensemble of two level atoms excited by a monochromatic field of frequency ω. Here, $\hbar\omega_1$ and $\hbar\omega_2$ are the energies of levels $|1\rangle$ and $|2\rangle$ which are coupled by a laser field with a Rabi frequency of Ω_0 and a detuning of δ. The Schrodinger equation under rotating wave approximation and transformation can be expressed as[51]:

$$i\hbar\,\partial|\Psi\rangle/\partial t = H|\Psi\rangle;\quad H/\hbar = -\delta|2\rangle\langle 2| + \Omega_0/2(|1\rangle\langle 2| + |2\rangle\langle 1|) \tag{9.1}$$

where $\delta \equiv \omega - (\omega_2 - \omega_1)$ and $|\Psi\rangle = c_1|1\rangle + c_2|2\rangle$. If we want to include the effect of decay due to spontaneous emission at the rate of Γ from state $|2\rangle$, we must make use of the density matrix equation of motion[5151]. However, the practical system we propose, as described in ref 52, is an effective two level system involving two metastable states. As such, we can set $\Gamma = 0$ in this model. This allows to us to make use of the amplitude equations (eqns. (**9.1**)) to find the temporal evolution of each atom. The general solution of this equation can be expressed as follows:

$$\begin{bmatrix} c_1(t+\Delta t) \\ c_2(t+\Delta t) \end{bmatrix} = e^{\frac{i\delta t}{2}} \begin{bmatrix} \cos(\Omega'\Delta t/2) - i\,\delta/\Omega'\sin(\Omega'\Delta t/2) & -i\,\Omega_0/\Omega'\sin(\Omega'\Delta t/2) \\ -i\,\Omega_0/\Omega'\sin(\Omega'\Delta t/2) & \cos(\Omega'\Delta t/2) + i\,\delta/\Omega'\sin(\Omega'\Delta t/2) \end{bmatrix} \begin{bmatrix} c_1(t) \\ c_2(t) \end{bmatrix} \tag{9.2}$$

where, $\Omega' = \sqrt{\Omega_0^2 + \delta^2}$. To simulate the process of stimulated photon echo, for example, we start with an ensemble of two-level systems with a ground state $|1\rangle$ and an excited state $|2\rangle$, as depicted in the inset of Figure 9.1(a). The spectral atomic distribution has a Gaussian profile with a width of σ due to Doppler shift. The effective detuning seen by an atom moving with a velocity v in the direction of the laser beam is given by $\delta = \delta_o - kv$, where $\delta_o = \omega - \omega_o$ is the detuning of the laser for a stationary atom, and $\omega_o = \omega_2 - \omega_1$ is the resonance frequency of the atom. Note that, alternatively, this is equivalent to a laboratory frame picture in which the laser frequency is fixed at ω, an atom with velocity v has a resonance frequency of $\omega_v = \omega_o + kv = (\omega_2 - \omega_1) + kv$, and the detuning experienced by this atom is $\delta = \omega - \omega_v = \omega - (\omega_o + kv) = \delta_o - kv$. For our simulations, we assume the laser to be resonant with the stationary atoms, so that $\omega = \omega_o$, and $\delta_o = 0$.

We suppose that, initially, all N atoms are prepared in the state $|1\rangle$. We first determine the quantum state of a band of atoms, with velocity v, after it has interacted with several laser pulses in sequence. This state is then used to determine the amplitude and phase of the induced dipole moment [proportional to $\rho_{12} \equiv c_1(t)c_2^*(t)e^{i[\delta t+(\omega_2-\omega_1)t]}$] oscillating at the frequency of $\omega_v = \omega_o + kv$. We then calculate the response of all the atoms with different velocities and add them together, weighted by the Gaussian distribution as a function of velocity. The resulting net dipole moment induced by the photon echo is given by: $P(t_4) = \int_{-\infty}^{\infty} \rho_{12}(t_4,\delta)\, g(\delta)\, d\delta$; where $g(\delta)$ is the Gaussian spectral distribution. The electric field of the resulting optical pulse is proportional to this net dipole moment. To ensure high-fidelity recording of the pulses in the atomic medium, we constrain the duration of the shortest light pulse to be at least five times longer than the inverse of σ. The simulation results using the quantum mechanical amplitude equations are described in details in section 9.5.

9.4 Derivation of the Transfer Function for the Temporal Correlator (Analytical Model)

To illustrate the temporal correlation process mathematically, we derive here the transfer function of the temporal correlator. In a later section, we show that this model is in close agreement with the response determined by solving the equations of motion of the atomic medium explicitly.

Let us consider the case where three temporal signals denoted as A(t), B(t) and C(t), are encoded on a laser beam with a modulator, as shown in Figure 9.1(b). These functions represent the complex envelope of the electric field amplitude, with a central frequency of ω_L. Explicitly, we can write:

$$E_Q(t) = Q(t)\exp\left(i(\omega_L t - kz)\right) + cc = \left|Q(t)\right|\exp(i\phi_Q); \quad (Q = A, B, C) \tag{9.3}$$

After applying the rotating wave approximation and the rotating wave transformation[51] (which is augmented to transform out the common phase factor *kz* as well), we find that the effective Hamiltonian for each of these fields can be expressed as:

$$H_Q(t)/\hbar = \omega\left|2\right\rangle\left\langle 2\right| + \Omega_Q(t)/2\left|1\right\rangle\left\langle 2\right| + \Omega_Q^*(t)/2\left|2\right\rangle\left\langle 1\right|; \quad (Q = A, B, C) \tag{9.4}$$

where the complex and time dependent Rabi frequency for each field is given by: $\Omega_Q(t) = \mu Q(t) = \left|\Omega_Q(t)\right|\exp(i\phi_Q), (Q = A, B, C)$; with μ being the dipole moment of the two level system, and the detuning of the center frequency of the laser (ω_L) from the resonance frequency of the atom (ω_{Atom}) is defined as $\omega \equiv \omega_{Atom} - \omega_L$.

As shown in Figure 9.1(b), the three temporal signals have finite durations in time, and are separated from one another. Specifically, we assume that the three signals, A, B and C, arrive at the atomic medium at t=T_1, T_2 and T_3, respectively. Therefore, the Rabi frequencies seen by the atomic medium can be expressed as $\Omega_q(t) = \mu q(t) = \left|\Omega_q(t)\right| \exp(i\phi_Q); (Q = A, B, C; q = a, b, c)$, where $a(t) = A(t - T_1);\ b(t) = B(t - T_2);\ c(t) = C(t - T_3)$. Before proceeding further, we define explicitly the time domain Fourier Transform, $\tilde{g}(\omega)$ of a function $g(t)$ as follows:

$$g(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \tilde{g}(\omega) \exp(-i\omega t) d\omega\ ; \qquad \tilde{g}(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} g(t) \exp(i\omega t) dt \tag{9.5}$$

From this definition, it then follows immediately that:

$$\tilde{\Omega}_a(\omega) = \mu \tilde{A}(\omega) \exp(i\omega T_1); \quad \tilde{\Omega}_b(\omega) = \mu \tilde{B}(\omega) \exp(i\omega T_2); \quad \tilde{\Omega}_c(\omega) = \mu \tilde{C}(\omega) \exp(i\omega T_3) \tag{9.6}$$

In the time domain, the atoms see the pulses at different times. However, the equivalent picture in the frequency domain is that the atoms see the Fourier components of all the pulses *simultaneously*, during the time window within which all three pulses are present. Thus, for $t \geq T_3$, the response of the atomic medium can be computed by assuming that it has interacted with all the fields simultaneously. To evaluate this response, we denote first as $N(\omega)$ the distribution of the atomic frequency detunings (i.e., the inhomogeneous broadening). Thus, the quantity $N(\omega)d\omega$ represents the number of atoms which have detunings ranging from $\omega - d\omega/2$ to $\omega + d\omega/2$, representing a spectral band of width $d\omega$. In the spectral domain view, a good approximation to make is that the atoms interact only with those components of the field that are resonant with the atoms, within a small band, justified by the fact that the spectral component within a vanishingly small band, for short enough pulses, is very small. Thus, the Schrodinger Equation for the amplitude of *this band of atoms,* in the rotating wave frame, is given by:

$$\frac{\partial}{\partial t}\begin{bmatrix} C_1(\omega) \\ C_2(\omega) \end{bmatrix} = -i\begin{bmatrix} 0 & \tilde{\Omega}(\omega)/2 \\ \tilde{\Omega}^*(\omega)/2 & 0 \end{bmatrix}\begin{bmatrix} C_1(\omega) \\ C_2(\omega) \end{bmatrix} \tag{9.7}$$

where the net, complex Rabi frequency within this band is given by $\tilde{\Omega}(\omega) = \tilde{\Omega}_a(\omega) + \tilde{\Omega}_b(\omega) + \tilde{\Omega}_c(\omega) = \mu(\tilde{a}(\omega) + \tilde{b}(\omega) + \tilde{c}(\omega)) \equiv \left|\tilde{\Omega}(\omega)\right| \exp(-i\phi(\omega))$

Assuming that all the atoms are in the ground state before the first pulse is applied, the solution for this equation, physically valid for $t \geq T_3$, is given by: $C_1(\omega) = Cos(\left|\tilde{\Omega}(\omega)\right| t/2); C_2(\omega) = -iSin(\left|\tilde{\Omega}(\omega)\right| t/2)\exp(i\phi(\omega))$. The amplitude of the electromagnetic field produced by the atoms *in this band* is proportional to the induced dipole moment, which in turn is proportional to the induced coherence, given by:

$$\begin{aligned} \rho_{12}(\omega,t) &= C_1 C_2^* \exp(-i\omega_{atom} t) = C_1 C_2^* \exp(-i\omega_L t - i\omega t) \\ &= (i/2)\exp(-i\omega_L t) Sin[\left|\tilde{\Omega}(\omega)\right| t]\exp(-i\omega t) \times \exp(i\phi(\omega)) \end{aligned} \tag{9.8}$$

As we argued above, the component of the Rabi frequency within a very small band is very small, so that we can make use of the approximation that $\mathrm{Sin}(\theta) \approx \theta - \theta^3/6$. Noting that the interaction occurs for a time window of duration $T \approx T_3 - T_1$, we can thus write that:

$$\rho_{12}(\omega,t) \approx (i/2)\exp(-i\omega_L t) \times [\tilde{\Omega}(\omega)T - \left|\tilde{\Omega}(\omega)\right|^2 \tilde{\Omega}(\omega)T^3/6]\exp(-i\omega t) \tag{9.9}$$

The signal (i.e., the electric field) produced by all the atoms can be expressed as:

$$\Sigma(t) = \alpha \exp(-i\omega_L t)\int_{-\infty}^{\infty} d\omega N(\omega)[\tilde{\Omega}(\omega)T - \left|\tilde{\Omega}(\omega)\right|^2 \tilde{\Omega}(\omega)T^3 / 6]\exp(-i\omega t) \tag{9.10}$$

where the proportionality constant, α, depends on the dipole moment of the two level system and the density of the atomic medium. For extracting the essential result, we assume that the width of the atomic spectral distribution is very large compared to that of $\tilde{\Omega}(\omega)$, so that $N(\omega)$ can be replaced by a constant, N. Furthermore, we define $\Sigma'(t) = \Sigma(t)\exp(i\omega_L t)$ as the envelope of the signal centered at the laser frequency, and $\beta = -\alpha N$, so that we can write:

$$\Sigma'(t) = \beta\int_{-\infty}^{\infty} d\omega[\left|\tilde{\Omega}(\omega)\right|^2 \tilde{\Omega}(\omega)T^3 / 6 - \tilde{\Omega}(\omega)T] \times \exp(-i\omega t) \tag{9.11}$$

Note that the time dependent value of the off diagonal density matrix element $\rho_{12}(t)$, integrated over all atoms, is simply proportional to this signal: $\rho_{12}(t) = \xi\Sigma(t)$, where ξ is a proportionality constant. This, of course, is proportional to the density matrix element in the rotating wave frame:

$$\tilde{\rho}_{12}(t) \equiv \rho_{12}e^{i\omega_L t} = \xi\Sigma(t)e^{i\omega_L t} \tag{9.12}$$

The linear terms in eqn, (**9.11**) represent the so-called free-induction decay which occurs immediately after each pulse leaves the atomic medium, as can be shown easily, and do not contribute to the correlation signal. Since the net Rabi frequency has three components, corresponding to the three pulses, there will be a total of twenty seven components corresponding to the non-linear term. However, some of these terms are identical to one another except for numerical coefficients, leading to eighteen distinct terms. They are tabulated in table 1, where the coefficient in front of each term indicates the number of

Table 9.1: List of nonlinear terms from third order expansion

	Terms Appearing at $t \leq T_3$		Terms Appearing at $t > T_3$	
	Nonlinear Terms	Temporal Position, t	Nonlinear Terms	Temporal Position, t
Causal	$\tilde{a}^2\tilde{a}^*$	T_1	$2\tilde{a}^*\tilde{b}\tilde{c}$	$-T_1+T_2+T_3$
	$\tilde{b}^2\tilde{b}^*$	T_2	$\tilde{a}^*\tilde{c}^2$	$-T_1+2T_3$
	$\tilde{c}^2\tilde{c}^*$	T_3	$\tilde{b}^*\tilde{c}^2$	$-T_2+2T_3$
	$2\tilde{a}\tilde{a}^*\tilde{b}$	T_2		
	$2\tilde{b}\tilde{b}^*\tilde{c}$	T_3		
	$2\tilde{a}\tilde{a}^*\tilde{c}$	T_3		
	$\tilde{a}^*\tilde{b}^2$	$-T_1+2T_2$		
Acausal	$\tilde{a}^2\tilde{b}^*$	$(2T_1-T_2) < T_1$	A: Causal & Appearing at $t \leq T_3$ B: Acausal C: Causal & Appearing at $t > T_3$	
	$\tilde{a}^2\tilde{c}^*$	$(2T_1-T_3) < T_1$		
	$\tilde{b}^2\tilde{c}^*$	$(2T_2-T_3) < T_2$		
	$2\tilde{a}\tilde{b}\tilde{c}^*$	$(T_1+T_2-T_3) < T_1$		
	$2\tilde{a}\tilde{b}\tilde{b}^*$	T_1		
	$2\tilde{b}\tilde{c}\tilde{c}^*$	T_2		
	$2\tilde{a}\tilde{c}\tilde{c}^*$	T_1		
	$2\tilde{a}\tilde{b}^*\tilde{c}$	$T_1-T_2+T_3 < T_3$		

times it occurs. These terms can first be divided into two categories: causal and acausal. The acausal terms occur at a time that is earlier than the time of application of at least one of the three constituent input signals. These appear due to the fact that we have used Fourier Transforms instead of Laplace Transforms in our analysis. It is to be understood that these terms are unphysical. However, using Fourier Transform in the analysis of the temporal correlator makes it convenient to multiplex with the spatial correlator for use in 3D STC, as illustrated in chapter 10.

The causal category can be broken up into two groups: those appearing at $t \le T_3$ and those appearing after $t > T_3$. For reference, we thus have three different groups of signals, designated as follows: Group A: Causal and appearing before or at T_3; Group B: Acausal; Group C: Causal and appearing after T_3. This grouping is indicated in the lower right corner of Table 9.1. In grouping these terms, we have assumed that $(T_2 - T_1) < (T_3 - T_2)$. Under the assumptions made here in deriving these results, the only physically meaningful terms are those in Group C, since we are calculating the response of the atoms to the combined field of all three pulses. For the STC, the only relevant terms are also those in group C. These correspond to Inverse Fourier Transform of $\tilde{a}^*\tilde{b}\tilde{c}, \tilde{b}^*\tilde{c}^2, \tilde{a}^*\tilde{c}^2$, appearing at time t = -T$_1$ +T$_2$ +T$_3$, t = -T$_2$ + 2T$_3$ and t = -T$_1$ + 2T$_3$. The term that corresponds to the desired correlation signal is $\tilde{a}^*\tilde{b}\tilde{c}$. In the next chapter , we discuss in detail how to choose various parameters in a way that ensures no overlap between this correlation signal and the signal corresponding to the other two nonlinear terms in group C.

We now consider explicitly the signal produced by the term $\tilde{a}^*\tilde{b}\tilde{c}$. It can be expressed as: $\Sigma'_C(t) = [\beta\mu^3/6]\int_{-\infty}^{\infty} d\omega[\tilde{a}^*(\omega)\tilde{b}(\omega)\tilde{c}(\omega)]\exp(-i\omega t)$. For simplicity, we now define the normalized signal as $\sigma(t) \equiv \Sigma'_C \sqrt{2\pi} * 6/[\beta\mu^3]$, so that we can write:

$$\sigma(t) = 1/\sqrt{2\pi}\int_{-\infty}^{\infty} d\omega[\tilde{a}^*(\omega)\tilde{b}(\omega)\tilde{c}(\omega)]\exp(-i\omega t) \tag{9.13}$$

It then follows that the FT of the normalized correlation signal is:

$$\tilde{\sigma}(\omega) = \tilde{a}^*(\omega)\tilde{b}(\omega)\tilde{c}(\omega) = \tilde{A}^*(\omega)\tilde{B}(\omega)\tilde{C}(\omega)\exp(j\omega(T_3 + T_2 - T_1)) \tag{9.14}$$

If we define $\tilde{S}(\omega)=\tilde{A}^{*}(\omega)\tilde{B}(\omega)\tilde{C}(\omega)$ then then it follows that S(t) is the cross-correlation between A(t) and the convolution of B(t) and C(t). Since A(t) is essentially a delta function in time, S(t) is effectively the convolution of B(t) and C(t). Explicitly, if we consider $A(t)=A_0\delta(t)$, we get $S(t)=A_0\int_{-\infty}^{\infty}B(t')C(t-t')dt'$.

If we take into account the finite temporal width of A(t), this signal S(t) will be broadened by this added width. Finally, we note that $\sigma(t)=S(t-(T_3+T_2-T_1))$, which means that this correlation signal occurs at $t=T_3+(T_2-T_1)$, as already noted above. These results have been compared with actual numerical simulation in section 9.5, with close agreement.

9.5 Numerical Simulations of the Stimulated Photon Echo and Temporal Correlator

In section 9.2 through 9.4, we have introduced the numerical and the analytical model and claimed that they are essentially equivalent. In this section, we verify this claim by comparing the simulation results of both models. For simplicity, first we show the simulation of a stimulated photon echo (SPE) process which is the simplest version of the temporal correlator.

Figure 9.2(a) shows the sequence of pulses associated with the SPE. A short pulse a(t) is applied as the writing beam at time T_1, followed by the query pulse b(t) at time T_2. At time T_3, the reference pulse is applied to this memory. An echo pulse is observed in a temporally shift invariant manner at time $t = -T_1 +T_2+T_3$. If the SPE process is viewed as a temporal correlator, then the echo pulse represents the correlation peak. The photon echo process described above is simulated using the quantum mechanical amplitude equations (details in section 9.3), and the result is shown in Figure 9.2(b). Here, we have used an idealized, decay-free two level system of atoms with an inhomogeneous broadening that is atleast five times larger than the inverse of the temporal width of each of the input pulse (all of which have the same width). The atoms are initially prepared in the ground state, so that we have $c_1(t_0 = 0) = 1$ and $c_2(t_0 = 0) = 0$. Here, the first π/10 pulse is followed, at time T_1, by two other π/10 pulses appearing at T_2 and T_3 respectively. The echo appears at $t = T_2 + (T_3 - T_1)$. The simulation of the atomic model has been performed in a super computer for faster calculation. In addition to the correlation term other nonlinear terms appear in the simulation which have been discussed in detail in section 9.4. In the analysis of section 9.4 , we have shown that out of twenty seven nonlinear terms only three terms are relevant which correspond to the Inverse Fourier Transform of $\tilde{a}^*\tilde{b}\tilde{c}, \tilde{b}^*\tilde{c}^2, \tilde{a}^*\tilde{c}^2$ and they are categorized in group C (in Table 9.1). In the simulation of the temporal correlator using the numerical model in Figure 9.2(b), we see that the above mentioned terms are appearing at time $t = -T_1 + T_2 + T_3$, $t = T_2 + 2T_3$ and $t = -T_1 + 2T_3$. Here, the desired correlation signal, which is denoted as s(t), corresponds to the term $\tilde{a}^*\tilde{b}\tilde{c}$ and it appears at time $t = -T_1 + T_2 + T_3$ as expected. It should be noted that there is a signal that appears at time $t = -T_1 + 2T_2$. This can be understood as the conventional photon echo signal. Since it appears before $t = T_3$, it is of no interest for the temporal correlation process.

To simulate the SPE using the analytical model, it is necessary to modify the transfer function of eqn.(**9.14**) by adding additional non-linear terms in Table 9.1. As we have discussed in detail in section

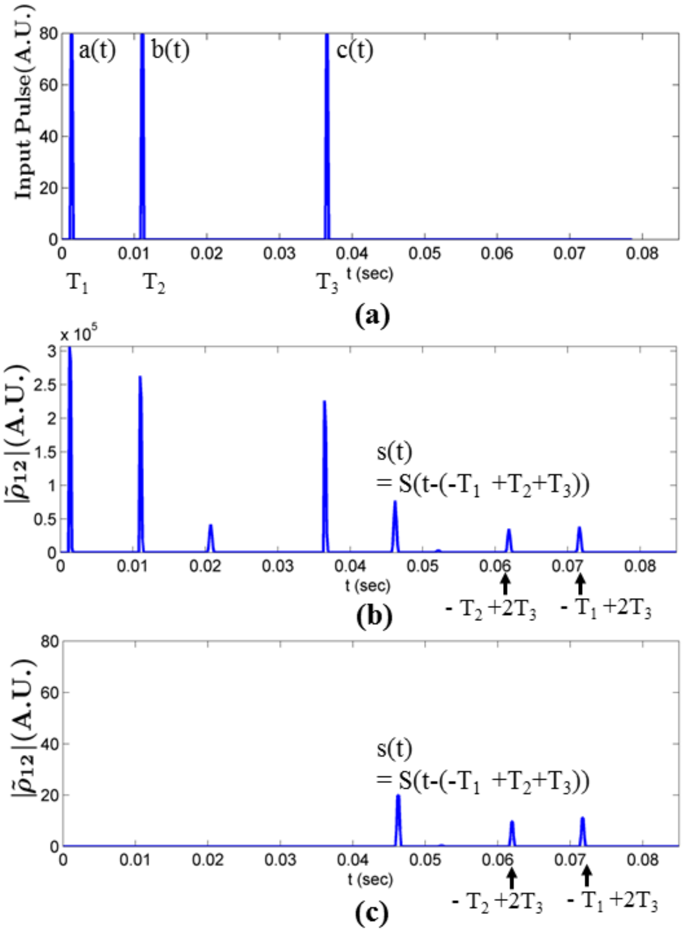


Figure 9.2: (a) Pulse sequence associated to the three pulse photon echo in two-level atoms. (b) Simulation results of photon echo using the numerical model which employs quantum mechanical amplitude equation and (c) Simulation results using the analytical model. Both model yields the same result with faster simulation time for the analytical model. [A.U. = Arbitrary Unit]

9.4, the only terms in Table 9.1 that are physically meaningful are those appearing in group C, which occur at $t > T_3$. Thus, the modified version of the transfer function can be expressed as:

$$\tilde{\sigma}(\omega_T) = 2\tilde{a}^*(\omega_T)\tilde{b}(\omega_T)\tilde{c}(\omega_T) + \tilde{b}^*(\omega_T)\tilde{c}^2(\omega_T) + \tilde{a}^*(\omega_T)\tilde{c}^2(\omega_T) \tag{9.15a}$$

$$\sigma(t) = \mathrm{IFT}\{\tilde{\sigma}(\omega_T)\}, \text{ for } t > T_3 \qquad (9.15b)$$

where the three terms corresponds to group C in table 9.1. The quantity $\sigma(t)$ is proportional to the signal produced by the system, for $t > T_3$ under the simplifying assumption that the inhomogeneous broadening is much larger than the spectral spread of the terms in (**10.6**), as shown in section 9.4 [eqns (**9.8**) - (**9.12**)]. We have also shown in section 9.4 that this term is proportional to $\tilde{\rho}_{12}(t)$, the off-diagonal density matrix element in the rotating wave basis. In Figure 9.2(c) we have shown that implementing the transfer functions of eqn. (**10.6**) essentially yields the same result as the numerical model (for $t > T_3$, which is the relevant time span for the correlator), but at about 2×10^5 times faster simulation time for this particular case. To compare the correlation peak values for both models, let us define the ratio of the magnitude of the writing pulse, a(t) to the magnitude of the output correlation peak, s(t) as η. For the numerical model, η is 3.95 and for the analytical model, η is 3.99. The other nonlinear terms also maintain same magnitude ratio in both models.

Now, consider the situation where the reference pulse and the query pulse are each replaced by a short sequence of pulses as shown in Figure 9.3(a). Figure 9.3(b) and Figure 9.3(c) show the simulation results of the corresponding temporal correlator using the numerical model and the analytical model, respectively. Here, first we apply a very weak write pulse, a(t) = A(t- T_1). Next we apply the query stream, b(t) = B(t- T_2) and after some time lag the reference stream, c(t) = C(t-T_3) is applied. The final correlation signal appears at time $t = -T_1 + T_2 + T_3$ inside the dotted box. It is obvious from the simulation

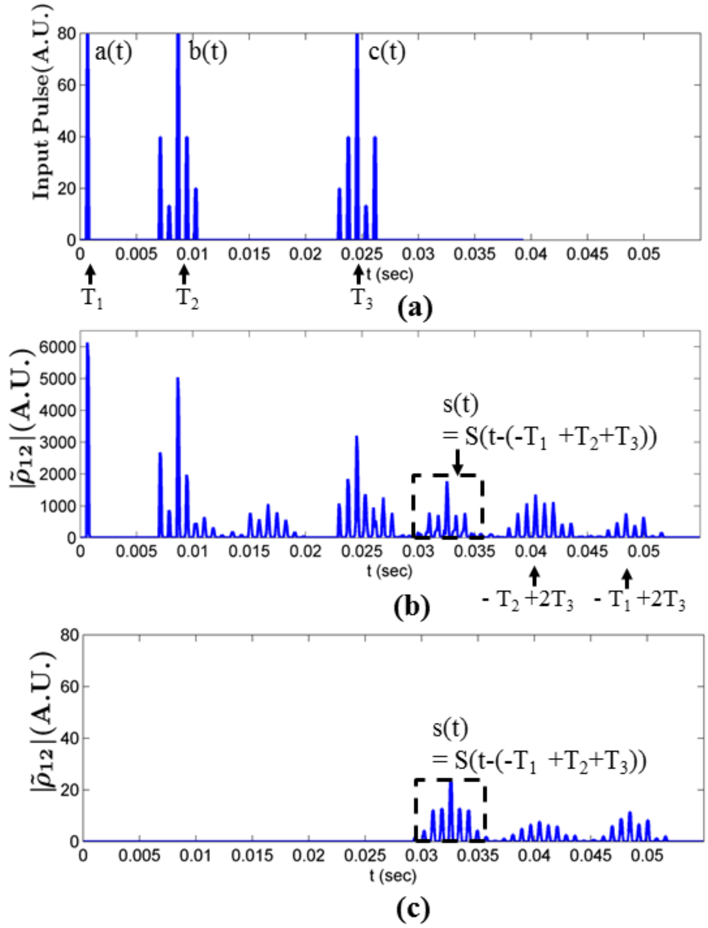


Figure 9.3: (a) Pulse train associated to the temporal correlator. (b) Simulation results of a temporal correlator using the numerical model which employs mechanical amplitude equation and (c) the analytical transfer function model. [A.U. = Arbitrary Unit]

that both models yield essentially the same result. However, the analytical model is much faster. Note that the other nonlinear terms appearing in this case do not interfere with the final correlation signal. This issue is discussed in greater detail later on in the next chapter.

Chapter 10 Architecture and Simulation Results of the Spatio Temporal Correlator (STC)

10.1 Introduction

Automated target recognition (ATR) has been a very active field of research for several decades. Significant advances in ATR have been made using analog approaches employing holographic correlators, as well as computational approaches using dedicated digital signal processing (DSP) chips or software. However, these techniques are inadequate for the task of automatic event recognition (AER). AER is defined as the task of identifying the occurrence of an event within a large video data base. Consider, for example, a data base that contains video surveillance gathered by a camera-equipped drone or a satellite, which is monitoring a site for suspicious activities, such as a truck of a particular size entering or exiting a facility. For data gathered over a few hours, this event may have occurred several times. The goal of an AER system is to determine if these events occurred, when they occurred, and how many times. In principle, this can be achieved by searching through each frame in the data base, and comparing them with reference images. This process is prohibitively time consuming, even with a very efficient optical image correlator, a software or a DSP based image recognition system. However, by employing the properties of atoms[20, 21, 22, 24, 53] it is possible to realize an AER that can recognize rapidly the occurrence of events, the number of events, and the occurrence times. Separate aspects of the overall technology needed for the AER have also been demonstrated by us earlier[20, 21, 25, 26, 27]. In this chapter, we describe quantitatively the design of an AER system using Rb vapor in nano-porous glass with paraffin coating and show the simulation results of such a system. The AER will be realized by the technique of spatio-temporal holographic correlation[22, 23, 28, 29, 30]. This combines the process of translation invariant spatial holographic correlation with the process of translation invariant temporal correlation. The resulting system is a Spatio-Temporal Correlator (STC). In previous chapters, we have already described different aspects of the shift invariant holographic correlator and the temporal correlators. In this chapter, we will show how temporal correlator can be combined with the holographic correlator to realize the AER system.

The Spatio-Temporal Correlator (STC), shift invariant in space and time, can recognize rapidly an event that may be present in a video file, and determine the temporal location of the event. In general, modeling the STC requires determining the temporal dynamics of a large number of inhomogeneously broadened atoms, multiplexed with free-space wave propagation equations. Here, along with modeling the STC using the Schrodinger equation for the temporal evolution of atoms excited by optical fields, we show that the response of the STC can be determined by modeling the response of the interaction medium as a simple, three-dimensional, multiplicative transfer function in the spatio-temporal Fourier domain. We explain the physical origin of this model, and then establish the validity of this model by comparing its prediction with that determined via the quantum mechanical dynamics. We then show that the transfer function based analytical model can be used to find the response of the STC at a speed that is many orders of magnitude faster than solving the SE.

10.2 Automatic Event Recognition (AER) via Translation Invariant Spatio-Temporal Correlator

The natural extension to searching for images in spatial domain and searching for signals in time is to search for an image that is changing in time, which is simply a video clip corresponding to an event. Consider a video signal, either from a live camera feed or from a DVD player, for example. The video signal is an image that is changing in time, i.e. it has both spatial and temporal properties. An AER system can recognize a short clip within that video feed, by using a Translation Invariant Spatio-Temporal Correlator (TI-STC) which is a combination of the TI-SHC and the TI-TC.

The AER process is illustrated schematically in Figure 10.1. In Figure 10.1(a), we show a series of 30 consecutive frames in a video. Frames 1-10 are of a static, unchanging scene. In frame 11 a car begins to drive across the scene, and stops in frame 20. Frames 21-30 are again static. This is illustrated in Figure 10.1(b). If we consider a single pixel from this video signal, as indicated in Figure 10.1(c), the resulting pattern as the car moves from one place in frame 11 to another in frame 20 is shown in Figure 10.1(d). This signal is akin to the signals we used in the temporal signal correlation example in section 9.2. Each pixel in the video will have a corresponding bit stream, which could all be recognized separately in a temporal correlator. However, by combining the spatial correlator, we can now recognize a group of pixels that form an image, as they change in time.

As an example, consider the case where the event to be recognized is the car driving from one place to another. In the AER system, the whole video will be the database, and the ten clips corresponding to the movement of the car will be the query event. The output correlation signal in the AER system will

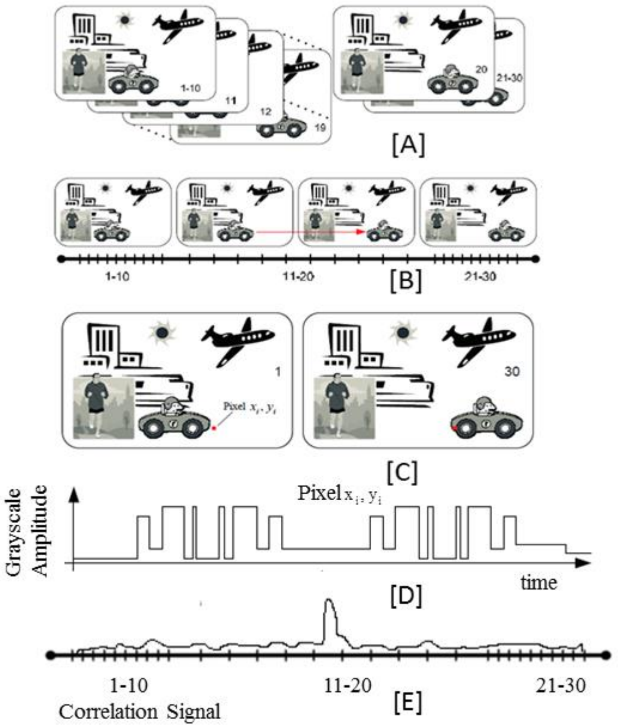
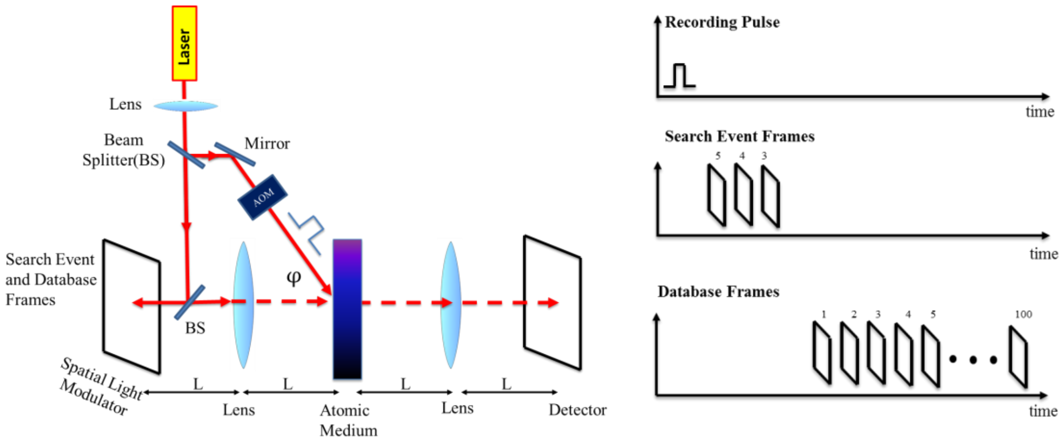


Figure 10.1: Schematic illustration of the automatic event recognition process. See text for details.

contain a correlation peak corresponding to the time between frames 15 and 16, where middle of the query event occurs, as illustrated in Figure 10.1(e). Because the system is translation invariant spatially, we can find the car driving across the scene no matter where it occurs within the frame. Furthermore, because the system is translation invariant in time as well, no a priori knowledge regarding the start time of the query event within the database video is necessary. Furthermore, the time at which the correlation peak will appear can be used to infer the location of the query event within the database video. It should also be noted that if the same event occurs N times within the database video, N different correlation peaks will be observed. Finally, if the polar Mellin transform[12,13,15,16,31] is used to pre-process each of the frames in the database video as well as the query video, it would be possible to recognize the event even when the images in the query clip are scaled and rotated with respect to the database video.

10.3 Architecture of the Spatio-Temporal Correlator (STC)

The experimental configuration for realizing the AER system is illustrated schematically in Figure 10.2. The architecture for the AER system is similar to that of a conventional spatial holographic correlator except that the write pulse is replaced by a plane wave of certain duration in space and time. The recording medium is replaced by the inhomogeneously broadened atomic medium (AM). The laser beam is directed to the reflection mode Spatial Light Modulator (SLM) with a polarizing beam splitter. The SLM reflects a pattern of light that is orthogonally polarized so that it passes through the same beam splitter. The pattern produced by the SLM is controlled by signals applied to it. The lens after the SLM



(a) (b)

Figure 10.2: (a) Schematic illustration of the experimental configuration for an Automatic Event Recognition system employing spatio-temporal correlator. The focal length of each lens is L. (b) The sequence at which the recording pulse, reference and query film are applied to the atomic media.

produces the two-dimensional spatial FT of the SLM pattern in the plane of the atomic medium. Similarly, the second lens produces the two dimensional spatial FT of the field, produced by the atomic medium, in the plane of the detector array. The recording pulse is a highly localized spot in the plane of the SLM. For example, if it is located at $\{x = 0; y = y_o\}$, (with $y_o << L$), then, in the plane of the atomic medium (AM), it will appear essentially as a plane wave, moving in the y-z plane, and making an angle of $\sin^{-1}(a / L)$ with the z-direction. Temporally, the duration of the recording pulse is chosen to be much shorter than the temporal separation between consecutive frames in the query event or in the database video. This ensures that its spectrum will be much wider than that of the query clip and the database video. The timing sequence of the pulses is as follows. The recording pulse is applied first, and the time of the arrival of the center of this pulse at the AM is defined as T_1. Following a delay, the frames corresponding to the query event are sent to the SLM. The time of the arrival of the center of this sequence at the AM is defined as T_2. As noted above, the first lens produces a two-dimensional spatial FT of each frame in the plane of the AM. Thus, the interference of the recording pulse and the query frames are stored in the AM, in the form of spatio-spectral gratings. After another delay, the database video frames are sent to the SLM, and the time of the arrival of the center of this sequence at the AM is defined as T_3. Again, the first lens produces a two-dimensional spatial FT of each frame in the plane of the AM. These frames effectively diffract from the spatio-spectral gratings generated by the interference between the recording pulse and the query frames. The resulting signal from the AM passes through the second lens and the detector array records the output signals as functions of time. The output signal, integrated spatially over the detector array, contains a correlation peak at a time corresponding to the position where the matching pattern occurs in the database video. This is, of course, the desired functionality of the AER

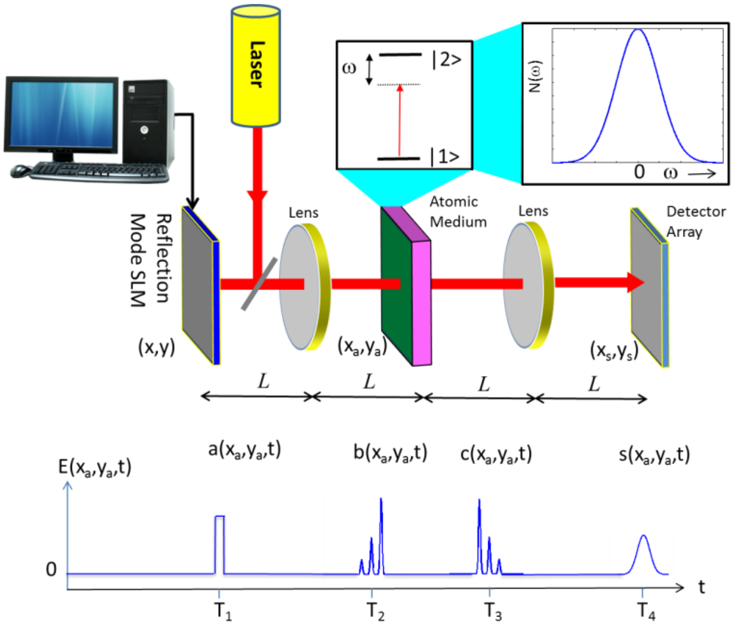


Figure 10.3: Illustration of the spatio-temporal correlator. Here, the focal length of each lens is L. See text for details

.

10.4 Analytical Transfer Function for the STC

We consider next the complete spatio-temporal correlator system, shown in Figure 10.3. Here, the signals are generated by modulating the field from a laser with a spatial light modulator (SLM), for example. Thus, each of three pulses are encoded with two dimensional spatial information. Specifically, we now have three functions containing information: $A(x,y,t)$, $B(x,y,t)$ and $C(x,y,t)$, where (x,y) are the rectilinear spatial coordinates in the plane of the SLM. The corresponding signals in the plane of the atomic medium are Fourier Transformed (FT'd) in the spatial domain due to the first lens. Before writing these functions down, we recall briefly the mathematical relations between a spatial function, $U(x,y)$, in the SLM plane, and the corresponding spatial function $V(x_a, y_a)$ in the atomic medium plane, which has coordinates (x_a, y_a). We consider first the two dimensional FT of the function $U(x,y)$, denoted as $\tilde{U}(k_x,k_y)$, which are related to each other as follows:

$$\tilde{U}\left(k_x,k_y\right) = \frac{1}{2\pi}\int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy U(x,y)\exp[-i(k_x x + k_y y)] \tag{10.1}$$
$$U(x,y) = \frac{1}{2\pi}\int_{-\infty}^{\infty} dk_x \int_{-\infty}^{\infty} dk_y \tilde{U}\left(k_x,k_y\right)\exp[i(k_x x + k_y y)]$$

According to the laws of Fresnel diffraction, and properties of ideal lenses, the function $V(x_a, y_a)$ is given by:

$$V(x_a, y_a) = \frac{\exp[i2kL]}{i\lambda L}\tilde{U}(k_x,k_y)\Bigg|_{k_x=\frac{2\pi x_a}{\lambda L},k_y=\frac{2\pi y_a}{\lambda L}} = \frac{\exp[i2kL]}{i\lambda L}\tilde{U}(\frac{2\pi x_a}{\lambda L},\frac{2\pi y_a}{\lambda L}) \tag{10.2}$$

where L is the focal length of the lens, and λ is the wavelength of the laser. Thus, aside from the inconsequential phase factor and the scaling factor, the spatial function in the plane of the atomic medium is the spatial FT of the spatial function in the plane of the SLM. Similarly, the spatial function of the field produced in the plane of the detector, $W(x_s, y_s)$ which has coordinates (x_s, y_s), will be the spatial FT of the spatial function in the plane of the atomic medium. Thus, in the absence of any interaction with the atomic medium, the spatial function of the filed in the plane of the detector becomes an exact reproduction of the spatial function of the field in the plane of the SLM, but inverted in both *x* and *y* directions, along with a phase shift factor of $\exp(i4kL)$:

$$W(x_s, y_s)\big|_{no\ atomic\ medium} = \exp(i4kL)U(-x,-y) \tag{10.3}$$

This is the well-known 4F imaging process. It should be noted that the prefactor of $(1/i\lambda L)$ in eqn. (**10.2**) gets compensated for during the second stage, and does not appear in eqn. (**10.6**).

We now define the three-dimensional (spatio-temporal) FT of a function $g(x,y,t)$ as follows:

$$\begin{aligned} \tilde{g}\left(k_x,k_y,\omega\right) &= 1/\left[2\pi\right]^{3/2} \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy \int_{-\infty}^{\infty} dt\, g(x,y,t) \exp[-i(k_x x + k_y y + \omega t)] \\ g(x,y,t) &= 1/\left[2\pi\right]^{3/2} \int_{-\infty}^{\infty} dk_x \int_{-\infty}^{\infty} dk_y \int_{-\infty}^{\infty} d\omega\, \tilde{g}\left(k_x,k_y,\omega\right) \exp[i(k_x x + k_y y + \omega t)] \end{aligned} \tag{10.4}$$

Noting that the spatial FTs of the signals A(x,y,t), B(x,y,t) and C(x,y,t) appear at the plane of the atomic medium at times T_1, T_2, and T_3, respectively. We also allow for the situation where each frame might be shifted from the center by an amount $(\mathrm{x_j}, \mathrm{y_j})$.The corresponding Rabi frequencies in the three dimensional spectral domain can be expressed as $\tilde{\Omega}_q(k_x,k_y,\omega) = \zeta\mu\tilde{Q}(k_x,k_y,\omega)\exp(ik_x x_q)$ $\exp(ik_y y_q)\exp(i\omega T_q)$; $[q = a,b,c; Q = A,B,C]$ where $\mathrm{a(x,y,t) = A(x-x_a, y-y_a, t-T_1)}$; $\mathrm{b(x,y,t) = B(x-x_b, y-y_b, t-T_2)}$; c(x, y, t) = C(x - $\mathrm{x_c}$, y - $\mathrm{y_c}$, t -$\mathrm{T_3}$) and $\zeta = \exp(i2kL)/(i\lambda L)$ (see the prefactor in eqn. (**10.3**)). It is to be understood that the spatial frequency components $\{k_x, k_y\}$ corresponds physically to spatial locations $\{(x_a/\lambda L), (y_a/\lambda L)\}$ in the plane of the atomic medium. Thus, the SE for the amplitude in the rotating wave frame, is given by:

$$\frac{\partial}{\partial t}\begin{bmatrix} C_1(k_x,k_y,\omega) \\ C_2(k_x,k_y,\omega) \end{bmatrix} = -i \begin{bmatrix} 0 & \tilde{\Omega}(k_x,k_y,\omega)/2 \\ \tilde{\Omega}^*(k_x,k_y,\omega)/2 & 0 \end{bmatrix} \begin{bmatrix} C_1(k_x,k_y,\omega) \\ C_2(k_x,k_y,\omega) \end{bmatrix} \tag{10.5}$$

where the net, complex Rabi frequency within this band is given by $\tilde{\Omega}(k_x,k_y,\omega)=\tilde{\Omega}_a(k_x,k_y,\omega)+\tilde{\Omega}_b(k_x,k_y,\omega)+\tilde{\Omega}_c(k_x,k_y,\omega)=\mu(\tilde{a}(k_x,k_y,\omega)+\tilde{b}(k_x,k_y,\omega)$ $+\tilde{c}(k_x,k_y,\omega))$. Here, we are making a more generalized form of the REA. Specifically we are assuming that the atoms in the Spatio-Temporal Fourier Transform (STFT) domain volume element $(\mathrm{k_x}-\Delta\mathrm{k_x}/2,\mathrm{k_y}-\Delta\mathrm{k_y}/2,\omega-\Delta\omega/2)\rightarrow(\mathrm{k_x}+\Delta\mathrm{k_x}/2,\mathrm{k_y}+\Delta\mathrm{k_y}/2,\omega+\Delta\omega/2)$ interacts only with those components of the field that occupy the same volume element. Using the same line of arguments and making use of the same set of approximations as presented for the temporal-only case, we then conclude that the normalized signal in the plane of the detector array, corresponding to the correlation signal, is given by:

$$\begin{aligned}\sigma(x_s,y_s,t)=&\frac{1}{[2\pi]^{3/2}}\int_{-\infty}^{\infty}dk_x\int_{-\infty}^{\infty}dk_y\int_{-\infty}^{\infty}d\omega[\tilde{A}^*(k_x,k_y,\omega)\tilde{B}(k_x,k_y,\omega)\tilde{C}(k_x,k_y,\omega)\\&\times\exp(jk_x(x_3+x_2-x_1)+jk_y(x_3+x_2-x_1))\exp(j\omega(T_3+T_2-T_1))]\exp(-j(k_xx_s+k_yy_s+\omega t))\end{aligned}\tag{10.6}$$

If we consider no shift in the spatial direction i.e. $\mathrm{x_q}=\mathrm{y_q}=0$ where q = a,b,c, then it follows that the three-dimensional FT of the normalized correlation signal is:

$$\tilde{\sigma}(k_x,k_y,\omega)=\tilde{A}^*(\omega)\tilde{B}(\omega)\tilde{C}(\omega)\exp(j\omega(T_3+T_2-T_1))\tag{10.7}$$

If we define

$$\tilde{S}(k_x,k_y,\omega)=\tilde{A}^*(k_x,k_y,\omega)\tilde{B}(k_x,k_y,\omega)\tilde{C}(k_x,k_y,\omega)\tag{10.8}$$

then it follows that $S(x_s,y_s,t)$ is the three dimensional cross-correlation between $A(x,y,t)$ and the three-dimensional convolution of $B(x,y,t)$ and $C(x,y,t)$. Since $A(x,y,t)$ is essentially a delta function in both temporal and spatial domains (i.e., it is a very short temporal pulse, and is a small point signal at the center of the SLM plane), $S(x_s, y_s, t)$ is effectively the three-dimensional convolution of $B(x, y, t)$ and $C(x, y, t)$. Explicitly, if we consider $A(x, y, t) = A_0\delta(x)\delta(y)\delta(t)$, we get:

$$S(x_s, y_s, t) = A_o\int_{-\infty}^{\infty} dt'\int_{-\infty}^{\infty} dx'\int_{-\infty}^{\infty} dy' B(x', y', t')C(x_s - x', y_s - y', t - t') \qquad (10.9)$$

If we take into account the finite width of $A(x, y, t)$ in all three dimensions, this signal $S(x_s, y_s, t)$ will be broadened by these added widths in each dimension. Finally, we note that $\sigma(x_s, y_s, t) = S(x_s, y_s, t - (T_3 + T_2 - T_1))$, which means that this correlation signal occurs at $t = T_3 + (T_2 - T_1)$, in the plane of the detector array.

As noted above, the correlation signal produced in the detector plane can also be calculated explicitly by solving the equation of motion for the atoms explicitly. For a large database, such a computation is exceedingly time consuming, since the atomic medium is inhomogeneously broadened. However, the analytical model presented here makes the computation much faster.

10.5 Numerical Simulations Results of the Automatic Event Recognition System

Figure 10.4 shows the limiting case of the STC where the query event and the database event are each a single frame and they match exactly in the spatial domain. Specifically, the write pulse, A(x,y,t) is applied (centered) at $\mathrm{t = T_1}$. In the time domain, it is a $\pi/2$ pulse, while spatially it is a gaussian spot (centered) at $\mathrm{x = 0, y = 0}$. The query image, B(x,y,t) is applied (centered) at $\mathrm{t = T_2}$

,and the reference image C(x,y,t) is applied (centered) at $t = T_3$. As expected, a correlation peak appears at $t = T_2 + T_3 - T_1$ .The correlation signal is computed in two different ways: first by using the explicit numerical model described in 9.3, and then using the analytical model defined by equations (**10.7**), (**10.8**) and (**10.9**). Here, we can see that the numerical and the analytical models yield essentially identical results but at a faster simulation time. The numerical model is simulated using a parallel computer[54,55,56] with 20 processors whereas the analytical model is simulated using a single processor (Intel Core I7-4600U CPU @2.1GHz). However, the simulation time of the analytical model is ~10^5 faster than the numerical model. Both models also show additional signals appearing at different times, corresponding to

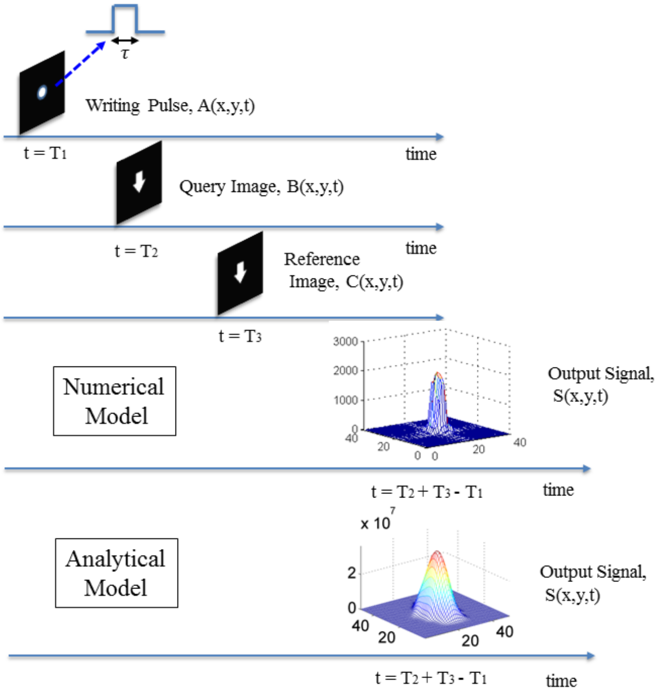


Figure 10.4: Simulation of the STC using the numerical model and analytical model. The output signal yields the same result for both models.

the additional non-linear interactions, similar to the results shown in Figure 9.2. Here, for clarity, we have shown only the correlation signals, and excluded these additional signals.

From the above simulation results, it is obvious that the numerical model and the analytical model are in good agreement with each other. Hence, we can use the analytical model reliably for simulating a three dimensional STC. Figure 10.5 shows the simulation result of the STC using the analytical model, where we have five query frames and ten reference frames. At time $T_{1,}$ we apply the write pulse with a very short duration in time and space. In the spatial domain, the writing pulse is a small spot localized in the center of the frame. After some delay, a query video clip of five frames, denoted as B(x,y,t), are applied, at a frame rate of 30 fps (frame per second). The center of the query clips occur at time T_2.

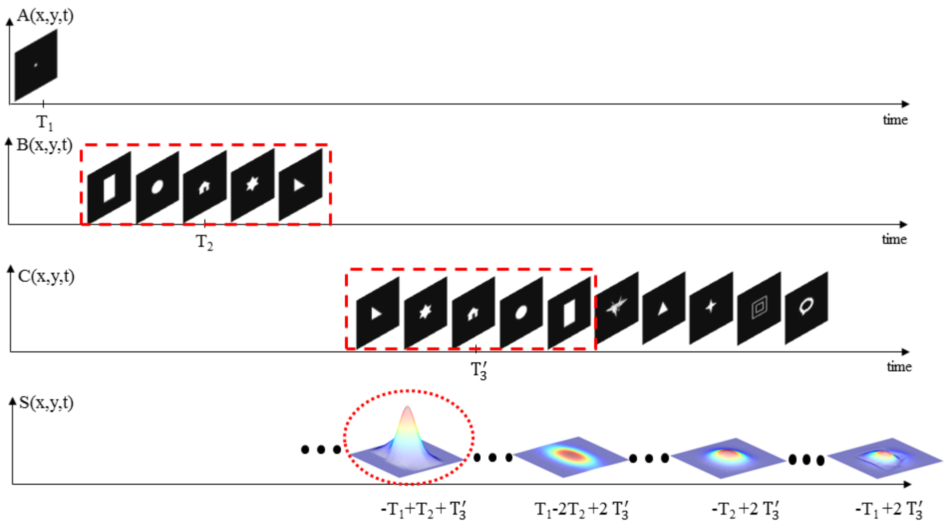


Figure 10.5: Simulation result of the Spatio Temporal Correlator using the analytical model. A(x,y,t) is the writing pulse, B(x,y,t) is the query frame set, and C(x,y,t) is the database frames. S(x,y,t) shows the results of the correlator at the detector plane at different time.

After additional delay, the database clips, C(x,y,t) are applied, with the same frame rate as the query frames.

The database clips contain ten frames, within which only first five frames match with the query clips, as shown by the dotted box. The center of the matched clip in the reference database appears at times T_3'. Now, using eqn (**10.7**), eqn (**10.8**) and eqn (**10.9**), we get the signal S(x,y,t), which has the highest peak at time: $t = T_3' + T_2 - T_1$, as shown in the encircled feature in Figure 10.5. Other peaks from the nonlinear terms appear at time $t = -T_2 + 2T_3$ and $t = -T_1 + 2T_3$. Now, the result can be more clearly interpreted if we integrate each frame of S(x,y,t) over space and plot it with respect to time, as shown in Figure 10.6. Figure 10.6(a) shows the result for the STC considered in Figure 10.5 where the query

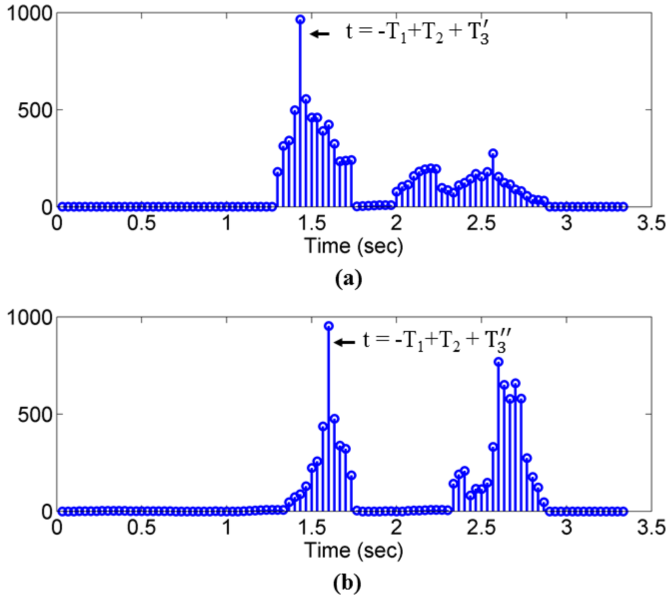


Figure 10.6: Simulation result of the STC where each frame integrated over space and plotted with respect to time.

frames match the first half of the reference frames. It is clear from Figure 10.6(a) that the highest peak occurs at time $t = T_3' + T_2 - T_1$. If the query frames occurred in the last half of the reference frames, the correlation peak would move 5 frames to time $t = T_3'' + T_2 - T_1$ (where $T_3' - T_3'' = (5/30)\mathrm{sec}$) as seen in Figure 10.6(b). Next, we describe how the STC can be operated in a manner such that the correlation window does not overlap the reference frames and also the nonlinear terms do not overlap the correlation signal window. The input and the output signals appearing in the STC within these constraints are illustrated schematically in Figure 10.7. In the STC process, the center of the input pulses A(x,y,t), B(x,y,t) and C(x,y,t) appear at times T_1, T_2 and T_3 respectively. The write pulse, A(x,y,t) consists a single frame whereas the query, B(x,y,t) and the reference, C(x,y,t) contain multiple frames. Let us assume that, the query and the reference contain X and Y number of frames, respectively, where Y = nX (assume $n \geq 1$). The peak of the correlation signal (denoted as S_1) appears at times $t = -T_1 + T_2 + T_3'$ where T_3' is time corresponding to the center of the matching clip within the reference frames. Thus, if $T_3' = T_3$, then the correlation peak will appear at time $t_0 = -T_1 + T_2 + T_3$. However, if $T_3' < T_3 (T_3' > T_3)$ then the correlation peak will appear at a time $t < t_0 (t > t_0)$. If we assume that the matching clip is fully contained within the reference frames, then it follows that the location of the correlation peak will be confined to a window of width $\sim (Y - X)\Delta t$, where Δt is the sum of the duration of each frame and the gap between the adjacent frames. However, convolution between two matching signals (i.e. the query and the matching clips), each with a duration of $\sim X\Delta t$, produces a signal with a temporal width of $\sim 2X\Delta t$. Thus, the overall window for observing the correlation signal will have a width $\sim (X + Y)\Delta t \equiv \Delta\tau_1$, centered at time $t_0 = -T_1 + T_2 + T_3$.

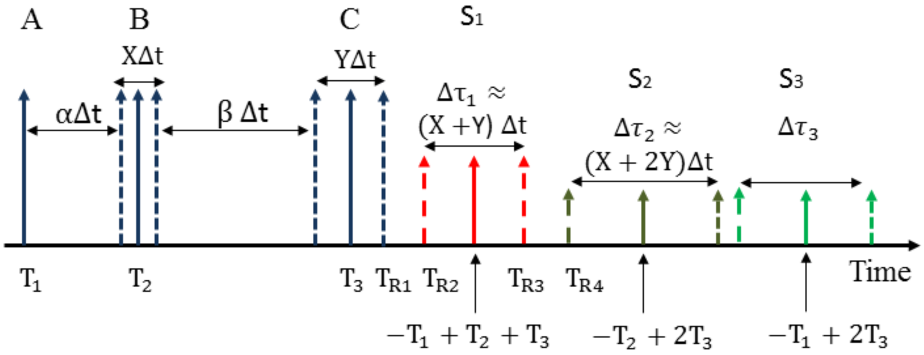


Figure 10.7: Sequence of pulses appearing in the Spatio-Temporal Correlator.

Let us denote as $\alpha\Delta t$ the time lag between the writing pulse, A(x,y,t), and the first frame in the query clip, B(x,y,t). The value of α has to be chosen to ensure that the observation window for the correlation signal does not overlap the reference frames, C(x,y,t). Let us assume that, the rightmost frame of the reference, C(x,y,t) occurs at time T_{R1} and the leftmost edge of the correlation window occurs at time T_{R2}. We thus need to ensure that $T_{R2} - T_{R1} \geq 0$. Considering $T_1 = 0$ for simplicity and without loss of generality and noting that the width of the correlation window is $\Delta\tau_1 \approx (X + Y)\Delta t$, it is easy to show that $\alpha \geq nX$ satisfies this constraint. For example, if the number of query frames is X = 3 and the number of reference frames is Y = 9, then n = 3 and α should be chosen greater than or equal to 9.

As mentioned earlier, we have investigated the contribution of additional non-linear terms in section 9.4. Specifically, in Table 9.1 of section 9.4, we have cataloged all the non-linear terms. Given that we look for signal only after applying the reference frames, the only terms of relevance are those shown in group C of this table. There are three terms here. One of these is the desired correlation signal, denoted as $S_1 = \mathrm{IFT}(\tilde{a}^*\tilde{b}\tilde{c})$ where $\tilde{a}, \tilde{b}, \tilde{c}$ are the three dimensional FT's of A(x,y,t), B(x,y,t) and C(x,y,t) respectively. The other two are: $S_2 = \mathrm{IFT}(\tilde{b}^*\tilde{c}^2)$, $S_3 = \mathrm{IFT}(\tilde{a}^*\tilde{c}^2)$ which will appear centered at $t = 2T_3 - T_2$ and $t = 2T_3 - T_1$, respectively. Using the same set of arguments we used earlier to determine the width of the observation window for the correlation signal, we can show that the width of the window for S_2 is ~ $(X + 2Y)\Delta t \equiv \Delta\tau_2$ and the width of the window of window for S_3 is $\sim (Y + 1)\Delta t \equiv \Delta\tau_3$. Thus, if we ensure that the window for S_2 does not overlap that of S_1, that will also ensure that there is no overlap between the windows for S_3 and that of S_1.

Let us denote $\beta\Delta t$ as the temporal separation between the rightmost frame of the query clip and the leftmost frame of the reference. We denote T_{R3} as the right edge of the window for S_1, and T_{R4} as the left edge of the window for S_2. We need to choose β to have a value such that $T_{R4} \geq T_{R3}$. It is easy to show that this condition is satisfied if $\beta \geq (2nX + X)$. For example, if the number of query frames is X=3 and the number of reference frames is Y = 9, then n = 3 and β should be chosen to be greater than or equal to 21 to avoid overlap between the windows for S_1 and S_2.

Chapter 11 Practical Considerations in AER

11.1 Introduction

In this chapter we address practical issues in realizing the Spatio-Temporal Correlator (STC) based Automatic Event Recognition (AER). First, we show that optically off-resonant excitation of three-level atoms in the Λ configuration is suitable for realizing the STC. Furthermore, we show how a combination of a paraffin coated, nano-porous Rb vapor cell, a holographic video disc, and a lithium niobate crystal can be used to realize the STC for practical use.

11.2 STC Architecture using Three-level Λ system

As we have mentioned earlier, a conventional two level optical transition is unsuited for the STC because of the rapid decay of the excited state. This problem can be circumvented by making use of a three-level Λ system, as shown in Figure 11.1.

A particular example of such a system consists of levels $|1\rangle$ ($5^2S_{1/2}$; F=1), $|2\rangle$ ($5^2S_{1/2}$; F=2), and$|3\rangle$ ($5^2P_{1/2}$ manifold). If the two optical fields are highly detuned, this system behave effectively as a two level system consisting of states $|1\rangle$ and $|2\rangle$, as we have shown in detail in Appendix E. The effective 1-2 transition is inhomogeneously broadened due to Doppler shifts. If the two optical beams are co- propagating, this Doppler width at room temperature is very small. However, if the beams are counter- propagating, then this Doppler width is quite large. As such, we must make use of the counter-propagating scheme if we want to maximize the bandwidth of the system. The use of this Λ system for realizing the AER requires significant modifications of the conceptual architecture shown earlier in Figure 10.2. In Figure 11.2, we show the physical implementation of the AER architecture employing the Λ transition in ^{87}Rb atoms in a vapor cell. A pulsed auxiliary beam (at frequency ω_c) that couples level $|2\rangle$ to the $5^2P_{3/2}$ manifold is used to optically pump the atoms into level $|1\rangle$ before any correlation process starts. As noted above, in order to maximize the spectral broadening of the atomic medium, the two Raman transition laser beams at frequencies ω_a and ω_b are made to be counter-propagating. The beams (at frequency ω_a) that carry the image information as well as the recording pulse are applied along the 1-3 transition, detuned by an amount, δ. The beam (at frequency ω_b) that excites the 2-3 transition is applied at all times, is also detuned by the same amount, δ. The value of δ is chosen to be much larger than the decay time of level $|3\rangle$, in order to ensure that the three-level system function effectively as a two

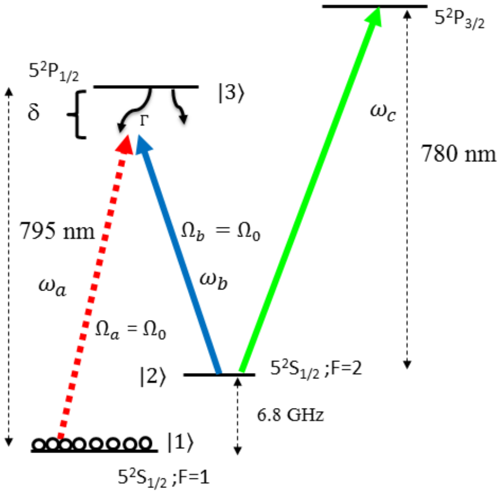


Figure 11.1: Raman interaction in a three level system

level system coupling state $|1\rangle$ to state $|2\rangle$ (see appendix E for details). The two Raman transition beams are polarized to be linear and orthogonal to each other. The optical pumping beam at frequency ω_c has the same linear polarization as that of the Raman beam at frequency ω_b. A dichroic mirror is used to combine these two frequencies, made possible by the fact that they differ in wavelengths by ~15 nm. The correlation signal appears at frequency ω_a, and passes through the second Polarizing Beam Splitter (PBS). In the detector plane, the correlation signal is detected by the detector array. The resulting voltage signals from all the detectors are integrated to produce the net signal of the AER system. Alternatively, a single detector with a large area to replace the detector array can be used to produce the net signal automatically.

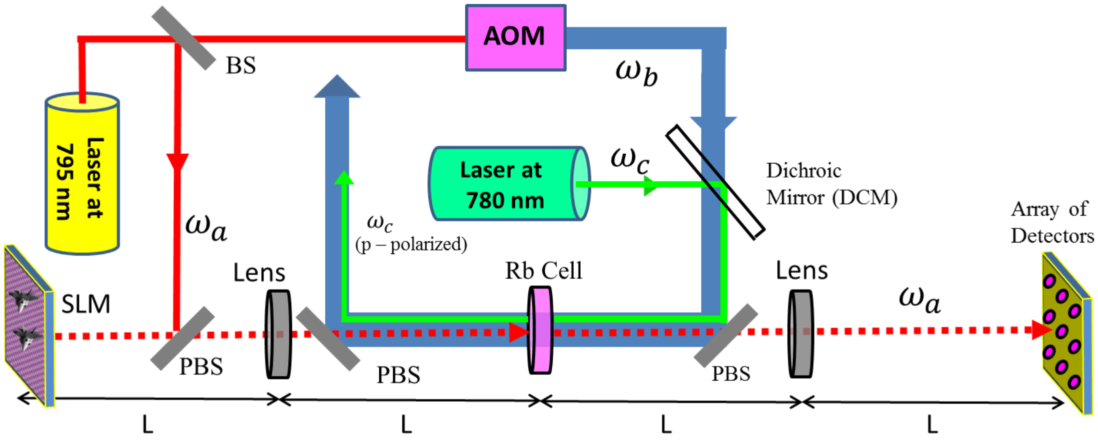


Figure 11.2:Fig 12: Schematic illustration of the physical implementation of the Automatic Event Recognition system employing vapor cell containing ^{87}*Rb atoms and neon buffer gas. [PBS = Polarizing Beam Splitter; AOM: Acousto-Optic Modulator; SLM: Spatial Light Modulator]. See text for details.*

11.3 STC Architecture using Counter-propagating scheme

In Figure 11.2 we have shown a configuration where the beam at frequency ω_b is propagating in a direction opposite to that of the image fields at frequency ω_a. As noted above, this geometry corresponds to maximization of the inhomogeneous broadening of the effective two level transition between states $|1\rangle$ and $|2\rangle$, which we denote here as Γ_{INH}. In general, for high fidelity operation of the system, it is necessary to ensure that this broadening is greater than the inverse of the duration of the shortest pulse in the data stream, which we denote here as Γ_P. On the other hand, if $\Gamma_{INH} >> \Gamma_P$, then only a small fraction (of the order of Γ_P/Γ_{INH}) contributes effectively to the correlation signal. Thus, under such a scenario, it is useful to employ a scheme where the angle of the beam at frequency ω_b (with respect to the direction of propagation of the beams at frequency ω_a) is varied to make the value of Γ_P/Γ_{INH} close to unity. Specifically, let us denote this angle as θ_{ab}, as shown in Figure 11.3. (Note that θ_{ab} =180 degrees in the configuration shown in Figure 11.2). In general, the value of Γ_{INH} as a function of θ_{ab} is given by $\Gamma_{INH} = 2[\omega_a - \omega_b \mathrm{Cos}(\theta_{ab})]u/c$, where *u* is the most probable velocity in the vapor cell, and *c* is the speed of light. Assuming a typical value of u=250 m/sec, and ignoring the small difference between the values of ω_a and ω_b, we get that $F_{INH} \equiv \Gamma_{INH}/2\pi \approx [1-\mathrm{Cos}(\theta_{ab})]*630\mathrm{MHz}$. Thus, for the configuration shown in Figure 11.2, where θ_{ab} is 180 degrees, the inhomogeneous width is ~1260 MHz. If we use a different, nearly co-propagating geometry, with θ_{ab} equaling 15 degrees for example, the inhomogeneous width would be ~21.5 MHz. For the case of exact co-propagation, we must take into account the small difference between ω_a and ω_b, and in that case we have $F_{INH} \approx 11\,kHz$. In practice, the

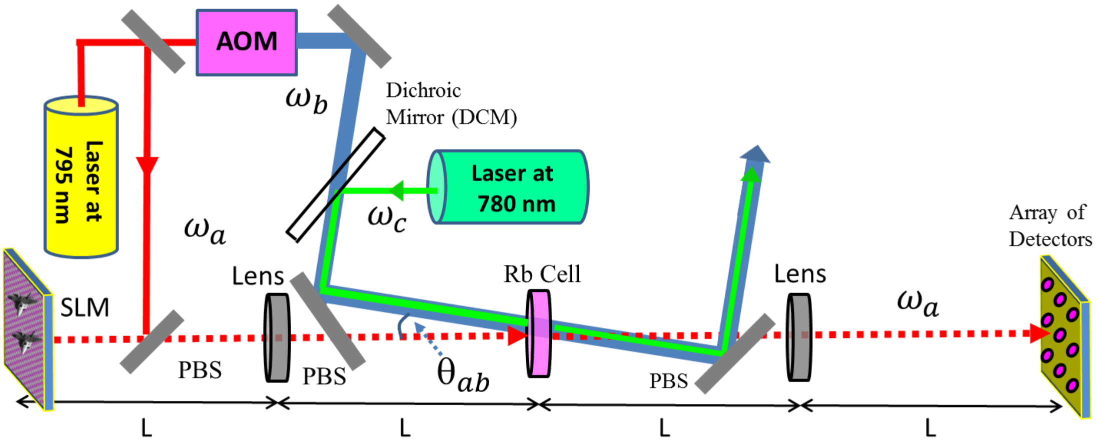


Figure 11.3: Schematic illustration of the physical implementation of the Automatic Event Recognition system employing paraffin coated porous glass vapor cell containing ^{87}Rb *atoms, with nearly co-propagating Raman interaction. [PBS = Polarizing Beam Splitter; AOM: Acousto-Optic Modulator; SLM: Spatial Light Modulator]. See text for details.*

choice of θ_{ab} would be dictated by the need to ensure that Γ_P/Γ_{INH} is close to unity in order to maximize the efficiency of the correlation process.

11.4 Image Retrieval time of the SLM

Consider next the speed with which the query clip and the video data base can be loaded into the AER system. For demonstrating the basic functionality of the AER system, one can consider making use of a ferro-electric liquid crystal based SLM. Each gray-scale frame in the video data base as well as the query clip will be converted to 512X512 image files with 8 bit representation of each pixel amplitude, in order to make them compatible with this SLM. Fundamentally, such an SLM can display images at the rate of

2500 frames per second. On the other hand, using a PCIe (Peripheral Component Interconnect express) data bus, the maximum rate at which one can transfer 8-bit (or 16-bit) deep 512X512 image frames is ~1667 frames per second (600 microseconds per frame). Thus, this will be the limiting frame sequencing rate if the SLM approach is used.

The shortest duration of each temporal pulse (for each pixel emerging from the SLM) in this case would be of the order of ~600 microseconds. The frequency (in Hz, rather than radian per second) corresponding to the inverse of this temporal width is ~10.5 kHz. However, the FT of these pulses is expected to have a spread that can encompass multiple harmonics of this frequency. Thus, a reasonable choice for the desired inhomogeneous width of the Raman transition would be ~100 kHz. From the discussion presented above, we see that this inhomogeneous width can be produced by using a relative angle θ_{ab} of ~ one degree between the beams at frequency ω_a and the beam at frequency ω_b, which follows from the expression that $F_{INH} \approx [1-\mathrm{Cos}(\theta_{ab})]*630\mathrm{MHz}$. Thus, in this case, it would be suitable to use the nearly co-propagating geometry, with this value of θ_{ab} (~one degree) as shown in Figure 11.3.

11.5 Memory time of the Atomic Medium

A key feature of the atomic medium is that it stores the spatial and temporal interference between the recording pulse and the query frames in the electro-nuclear spin coherence in the form of a coherent superposition between states $|1\rangle$ to $|2\rangle$. The lifetime of this coherence time can be ~ 1 second in a paraffin coated Rb vapor cell[57,58]. The spatio-temporal correlation process must be carried out within this time window. However, as we discuss later, this time window does not limit the maximum size of the data base video that one can search through.

Consider a situation where the query clip has a nominal duration (T_{NOM}) of 20 seconds, if played on a regular monitor. For a video frame rate (F_{VFR}) of 30 per second, this would contain 600 frames. For the SLM loading speed of 600 microseconds (defined as T_{SLM}) per frame, this clip can be loaded into the atomic medium in 360 ms, which is much shorter than the nominal duration (20 seconds) of the query clip. Thus, in the context of the AER, the retrieval duration of a video clip is given by $T_{RET} = F_{VFR}*T_{NOM}*T_{SLM}$.

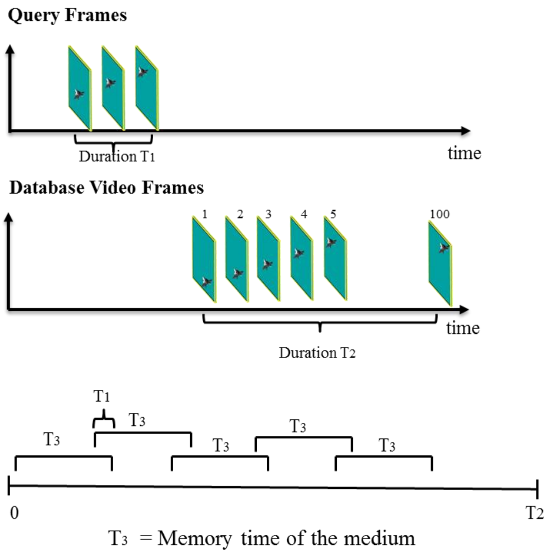


Figure 11.4: Schematic illustration of the process of the time sequencing necessary for the AER system. See text for details.

As we mentioned above, it is very important to note that the atomic coherence memory time of 1 second does not constrain the size of the database video one can search through. To see this, consider a situation where the memory time (i.e. the coherence time) is T_3, the time span of the query clip (expressed as its retrieval duration) is T_1, and the time span for the database video (expressed as its retrieval duration) is T_2. The correlator is operated for the duration T_3, during which a fraction (given by T_3/T_2) of the database has been searched. At this point, the AER system is reinitialized by using the optical pumping beam at frequency ω_c (see Figure 11.1, Figure 11.2 and Figure 11.3), and the same query clip is loaded again. The database is now loaded with a start time of (T_3-T_1), the AER is operated for another duration of T_3, and the process is repeated again, with a start time of $2T_3$-T_1, and so on, until the whole data base has been searched. This sequencing is illustrated schematically in Figure 11.4. The offset of T_1 in the start time of the database is to ensure that the AER would be able to detect the presence of the query clip even if it occurs in-between each segment searched within each memory window.

As noted earlier, the smallest loading time that can be accommodated by the atomic memory is determined by the effective inhomogeneous width (F_{INH}) of the two-photon transition coupling level $|1\rangle$ to level $|2\rangle$. When the two beams that cause this transition (at ω_a and ω_b) are counter-propagating, as shown in Figure 11.2, the inhomogeneous width becomes ~1.26 GHz, and this can accommodate a loading time as fast as 1 ns.

11.6 Implementing AER employing a vapor cell, a holographic video disc, and a lithium niobate crystal:

One possible way to make use of such a short loading time is to employ a very fast SLM. In recent years, significant progress has been made in developing technologies that could lead to the development of high speed SLMs. One example of such a development is a quantum well structure[40] that can be modulated very fast (>50 GHz) and efficiently to produce phase modulation. Using interferometric configurations for each pixel, this technology can be used to realize an SLM that can provide a frame loading rate approaching one per microsecond or even faster. However, in order to achieve this frame loading rate in practice, it is necessary to create a computer architecture and a data bus that can retrieve a page of data in a microsecond and transfer it to the SLM. While in principle it is possible to build such a system, it does

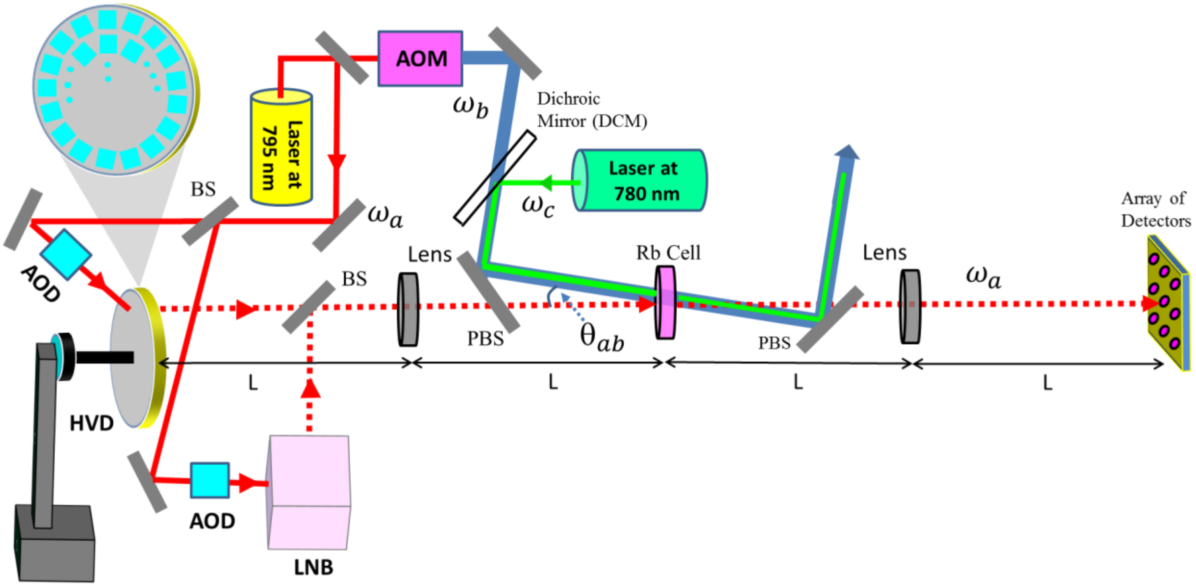


Figure 11.5: Schematic illustration of the physical implementation of the Automatic Event Recognition system employing paraffin coated porous glass vapor cell containing ^{87}Rb atoms, with nearly co-propagating Raman interaction, and using a holographic video disc and a lithium niobate crystal for rapid loading of data base and query videos. [PBS = Polarizing Beam Splitter; AOM: Acousto-Optic Modulator; AOD: Acousto-Optic Deflector; HVD: Holographic Video Disc; LNB: Lithium Niobate Crystal]. See text for details.

not currently exist, to the best of our knowledge. Thus, in order to make use of the fast loading time allowed by the atomic medium, one can explore the use of holographic techniques. The basic approach to be employed for this is as follows. The gray-scale video database will be converted to 512X512 image files with 8 bit representation of each pixel amplitude. These images will be recorded in a holographic video disc (HVD), of the type described in chapter 2. The recording will be done by transferring each image to an SLM, and interfering it with a reference plane wave. At each location of the disc, angle multiplexing of the reference plane wave will enable the storage of many images. As described in chapter 2, using a disc with a diameter of 15 cm, thickness of 1 cm, and 2000 segments, it should be able to store as many as 10 million images. For a video frame rate of 30 frames/second, this would correspond to ~92.5 hours of video data. A more modest system with 500 segments would allow storage of 2.5 million images, corresponding to more than 23 hours of video data. Similarly, the gray scale query clips will also be converted to 512X512 image files with 8 bit representation of each pixel amplitude. Since the query clips have to be periodically refreshed, based on the event of interest, one can make use of a

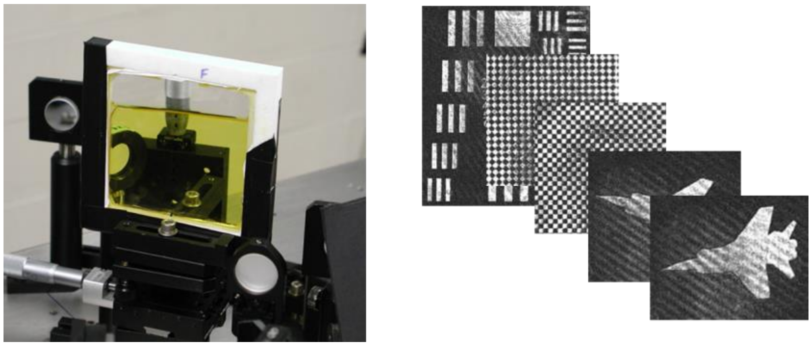


Figure 11.6: Left: Picture of a PQ (phenanthrenequinone) dye loaded PMMA (polymethyl methalycrate) based 10cmX10cmX1cm holographic recording sample made at the laboratory of the PI. Right: A set of angle multiplexed imaged stored and recalled from one location in this sample.

rewritable memory, namely a lithium niobate crystal[33, 59] .As we have noted earlier, it is easily possible to record and retrieve as many as 5000 images in a single location of such a crystal. For a frame rate of 30 per second, this would allow the storage of a single query clip of nearly 3 minutes in duration. In practice, it is unlikely that one would be interested in identifying a match to such a long clip. It is more likely that a typical clip of interest would be no more than 30 seconds long (physical duration). Thus, as many as six clips of interest can be loaded into the crystal at one time. Of course, when necessary, these can be erased by exposing the crystal to ultra-violet radiation, and then record new query clips.

During the operation of the AER system, the data read from the crystal will be made co-linear with the data read from the holographic video disc using a 50/50 beam splitter, as illustrated in Figure 11.5. To access different frames within the query clip, an acousto-optic deflector (AOD) can be used to scan the read angle. Typical speed of operation for such a device is about 20 MHz, with the number of resolvable angles being as high as 5000. Thus, it should be possible to retrieve images from the query clip at the rate of 1 frame per microsecond (which is much less than the maximum possible rate of 20 frames per microsecond corresponding to a bandwidth of 20 MHz). To access different frames from the holographic video disc (i.e., the video data base), at a given location on the disc, again one can use an acousto-optic deflector, with the same retrieval rate (i.e., about 1 frame per microsecond). To access different locations sequentially, a combination of a motorized rotator and a motorized translator can be used. For a 1 cm thick disc, the Bragg angular selectivity can be about 0.002 degrees. The angular range covered by a well-designed AOD can be ~10 degrees, so that about 5000 images can be read-out from one location. This is much larger than the 900 images in a query clip of 30 seconds of physical duration.

Thus, using the sequencing process illustrated in Figure 11.4, the AER would be operated for a duration that is long enough to load the 5000 images from one location, corresponding to ~ 5 msec (i.e., T_3=5 msec in the notation used in Figure 11.4). The atomic medium will then be re-initialized by using the optical pumping beam at frequency ω_c, while the rotation stage is used to move the system to the next location on the disc. The AER operation will then be repeated with the data retrieved from the new location. Once all locations at the same radial distance from the center of the disc have been searched through, the translation stage can be used to enable the AOD to access data from the locations at a different (adjacent) radial distance. If we assume conservatively that moving from one location to another can be accomplished in a stable manner in about 30 msecs, then all 500 locations can be searched in less than 20 seconds.

To accommodate operation at the speed of 1 frame per microsecond within each sequence, it is necessary to ensure that the inhomogeneous broadening is sufficiently large. Using the same arguments we made earlier, a safe value of F_{INH} will have to be about 10 MHz. From the expression that $\mathrm{F_{INH}} \approx [1-\mathrm{Cos}(\theta_{ab})]*630\mathrm{MHz}$, one will need to operate the AER system such that the angle θ_{ab} between the beam at ω_b and the image propagation direction will be about 10.2 degrees.

In order to realize this approach it is necessary to make holographic video discs with a thickness of 1 cm. Our group at Northwestern University developed material for making such thick holographic plates for data storage and other applications[60]. Briefly, this material is made of PMMA (polymethyl methalycrate), doped with the dye called phenanthren-quinone. The idea of a PMMA-based dye-doped photopolymer material for holographic materials was originally proposed and studied by Veniaminov et al. at the Vavilov Optical Institute[61,62,63,64,65]. Further development for various applications of this medium has continued at Northwestern University under the group of the PI[3,6,60,66,67,68], as well as at the California Institute of Technology[69,70] and at the National Chiao Tung University (NCTU)[71,72,73]. Very recently, gratings made with this medium has also been used for making DFB diode lasers[74]. Figure 11.6 shows a picture of a 10cm X 10cm X 1cm sample we made with this material, along with holographic images stored and recalled from it. One can use this technology to make a disc of 1 cm thickness and 15 cm diameter .

For a typical imaging system, the spatial resolution is about 10 um X 10 um. Thus, it is important to ensure that the atoms stay localized to such an area during each stage of the correlation sequence shown in Figure 11.4. In order to meet this constraint, one can make use of a vapor cell consisting of nano-porous glass[75,76,77]. Specifically, the process will make use of porous glass (PG) which is about 96% silica, and contains randomly inter-connected pores with a mean diameter that can be controlled to be anywhere from 5 nm to 40 nm. The technique for creating such a porous glass is well-known[77], and commercial versions of it are currently available from Schott, Inc., with various pore diameters. Since the pore size is much smaller than the wavelengths of interest (795 nm and 780 nm), the PG is transparent, does not produce significant scattering, and is of high optical quality. . Specifically, as shown in ref. 77, the transmissivity for wavelengths longer than 660 nm is ~92%. This is essentially the same as the transmissivity of conventional glass without antireflection coating, since the Fresnel reflection from each surface is about 4%. A typical PG plate available from Schott, Inc., has a dimension of the order of 2 cm X 2 cm X 1 mm. Using a simple variation of the technique described in ref. 58, the inner walls of the pores in the PG can be coated with paraffin in order to ensure that the Rb vapor does not lose coherence when they collide with the walls of the pores[57,58] . During the operation of the correlator, the PG should be exposed to radiation from a diode laser at 808 nm, applied at an angle with respect to the beams at 780 nm and 795 nm used for the correlation process. The presence of the beam at 808 nm would prevent the formation of Rb clusters, and ensure that the pores contain Rb in the vapor phase only[75,76].

11.7 Detection of the Correlation Signal

For detection of the correlation signal, one can make use of an image intensified high CCD camera, which has a speed of 200 psec. The signal captured by the camera can be acquired rapidly using a PCIe bus. As noted earlier, the signature of a positive match for a clip is given by the signal integrated over the detector array. As such, the signals captured from the camera can be integrated by the computer, and displayed on the screen. Alternatively, to eliminate the speed bottleneck imposed by this integration process, one can use a single detector with a large area to produce the signal. In order to characterize the performance of the AER system as well as to compare with theoretical models, one can choose to study the spatial distributions of the signals in all the frames captured by the computer. Of course, a detector array must be used in this case.

11.8 Conclusion

We describe an automatic event recognition (AER) system based on a three-dimensional spatio-temporal correlator (STC) that combines the techniques of holographic correlation and photon echo based temporal pattern recognition. The STC is shift invariant in space and time. It can be used to recognize rapidly an event (e.g., a short video clip) that may be present in a large video file, and determine the temporal location of the event**.** Using polar Mellin transform, it is possible to realize an STC that is also scale and rotation invariant spatially. Numerical simulation results of such a system are presented using quantum mechanical equations of evolution. For this simulation we have used the model of an idealized, decay-free two level system of atoms with an inhomogeneous broadening that is larger than the inverse of the temporal resolution of the data stream. We show how such a system can be realized by using a lambda-type three level system in atomic vapor, via adiabatic elimination of the intermediate state. We have also developed analytically a three dimensional transfer function of the system, and shown that it agrees closely with the results obtained via explicit simulation of the atomic response. The analytical transfer function can be used to determine the response of an STC very rapidly. In addition to the correlation signal, other nonlinear terms appear in the explicit numerical model. These terms are also verified by the analytical model. We develop a way to detect the correlation signal without interference from these additional nonlinear terms. We also show how such a practical STC can be realized using a combination of a porous-glass based Rb vapor cell, a holographic video disc, and a lithium niobate crystal.

Chapter 12 Conclusion

Target identification and tracking is important in many defense and civilian applications. Optical correlators provide a simple technique for fast verification and identification of data. The primary limitation of such Conventional Holographic Correlators (CHC) is the poor nature of the material used for recording. We develop the concept and design of a Hybrid Opto-electronic Correlator (HOC) using only photo-detectors, Spatial Light Modulator (SLM) and VLSI chips that overcomes the material problems of the CHC. In the HOC, the amplitude and phase information of the query as well as the reference image is recorded with photo-detectors via interference with plane waves. The output of such an HOC has four terms: two convolution signals and two cross-correlation signals. Here, we show through simulation and experimental results, that the correlation signal depends non-trivially on both the *relative* and the *absolute* phases of these two plane waves at the planes of interference. We develop a technique for controlling and optimizing the output signal, such that the convolution terms can be eliminated, by using a single servo, aided by repeated calibrations with known objects and reference images, interlaced with the actual correlator operation. To achieve the ultimate speed of such a correlator, we also propose an integrated graphic processing unit which would perform all the electrical processes in a parallel manner. We show that the HOC is capable of detecting objects in a scale and rotation invariant manner, along with the shift invariance feature, by incorporating Polar Mellin Transform (PMT). Furthermore, we also demonstrate how to carry out shift, rotation and scale invariant detection of multiple matching objects simultaneously, a process previously thought to be incompatible with the PMT based correlators.

We also develop the concept and design of an Automatic Event Recognition (AER) System based on a three-dimensional Spatio-Temporal Correlator (STC), that combines the techniques of holographic correlation and photon echo based temporal pattern recognition to recognize a video-clip contained in a video file using atoms stored in a porous-glass material. By employing the nonlinear properties of inhomogeneous broadened atomic medium we show that it is possible to realize an AER system that can recognize rapidly the occurrence of events, the number of events, and the occurrence times. To model the response of such a system, one requires solving the Schrodinger Equation (SE), which is a computationally extensive task. We develop an analytical model to find the response of the STC and show that the analytical model agrees closely with the results obtained via explicit numerical simulation, but at a speed that is many orders of magnitude faster than the numerical model. We also show how such a practical STC can be realized using a combination of a porous-glass based Rb vapor cell, a holographic video disc, and a lithium niobate crystal.

Appendix A Translation Invariance of Fourier Transform in One and Two Dimensions

Consider first the Fourier Transform (FT), $\tilde{U}\left(k_x\right)$ of a one dimensional function, $U\left(x\right)$:

$$\tilde{U}\left(k_x\right) \equiv \int_{-\infty}^{\infty} dx U(x) e^{-jk_x x} \tag{A.1}$$

Consider now a different function, $V\left(x\right)$, which is a shifted version of $U\left(x\right)$:

$$V\left(x\right) \equiv U(x-a) \tag{A.2}$$

The FT , $\tilde{V}\left(k_x\right)$, of this function is then given by:

$$\begin{aligned} \tilde{V}\left(k_x\right) &= \int_{-\infty}^{\infty} dx V(x) e^{-jk_x x} = \int_{-\infty}^{\infty} dx U(x-a) e^{-jk_x x} \\ &= \int_{-\infty}^{\infty} dx' U(x') e^{-jk_x\left(x'+a\right)} = e^{-jk_x a} \int_{-\infty}^{\infty} dx' U(x') e^{-jk_x x'} = e^{-jk_x a} \tilde{U}\left(k_x\right) \end{aligned} \tag{A.3}$$

Thus, we see that the amplitude of the FT is translation invariant.

This set of arguments apply equally well for two-dimensional functions. Consider the FT, $\tilde{U}\left(k_x, k_y\right)$ of a two dimensional function, $U\left(x, y\right)$:

$$\tilde{U}\left(k_x, k_y\right) \equiv \int_{-\infty}^{\infty} dy \int_{-\infty}^{\infty} dx U(x, y) e^{-jk_x x} e^{-jk_y y} \tag{A.4}$$

It is customary to express this FT using the vectors $\vec{\rho} \equiv x\hat{x} + y\hat{y}$ and $\vec{k} \equiv k_x \hat{x} + k_y \hat{y}$:

$$\tilde{U}\left(\vec{k}\right) \equiv \int_{-\infty}^{\infty} d\vec{\rho}U(\vec{\rho})e^{-j\vec{k}\cdot\vec{\rho}} = \tilde{U}\left(k_x, k_y\right) \equiv \int_{-\infty}^{\infty} dy \int_{-\infty}^{\infty} dx U(x,y) e^{-jk_x x} e^{-jk_y y} \tag{A.5}$$

Consider now a different function, $V\left(\vec{\rho}\right)$, which is a shifted version of $U\left(\vec{\rho}\right)$:

$$V\left(\vec{\rho}\right) \equiv U(\vec{\rho} - \vec{S}) \tag{A.6}$$

where $\vec{S} \equiv a\hat{x} + b\hat{y}$ is the shift vector. The FT, $\tilde{V}\left(\vec{k}\right)$, of this function is then given by:

$$\tilde{V}\left(\vec{k}\right) \equiv \int_{-\infty}^{\infty} d\vec{\rho}V(\vec{\rho})e^{-j\vec{k}\cdot\vec{\rho}} = e^{-j\vec{k}\cdot\vec{S}} \int_{-\infty}^{\infty} d\vec{\rho}\,' U(\vec{\rho}\,')e^{-j\vec{k}\cdot\vec{\rho}\,'} = e^{-j\vec{k}\cdot\vec{S}} \tilde{U}\left(\vec{k}\right) \tag{A.7}$$

Again, we see that the amplitude of the FT is translation invariant.

Appendix B Generation of Two-dimensional Fourier Transform Using a Lens

Consider the configuration, shown in figure B.1 , where the distance between the input plane and the output plane is twice the focal length, *L*, of the lens placed at the center. Under the paraxial approximation, and assuming that the lens diameter is much larger than the lateral extent of the images, the image at the output plane, $U_O\left(x_2, y_2\right)$, is proportional to the Fourier Transform of the input image at the input plane, $U_I\left(x_1, y_1\right)$:

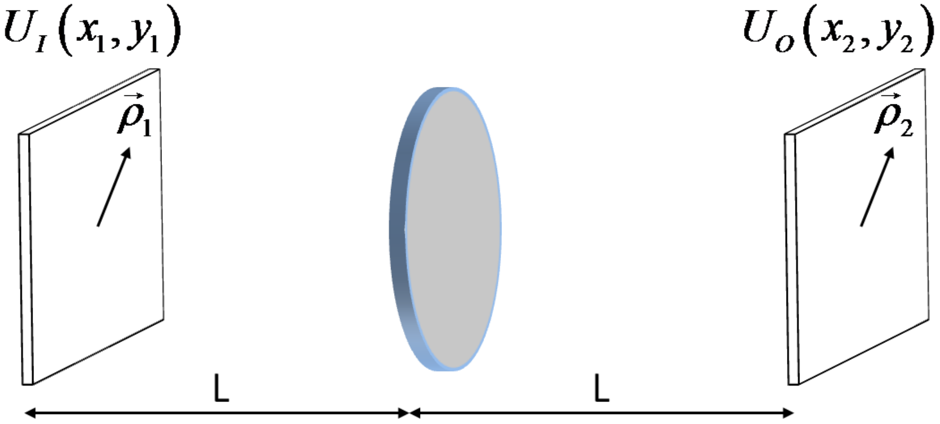


Figure B.1: Schematic illustration of Fourier Transformation by a Lens

$$U_O\left(x_2, y_2\right) = \frac{e^{j2kL}}{j\lambda L} \cdot \tilde{U}_I\left(\frac{kx_2}{L}, \frac{ky_2}{L}\right) \tag{B.1}$$

where λ is the wavelength and $k = 2\pi / \lambda$ is the wave number of the light used. The function $\tilde{U}_I\left(k_x, k_y\right)$, is the two dimensional FT of the input image, $U_I\left(x_1, y_1\right)$, as defined in Appendix A.

Appendix C Basic Aspects of a Holographic Memory

C.1: Thin Holographic Memory:

The architecture shown in **Error! Reference source not found.** can also be used to illustrate the basic process behind holographic image storage and recall. The first step in the process is the same as what is described in Chapter 1, producing a photographic plate with a transmission coefficient given by eqn. (**2.6**). Consider now a situation where this plate is illuminated by a plane wave, at the same angle as that used for the writing beam described by eqn. (**2.2**). We can write this read beam as:

$$U_R\left(x_M, y_M\right) = U_{R0} e^{-jk_\phi y_M}; \quad k_\phi = k \sin\phi \tag{C.1}$$

A simple analysis, taking into account the Fourier Transforming property of the second lens, shows that term in eqn (**2.6**) proportional to t_4 reproduces a copy of the original image, U_{A1}, in the output plane, centered around $(x_2 = 0, y_2 = 0)$. The other terms in eqn. (**2.6**) also produce signals, but they do not overlap with this retrieved image if ϕ is chosen to be sufficiently large. Of course, there are many other geometries for storage and recall of holographic images.

C.2: Angle-Multiplexed Thick Holographic Memory:

If the thin photographic plate is replaced by a ***thick holographic substrate***, it is then possible to store many images at the same location, using angular multiplexing. This is due to the sharp angular selectivity of Bragg scattering from a thick grating, which varies inversely with the

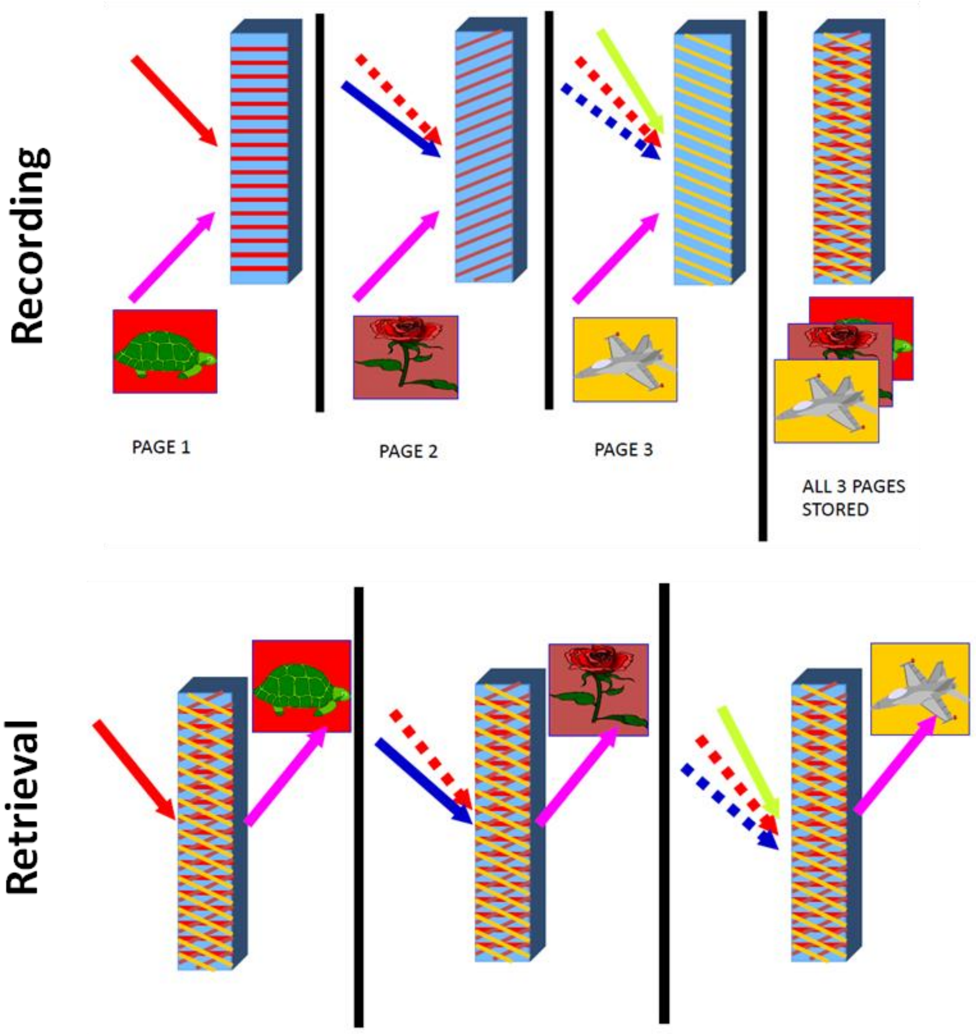


Figure C.1 : Illustration of angle-multiplexed recording and retrieval of images in a thick holographic substrate.

thickness of the substrate. Let us denote the Bragg selectivity bandwidth as $\Delta\phi_B$. The recording process of multiple images then proceeds as follows. First, image 1 is recorded by using a writing beam incident at an angle of ϕ_1. Then, the angle of the writing beam is changed to $\phi_2 = \phi_1 + \Delta\phi_B$, and the second image is applied to the input plane for recording, and so on. During the readout process, different images are read out by applying the read beam at different angles. Of course, in order to avoid cross-talks between the images, it is necessary to ensure that the Bragg selectivity bandwidth is larger than the spatial-frequency bandwidth of each of the stored images. The recording and retrieval process is illustrated schematically in figure C.1, in a simplified view without the lens.

Appendix D Relevant Properties of Mellin Transforms

D.1 Basic Definition of Mellin Transform:

In one dimension, the Mellin Transform, $M(z)$, of a function $f(x)$, is defined as:

$$M(z) \equiv \int_0^\infty dx\, f(x) x^{z-1} \tag{D.1}$$

where z is in general a complex variable. The Inverse Mellin Transform is given by:

$$f(x) = \frac{1}{2\pi j} \int_{c-j\infty}^{c+j\infty} dz M(z) x^{-z} \tag{D.2}$$

where c is a real number within the strip of analyticity.Some simple Mellin transforms are shown in the Table below:

Table D.1: Summary of Mellin Transforms

$f(x)$	$M(z)$
$\delta(x-a)$	a^{z-1}
$u(x-a)$ (step function, $a>0$)	$-\frac{a^z}{z}$
$xu(x)$	$-1/[z+1]$
$x^{-a}u(x)$	$-1/[z-a]$

D.2 Mellin Transform along complex axes:

In one dimension, the Mellin Transform, $M(-ju)$, of a function $f(x)$, is defined as:

$$M(-ju) \equiv \int_0^\infty dx\, f(x) x^{-ju-1} \tag{D.3}$$

In two dimension, the Mellin Transform, $M(-ju,-jv)$, of a function $f(x,y)$, is defined as:

$$M(-ju,-jv) \equiv \int_0^\infty dx \int_0^\infty dy\, f(x,y) x^{-ju-1} y^{-jv-1} \tag{D.4}$$

D.3 Scale Invariance of Complex-Axes Mellin Transform:

It is easy to show that amplitude of the Mellin Transform along complex axes is scale invariant. Specifically, if $M_1(-ju,-jv)$ is the Mellin Transform of a function $f_1(x,y)$, and $M_2(-ju,-jv)$ is the Mellin Transform of a scaled function $f_2(x,y)= f_1(xa,ya)$, then:

$$M_2(-ju,-jv)=a^{-ju-jv}M_1(-ju,-jv) \tag{D.5}$$

In the subsequent subsections of this appendix, we restrict our discussions to one dimension for transparency and simplicity; extension to two dimensions is easy to see. In addition, we will use the compact notation that:

$$m(u)\equiv M(-ju)=\int_0^\infty dx\, f(x)x^{-ju-1} \tag{D.6}$$

D.4 Simple Relation between Mellin Transform and Fourier Transform:

Consider a function, $f(x)$, with its Mellin Transform given by a $m(u)$. Consider now a change of coordinates from x to ξ, so that

$$x=e^{\xi} \quad \Leftrightarrow \quad \xi=\ln(x) \tag{D.7}$$

We now define a new function:

$$g(\xi)= f(e^{\xi}) \tag{D.8}$$

Then, it is easy to show, from the definition of Mellin Transform, that:

$$m(u) = \int_{-\infty}^{\infty} g(\xi) e^{-ju\xi} d\xi \tag{D.9}$$

Thus, the complex-axes Mellin Transform of the function $f(x)$ is just the Fourier Transform of the function $g(\xi) = f(e^{\xi})$.

Appendix E Three Level System Representation

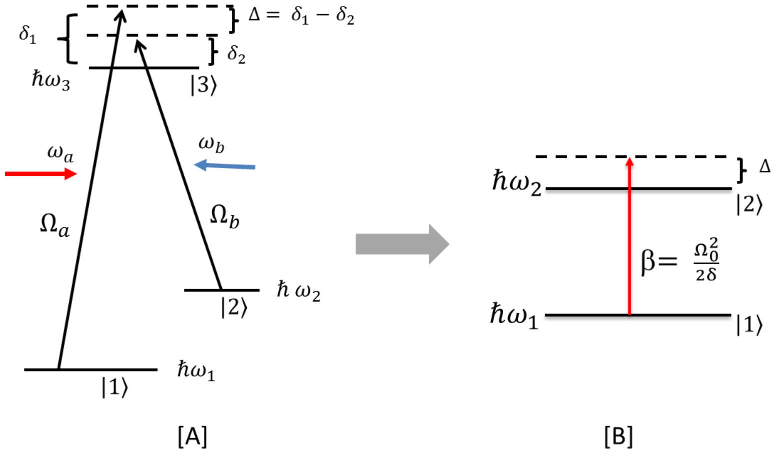


Figure E.1: [A] Energy diagram of the Λ system under stimulated Raman excitation. Here the excited state $|3\rangle$ *is short lived and the two low-lying ground states* $|1\rangle$ *and* $|2\rangle$ *are long lived. The light at frequency* ω_a *couples* $|1\rangle$ *to* $|3\rangle$*, while the counter propagating light at* ω_b *couples* $|2\rangle$ *to* $|3\rangle$*. [B] The effective two level system that results upon making the adiabatic approximation.*

Figure E.1[A] illustrates schematically the Λ system under stimulated Raman excitation. Here the excited state $|3\rangle$ is short lived and the two low-lying ground states $|1\rangle$ and $|2\rangle$ are long lived. The light at frequency ω_a couples $|1\rangle$ to $|3\rangle$, while the light at ω_b couples $|2\rangle$ to $|3\rangle$ as shown. Both couplings are electric-dipole interactions, whose strengths are given by the Rabi frequencies

$$\Omega_a = \frac{\mu_{13}.E_1}{\hbar};\Omega_b = \frac{\mu_{23}.E_2}{\hbar}; \tag{E.1}$$

where μ is the dipole moment operator matrix element for the corresponding transition. The laser detunings $\delta_1 = \omega_a - (\omega_3 - \omega_1)$ and $\delta_2 = \omega_b - (\omega_3 - \omega_2)$ are used to define the difference detuning as $\Delta = \delta_1 - \delta_2$ and the common mode detuning as $\delta = \frac{1}{2}(\delta_1 + \delta_2)$, where $\hbar\omega_1, \hbar\omega_2$ and $\hbar\omega_3$ are the energies of the three levels. Finally, the individual decay rates from state $|3\rangle$ to states $|1\rangle$ and $|2\rangle$ are Γ_{31} and Γ_{32}, respectively, and the total decay rate of state $|3\rangle$ is given by $\Gamma = \Gamma_{31} + \Gamma_{32}$. Here, we assume that $\Gamma_{31} = \Gamma_{32} = \Gamma/2$. In the atomic states basis, the Hamiltonian for the stimulated Raman interaction under the electric dipole and rotating wave approximations is given by:

$$H = \hbar \begin{bmatrix} \omega_1 & 0 & \Omega_a/2\,e^{i\omega_a t} \\ 0 & \omega_2 & \Omega_b/2\,e^{i\omega_b t} \\ \Omega_a/2\,e^{-i\omega_a t} & \Omega_b/2\,e^{-i\omega_b t} & \omega_3 \end{bmatrix} \tag{E.2}$$

Similar to what we showed in section 9.3, we can transform the Hamiltonian to the rotating wave basis by using the following matrix[51,52]: $Q = \begin{bmatrix} e^{i\alpha t} & 0 & 0 \\ 0 & e^{i\beta t} & 0 \\ 0 & 0 & e^{i\gamma t} \end{bmatrix}$ where $\alpha = \omega_1 - \frac{\Delta}{2}; \beta = \omega_2 + \frac{\Delta}{2}$. The Hamiltonian can then be expressed as:

$$\tilde{H} = \hbar\begin{bmatrix} \Delta/2 & 0 & \Omega_a/2 \\ 0 & -\Delta/2 & \Omega_b/2 \\ \Omega_a/2 & \Omega_b/2 & -\delta \end{bmatrix} \tag{E.3}$$

If the effect of decay from state $|3\rangle$ is to be taken into account, then we must use the density matrix approach to evaluate the response. However, we limit our system to the condition that $\delta \gg \Gamma, \delta \gg \Omega_0$. Under these conditions, the effect of decay from level 3 can be ignored[51,52] and using the Hamiltonian of eqn E.3, the amplitude equation can be written as*:*

$$\begin{bmatrix} \dot{\tilde{c}}_1(t) \\ \dot{\tilde{c}}_2(t) \\ \dot{\tilde{c}}_3(t) \end{bmatrix} = -i\begin{bmatrix} \Delta/2 & 0 & \Omega_a/2 \\ 0 & -\Delta/2 & \Omega_b/2 \\ \Omega_a/2 & \Omega_b/2 & -\delta \end{bmatrix}\begin{bmatrix} \tilde{c}_1(t) \\ \tilde{c}_2(t) \\ \tilde{c}_3(t) \end{bmatrix} \tag{E.4}$$

For simplicity consider the case where $\Omega_a = \Omega_b = \Omega_0$. We can make the adiabatic approximation ($\dot{\tilde{C}}_3 \approx 0$)[51,52], thus yielding the following relations:

$$\dot{\tilde{c}}_1(t) = \left(-i\frac{\Delta}{2} - i\frac{\Omega_0^2}{4\delta}\right)\tilde{c}_1(t) - i\frac{\Omega_0^2}{4\delta}\tilde{c}_2(t)$$

$$\dot{\tilde{c}}_2(t) = \left(i\frac{\Delta}{2} - i\frac{\Omega_0^2}{4\delta}\right)\tilde{c}_2(t) - i\frac{\Omega_0^2}{4\delta}\tilde{c}_1(t)$$

Defining $\varepsilon = \Omega_0^2/4\delta$ we can thus write:

$$\tilde{\tilde{H}}' = \hbar\begin{bmatrix} \Delta/2+\varepsilon & \varepsilon \\ \varepsilon & -\Delta/2+\varepsilon \end{bmatrix}; \left|\tilde{\Psi}\right\rangle = \begin{bmatrix} \tilde{c}_1(t) \\ \tilde{c}_2(t) \end{bmatrix} \tag{E.5}$$

where the reduced two level system satisfies the Schroedinger equation: $i\hbar \frac{\partial \left|\tilde{\Psi}'\right\rangle}{\partial t} = \tilde{\tilde{H}}' \left|\tilde{\Psi}'\right\rangle$. Let us further define: $|\psi'\rangle \equiv e^{i\theta t} |\tilde{\psi}'\rangle = \begin{bmatrix} C_1' \\ C_2' \end{bmatrix}$.Using this transformation we can find an effective two level Schrodinger equation of the form: $i\hbar \partial \left|\Psi'\right\rangle / \partial t = H' \left|\Psi'\right\rangle$; where $H' = \tilde{\tilde{H}}' - \hbar\theta$ $= \hbar \begin{bmatrix} \Delta/2 + \varepsilon - \theta & \varepsilon \\ \varepsilon & -\Delta/2 + \varepsilon - \theta \end{bmatrix}$; If we choose $\theta = \frac{\Delta}{2} + \varepsilon$, the Hamiltonian can be written as:

$$H' = \hbar \begin{bmatrix} 0 & \beta/2 \\ \beta/2 & -\Delta \end{bmatrix} \qquad \text{(E.6)}$$

where, $\beta = \frac{\Omega_0^2}{2\delta}$; The Hamiltonian of eqn. F.6 represents an effective two level transition between $|1\rangle$ and $|2\rangle$ with a Rabi frequency β and a detuning Δ, as shown in figure E.1[B].

The two level system used in section 9.3 now can be viewed as being realized by this reduced system. It should be noted that this effective system involves only the two metastable ground states, and does not include any decay. This justifies the model we used in section 9.3. Of course, collisional processes cause incoherent exchange of populations between states $|1\rangle$ and $|2\rangle$, and decay of the coherence between these two states. As we have noted in the main body of the thesis, the time constant for these processes can be as long as 1 second when paraffin coated vapor cells are used. If the duration of the correlation process is limited to a time much shorter than this time scale, the use of a two level system without any decay or dephasing is justified. It should be noted again that this time scale does not limit the maximum size of the video data base that can be searched, as explained in the main body of the thesis.

We also note that for this reduced two level system, the doppler broadening is manifested in the velocity dependence of Δ. Thus, the effective Doppler broadening is given by the difference (sum) of the Doppler broadening on the two legs of the Raman transition if the lasers are co-propagating, (counter-propagating). Since the inverse of the effective Doppler broadening is a lower limit of the duration of the pulses, it is important to maximize this broadening in order to ensure the highest possible operating speed. Of course, by using a geometry where the two beam are at a variable angle, it is possible to vary the effective inhomogeneous broadening, while ensuring that it is greater than the inverse of the shortest pulse widths, in order to maximize the efficiency of the correlation process, as discussed in the context of Figure 11.3 in the main body of the thesis.

Reference

---

[1] "Shift-Invariant Real-Time Edge-Enhanced VanderLugt Correlator Using Video-Rate Compatible Photorefractive Polymer," A. Heifetz, G.S. Pati, J.T. Shen, J.-K. Lee, M.S. Shahriar, C. Phan, and M. Yamomoto, Appl. Opt. 45, 6148-6153 (2006).

[2] "Translation-Invariant Object Recognition System Using an Optical Correlator and a Super-Parallel Holographic RAM", A. Heifetz, J.T. Shen, J-K Lee, R. Tripathi, and M.S. Shahriar, , Opt. Eng. 45, 025201 (2006).

[3] "Demonstration of a simple technique for determining the M/# of a holographic substrate by use of a single exposure," H. N. Yum, P. R. Hemmer, R. Tripathi, J. T. Shen, M. S. ShahriarOpt. Lett. 29, 1784-1786 (2004).

[4] "Shared hardware alternating operation of a super-parallel holographic optical correlator and a super-parallel holographic RAM", M.S. Shahriar, R. Tripathi, M. Huq, and J.T. Shen, Opt. Eng. 43, 1856-1861 (2004).

[5] "Super-Parallel Holographic Correlator for Ultrafast Database Search," M.S. Shahriar, M. Kleinschmit, R. Tripathi, and J. shenOpt. Lett. 28, 525-527 (2003).

[6] "Ultra-high density optical data storage," M. S. Shahriar, L. Wong, M. Bock, B. Ham, J. Ludman, and P. Hemmer, Symposium on Electro-Opticsl Present and Future, Optical Society of America Book Series on Trends in Optics, and Photonics, H.Haus, ed. (Optical Society of America, 1998), pp. 97–104.

---

[7] “Angular directivity of diffracted wave in Bragg- mismatched readout of volume holographic gratings,” A. Heifetz, J. T. Shen, S. C. Tseng, G. S. Pati, J.-K. Lee, and M. S. Shahriar, Opt. Commun. 280, 311–316 (2007).

[8] “Photore-fractive two-beam-coupling nonlinear joint-transform correlator,” J. Khoury, M. C. Golomb, P. Gianino, and C. Woods, J. Opt. Soc. Am. B 11, 2167–2174 (1994).

[9] “Multiple-object detection with a chirp encoded joint transform correlator,” Q. Tang and B. Javidi, Appl. Opt. 32, 5079–5088 (1993).

[10] “Signal detection by complex spatial filtering,” A. V. Lugt, IEEE Trans. Inf. Theory 10, 139–145 (1964).

[11] “Joint transform image correlation using a binary spatial light modulator at the Fourier plane,” B. Javidi and C. Kuo, Appl. Opt. 27, 663–665 (1988).

[12] "Scale-invariant optical correlation using Mellin Transforms," D. Casasent and D. Psaltis, Optics Comm. **17**, 59-63 (1976)

[13] "Position, rotation, and scale invariant optical correlation," D. Casasent and D. Psaltis, Appl. Opt. **15**, 1795-1799 (1976)

[14] "New Optical Transforms for Pattern Recognition", D. Casasent and D. Psaltis, Proceedings of the IEEE **65**, (1977).

[15] “Rotation, scale and translation resilient watermarking for images”, C.-Y. Lin, M. Wu, J. A. Bloom, I. J. Cox, M. L. Miller, and Y. M. Lui, IEEE Transactions on Image Processing **10**, 767-782 (2001).

---

[16] "Rotation-invariant optical recognition of three-dimensional objects," J. Esteve-Taboada, J. Garcıa, and C Ferreira, Appl. Opt. **39**, 5998-6005 (2000) .

[17] "Using Fourier-Mellin-based correlators and their fractional versions in navigational tasks", D. Sazbon, Z. Zalevsky, E. Rivlina, D. Mendlovic Pattern Recognition **35**, 2993 – 2999 (2002).

[18] "Rotation and scale invariance with polar and log-polar coordinate transformation", D. Asselin and H. H. Arsenault, Opt. Commun. **104**, 391-404 (1994).

[19] "Scale invariant pattern recognition with logarithmic radial harmonic filter", J. Rosen and J. Shamir Appl. Opt. **28**, 240-244 (1989).

[20] "Quantum interference and its potential applications in a spectral hole-burning solid," B.S. Ham, P.R. Hemmer, M.K. Kim, and M.S. Shahriar, Laser Physics 9 (4): 788-796 (1999).

[21] "Demonstration of a phase conjugate resonator using degenerate four-wave mixing via coherent population trapping in rubidium," D. Hsiung, X. Xia, T.T. Grove, P. Hemmer, and M.S. Shahriar, Opt. Commun., 154, 79-82 (1998).

[22] "Multidimensional holography by persistent spectral hole burning," A. Renn, U. P. Wild, and A. Rebane, J. Phys. Chem. A 106, 3045–3060 (2002).

[23] "From spectral holeburning memory to spatial-spectral microwave signal processing," W. R. Babbitt et al., Laser Physics 24, 094002 (2014).

[24] "Picosecond multiple-pulse experiments involving spatial and frequency gratings: a unifying nonperturbational approach", K. Duppen and D.A. Wiersma, J. Opt. Soc. Am. B, 3(4) (April 1986).

---

[25] "Time Domain Optical Data Storage using Raman Coherent Population Trapping," P.R. Hemmer, M.S. Shahriar, M.K. Kim, K.Z. Cheng and J. Kierstead, Opt. Lett. 19, 296 (1994).

[26] "Frequency Selective Time Domain Optical Data Storage by Electromagnetically Induced Transparency in a Rare-earth Doped Solid," B.S. Ham, M.S. Shahriar, M.K. Kim, and P.R. Hemmer, Opt. Letts. 22, 1849 (1997).

[27] "Observation of Ultraslow and Stored Light Pulses in a Solid," A. V. Turukhin, V.S. Sudarshanam, M.S. Shahriar, J.A. Musser, B.S. Ham, and P.R. Hemmer, Phys. Rev. Lett. 88, 023602 (2002).

[28] "All Optical Three Dimensional Spatio-Temporal Correlator for Video Clip Recognition," M. S. Monjur and M.S. Shahriar, in Proceedings of the Frontiers in Optics and Laser Science Conference, Tucson, AZ, October, 2014.

[29] "Three-dimensional Transfer-function of an Inhomogeneously Broadened Atomic Medium for All Optical Spatio-Temporal Video Clip Correlation," M. S. Monjur and M.S. Shahriar, accepted to appear in Proceedings of the Frontiers in Optics and Laser Science Conference, San Jose, CA, October, 2015.

[30] "Ultrawideband coherent noise lidar range-Doppler imaging and signal processing by use of spatial-spectral holography in inhomogeneously broadened absorbers ," Li Y, Hoskins A, Schlottau F, Wagner K H, Embry C and Babbitt W R 2006 Appl Opt. 45 6409–20 (2006).

[31] "Incorporation of Polar Mellin Transform in a Hybrid Optoelectronic Correlator for Scale and Rotation Invariant Target Recognition," Mehjabin S. Monjur, Shih Tseng, Renu Tripathi, and

---

M.S. Shahriar, J. Opt. Soc. Am. A, Vol. 31, No. 6, pp. 1259-1272 (June 2014)

[32] "Cell Identification with Computational 3-D Holographic Microscopy," I. Moon, M. Daneshpanah, A. Anand and B. Javidi, Optics and Photonics News, Vol. 22, No. 6, June 2011.

[33] "Non-volatile holographic storage in doubly doped lithium niobate crystals," K. Buse, A. Adibi, and D. Psaltis, Nature 393, 665-668 (1998)

[34] "Hybrid Optoelectronic Correlator Architecture for Shift- Invariant Target Recognition", M.S. Monjur, S. Tseng, R. Tripathi, J.J. Donoghue and M.S. Shahriar, J. Opt. Soc. Am. A, Vol. 31, No. 1, pp. 41-47 (Jan 2014).

[35] This, of course, would require slight modification of the optical architecture. Specifically, for the reference channel, an additional beam splitter is inserted to produce an isolated copy of the plane wave, C, to be recorded by FPA-1C, thus obviating the need for storage of |C|2 for later subtraction. The same is done for the query channel.

[36] "A short-wave infrared nano-injection imager with 2,500 A/W responsivity and low excess noise," O. G. Memis, J. Kohoutek, W. Wu, R. M. Gelfand, and H. Mohseni, IEEE Photon. J. 2, 858–864 (2010).

[37] "Signal-to-noise performance of a short-wave infrared nanoinjection imager," O. G. Memis, J. Kohoutek, W. Wu, R. M. Gelfand, and H. Mohseni, Opt. Lett. 35, 2699–2701 (2010).

[38] P. Gray, P. Hurst, S. Lewis, and R. Meyer, Analysis and Design of Analog Integrated Circuits (Wiley, 2009).

---

[39] "InP-based phase modulators and methods for making and using for same," H. Mohseni, U.S. patent 7,064,881 (June 20, 2006).

[40] "Enhanced electro-optic effect in GaInAsP/InP three-step quantum wells," H. Mohseni, H. An, Z. A. Shellenbarger, M. H. Kwakernaak, and J. H. Abeles, Appl. Phys. Lett. 84, 1823–1826 (2004).

[41] "High-performance surface-normal modulators based on stepped quantum wells," H. Mohseni, W. K. Chan, H. An, A. Ulmer, and D. Capewell, Proc. SPIE 5814, 191–198 (2005).

[42] "Tunable surface-normal modulators operating near 1550 nm with a high-extinction ratio at high temperatures," H. Mohseni, W. K. Chan, H. An, A. Ulmer, and D. Capewell, IEEE Photon. Technol. Lett. 18, 214–216 (2006).

[43] "Growth and characterization of InAs/GaSb photo- conductors for long wavelength infrared range," H. Mohseni, E. Michel, J. Sandven, M. Razeghi, W. Mitchel, and G. Brown, Appl. Phys. Lett. 71, 1403–1405 (1997).

[44] "Optoelectronic Correlator Architecture for Shift Invariant Target Recognition," M. S. Monjur, S. Tseng, R. Tripathi, J. Donoghue and M.S. Shahriar, J. Opt. Soc. Am. A, Vol. 31, Issue 1, pp 41-47 (2014).

[45] "New Optical Transforms for Pattern Recognition" D. Casasent and D. Psaltis, , Proceedings of the IEEE **65**, (1977).

[46] Recall that a PMT image is plotted as a function of ρ and θ. In the θ direction, any image will cover the whole range from 0 to 2π. As a result, two PMT images in two boxes adjacent in the vertical

---

direction will tend to merge into each other. This problem is circumvented by scaling each PMT image to 90% of its actual size, thus creating a guard band. This step does not affect the outcome of the correlator process.

[47] “Optical Header Recognition by Spectroholographic filter,” X. A. Shen and R. Kachru, Opt. Letts. 20, 2508 (1995).

[48] “Real-time optical waveform convolver cross correlator,” Y. S. Bai, W. R. Babbitt, N. W. Carlson, and T. W. Mossberg, Appl.Phys. Lett. 45, 714–716 (1984).

[49] “Laboratory measurement of the velocity of light,” W. P. Alford and A. Gold, Am. J. Phys. 26(7), 481–484. (1958).

[50] “Interference and the Alford and Gold Effect,” L. Mandel, J. Opt. Soc. Am. 52, 1335 (1962).

[51] “Evolution of an N-level system via automated vectorization of the Liouville equations and application to optically controlled polarization rotation,” Journal of Modern Optics, M.S. Shahriar, Ye Wang, S.Krishnamurthy, Y. Tu , G.S. Pati, and S. Tseng, Volume 61, Issue 4, pp. 351-367 (February 2014).

[52] “Dark-State-Based Three-element Vector Model for the Resonant Raman Interaction," M.S. Shahriar, P. Hemmer, D.P. Katz, A. Lee and M. Prentiss, Phys. Rev. A. 55, 2272 (1997).

[53] “From spectral hole burning memory to spatial-spectral microwave signal processing,” W. R. Babbitt et al., Laser Physics 24, 094002 (2014).

[54] N. Kartam and I. Flood, “Construction simulation using parallel computing environment,” Automation in Construction, Volume 10, Issue 1, pp 69–78 (2000).

---

[55] H. Adeli and O. Kamal, “Parallel Processing in Structural Engineering,” Chapman & Hall, 1993.

[56] J.M. Alonso, A. Arruabarrena, R. Beivide, and J.A.B. Fortes, “Evaluation of implementations of the CMB parallel simulation algorithm on distributed memory multicomputers,” Journal of Systems Architecture 44 6–7 1998 519–545.

[57] “Nonlinear Magneto-optic Effects with Ultranarrow Widths,” D. Budker, V. Yashchuk and M. Zolotorev, Phys. Rev. Lett. 81 5788 (1998).

[58] “Relaxation of Optically Pumped Rb Atoms on Paraffin-Coated Walls,” M. A. Bouchiat and J. Brossel, Phys. Rev. 147, 41(1966).

[59] “Angle-multiplexed storage of 5000 holograms in lithium niobate,” F. Mok, Optics Letters, Vol. 18, No. 11, 915 (1993).

[60] “Applications of Volume Holographic Gratings for Signal Processing,” John T. Shen, Doctoral Thesis, Department of Electrical Engineering and Computer Science, Northwestern University (2007).

[61] "Diffusional enhancement of holograms: phenanthrenequinone in polycarbonate," A. Veniaminov and E. Bartsch, Journal of Optics a-Pure and Applied Optics 4(4), 387-392 (2002).

[62] "The shape of the relaxation curve in diffusion measurements with the aid of photoinduced gratings," A. V. Veniaminov and E. Bartsch, Optics and Spectroscopy 101(2), 290-298 (2006).

[63] "Postexposure evolution of a photoinduced grating in a polymer material with phenanthrenequinone," A. V. Veniaminov, E. Bartsch, and A. P. Popov, Optics and Spectroscopy

---

99(5), 744-750 (2005).

[64] "Diffusion of phenanthrenequinone in poly(methyl methacrylate): Holographic measurements," A. V. Veniaminov and Y. N. Sedunov, Vysokomolekulyarnye Soedineniya Seriya a & Seriya B 38(1), 71-76 (1996).

[65] "Polymer and dye probe diffusion in poly(methyl methacrylate) below the glass transition studied by forced Rayleigh scattering," A. V. Veniaminov and H. Sillescu, Macromolecules 32(6), 1828-1837 (1999).

[66] "Very Thick Holographic Nonspatial Filtering of Laser Beams," J.E. Ludman, J. Riccobono, N. Reinhand, I., Yu. Korzinin, M.S. Shahriar, H.J. Caulfield, J. Fournier, and P.R. Hemmer, Optical Engineering 36, 6 (1997).

[67] "Coherent and incoherent beam combination using thick holographic substrates," M.S. Shahriar; J. Riccobono; M. Kleinschmit; J.T. Shen, Opt. Commun., 220, 1, pp. 75-83(2003).

[68] "Demonstration of a multiwave coherent holographic beam combiner in a polymeric substrate", H.N. Yum, P.R. Hemmer, A. Heifetz, J.T. Shen, J-K. Lee, R. Tripathi, and M.S. Shahriar, Optics Letter, 30, 3012-3014 (2005).

[69] "Characterization of phenanthrenequinone-doped poly(methyl methacrylate) for holographic memory," G. J. Steckman, I. Solomatine, G. Zhou, and D. Psaltis, Optics Letters 23(16), 1310-1312 (1998).

[70] "Comparison of the recording dynamics of phenanthrenequinone-doped poly(methyl methacrplate) materials," J. Mumbru, I. Solomatine, D. Psaltis, S. H. Lin, K. Y. Hsu, W. Z. Chen,

---

and W. T. Whang, Optics Communications 194(1-3), 103-108 (2001).

[71] "Experimental characterization of phenanthrenequinone-doped poly(methyl methacrylate) photopolymer for volume holographic storage," K. Y. Hsu, S. H. Lin, Y. N. Hsiao, and W. T. Whang, Optical Engineering 42(5), 1390-1396 (2003).

[72] "Analyses on physical mechanism of holographic recording in phenanthrenequinone-doped poly(methyl methacrylate) hybrid materials," Y. N. Hsiao, W. T. Whang, and S. H. Lin, Optical Engineering 43(9), 1993-2002 (2004).

[73] "Effect of ZnMA on optical and holographic characteristics of doped PQ/PMMA photopolymer," Y. N. Hsiao, W. T. Whang, and S. H. Lin, Japanese Journal of Applied Physics Part 1-Regular Papers Short Notes & Review Papers 44(2), 914-919 (2005).

[74] "Dye-doped PQ-PMMA phase holographic materials for DFB lasing," Y Gritsai, O Sakhno, L M Goldenberg, and J Stumpe, Journal of Optics, Volume 16, Number 3 (2014).

[75] "Optical recording in Rb loaded-porous glass by reversible photo-induced phase transformations," A. Burchianti, A. Bogi, C. Marinelli, E. Mariotti, and L. Moi, Optics Express, Vol. 16, No. 2, 1377 (2008).

[76] "Reversible Light-Controlled Formation and Evaporation of Rubidium Clusters in Nanoporous Silica," A. Burchianti, A. Bogi, C. Marinelli, C. Maibohm, E. Mariotti, and L. Moi, Phys. Rev. Lett. 97, 157404 (2006).

[77] "The possibility of using fullerene-saturated porous glasses for the optical limitation of laser radiation," I. M. Belusova, E. A. Gavronskaya, V. A. Grigor'ev, A. G. Skobelev, O. V. Andreeva,

---

I. E. Obyknovennaya, and A. S. Cherkasov, J. Opt. Technol. 68, 882–884 (2001).