

# Demonstration of shift, scale, and rotation invariant target recognition using the hybrid opto-electronic correlator

JULIAN GAMBOA,<sup>1,\*</sup> MOHAMED FOUDA,<sup>1</sup> AND SELIM M. SHAHRIAR<sup>1,2</sup>

<sup>1</sup>*Department of EECS, Northwestern University, Evanston, IL 60208, USA*

<sup>2</sup>*Department of Physics and Astronomy, Northwestern University, Evanston, IL 60208, USA*

\*[JulianGamboa2023@u.northwestern.edu](mailto:JulianGamboa2023@u.northwestern.edu)

**Abstract:** Previously, we had proposed a hybrid opto-electronic correlator (HOC), which can achieve the same functionality as that of a holographic optical correlator but without using any holographic medium. Here, we demonstrate experimentally that the HOC is capable of detecting objects in a scale, rotation, and shift invariant manner. First, the polar Mellin transformed (PMT) versions of two images are produced, using a combination of optical and electronic signal processing. The PMT images are then used as the reference and the query inputs for the HOC. The observed correlation signal is used to infer, with high accuracy, the relative scale and angular orientation of the original images. We also discuss practical constraints in reaching a high-speed implementation of such a system. In addition, we describe how these challenges may be overcome for producing an automated version of such a correlator.

© 2019 Optical Society of America under the terms of the [OSA Open Access Publishing Agreement](#)

## 1. Introduction

Target recognition and tracking has a wide range of applications in the modern world. Optical image recognition systems offer a fast alternative over traditional electronics-based systems. The simplest such optical system is the Vander Lugt correlator [1–3], which is able to compare two images using holographic filters. However, a key limitation to this technology is the use of a slow recording process for the filters. Other correlators have been designed to circumvent the recording process, such as the Joint Transform Correlator (JTC) [4–9], which uses dynamic materials to record and correlate at the same time. However, the material needed for such a correlator suffers from many practical problems, such as the need for applying a high voltage, and get damaged easily [10,11]. We recently proposed and demonstrated a new hybrid opto-electronic correlator (HOC) [12,13] that overcomes some of these limitations and replaces the JTC's nonlinear material with detectors. The advantage of such a correlator is discussed in more detail in [12]. Yet two key limitations inherent to optical target recognition remain in our originally proposed HOC architecture: the system is intolerant to changes in scale and rotation. There have been many proposals to overcome these limitations, many of which detail the implementation of coordinate transforms [14–18]. We recently proposed that the incorporation of the polar Mellin Transform (PMT) into the existing HOC architecture would result in a shift, scale, and rotation invariant correlator [19]. In this paper, we show the results of such an incorporation using commercially available instruments. In addition, we show that the output of a positive match can be analyzed to determine the rotation angle of the query image.

Today, computers are able to detect matched images with great accuracy thanks to advances in neural networks and image recognition algorithms. However, even state of the art systems take upwards of 26 ms to detect matched features [20]. This time quickly adds up when scanning large databases or processing real-time camera feeds. Our system, as proposed using specialized circuits for the electronic components, is capable of reaching correlation times on the order of a few microseconds [12]. The HOC is not meant to replace computers,

as they are capable of detecting much finer details and performing more complex algorithms. Instead, it is expected to work as a pre-processor that would filter out obvious matches and mismatches, and produce a vastly reduced set of images that may require further processing. Of course, in principle, this pre-processing could also be performed using electronic circuits, entirely removing the need for optical components. However, the current best 2D Fourier Transform (FT) electronic integrated circuits have execution times of over 6ms per image [21], highlighting the need for optical techniques.

To exemplify the usefulness of the HOC, consider a database with 1 million images, 100 of which are potential matches to a query. A computer using state of the art algorithms would take  $0.026 \times 10^6 = 26,000 \text{ seconds} = 7.2 \text{ hours}$  to compare each database image to the query image by using neural networks. If instead one uses electronic FT's for correlation pre-processing (requiring at least two FT's per correlation), it would take  $0.006 \times 2 \times 10^6 = 12,000 \text{ seconds} = 3.3 \text{ hours}$  to filter out the 100 potential matches, which then require a subsequent  $100 \times 0.026 = 2.6 \text{ seconds}$  to process with neural networks for more detailed results. Assuming a correlation time of  $5 \mu\text{s}$ , the HOC requires  $5 \times 10^{-6} \times 10^6 = 5 \text{ seconds}$  to perform the filtering, and then  $2.6 \text{ seconds}$  for the neural network processing. It is this kind of large-database image processing that would benefit most from the HOC. While electronic components are generally cheaper and more robust, the difference in performance between an all-electronic and the hybrid opto-electronic approach is large enough to outweigh the disadvantages.

The rest of the paper is organized as follows. Section 2 details the experimental setup and theory of operation of the system. An overview of the steps required to implement the PMT in the HOC is given in section 3. The results are presented and examined in section 4, where we show how the use of the PMT conforms to the theory. We conclude with a summary and outlook in section 5.

## 2. Experimental setup and working principle of the HOC

The details of the basic HOC architecture can be found in [12] and [13], while the augmentation thereof via incorporation of the PMT can be found in [19]. If commercially available components are used, the operating speed of the HOC is severely limited by the serial communication between the devices. For this reason we proposed a system called the Integrated Graphic Processing Unit (IGPU) which may allow the HOC to perform a correlation in a time scale as short as few microseconds. Much work remains to be done before the IGPU can be realized. As such, we have shown the working principle of the HOC using existing technology, without optimizing the speed of operation.

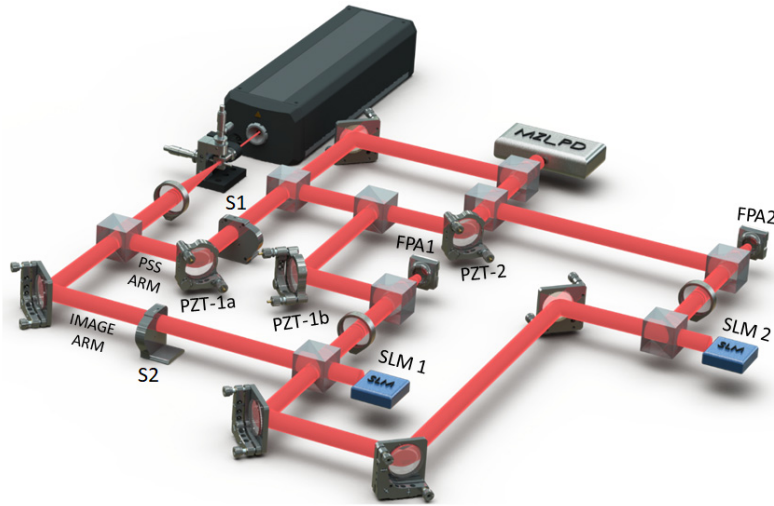


Fig. 1. Simplified architecture of the HOC for demonstrating the working principle using the PMT, shutters, and SLM's.

## 2.1 Overview of PMT augmented HOC

Like other optical correlators, the HOC takes advantage of the FT property of lenses. However, unlike traditional holographic correlators, it does not require a writing step where the information of the FT of the reference image is stored prior to its operation. Instead, the HOC captures the FT of the reference and query images, at the same time, on two separate arms. A Focal Plane Array (FPA) on each arm captures three intensity signals; the FT of the image, an auxiliary plane wave, and the interference between these two. The amplitude and phase information for the FT of the image is thus captured for each arm. We then subtract the intensity of the FT'd image beam and the auxiliary plane wave from the interference pattern for each arm. This yields two electronic FT-domain signals that are then multiplied together pixel-by-pixel resulting in a single output signal. By then transferring this signal back to the optical domain using an SLM, we can pass it through another lens and obtain its FT, which will correspond to the space-domain convolution and correlation of the two original images. This is further explained in section 2.3.

The amplitudes of the cross-correlation and convolution produced this way depend on the relative phase of the two auxiliary plane waves. Thus, for a practical implementation of this scheme we employ a Phase Stabilizer and Scanner (PSS), which is described in more detail later on.

The process as described above is able to recognize a match between a reference image and a query image in a shift invariant manner. However, it is not rotation and scale invariant. This limitation is eliminated by employing the PMT process. This involves the following additional steps in each arm before the interference with the auxiliary beams occurs. First, the FT of each image is detected with an FPA, then the amplitude of the FT is determined by taking the square root of the signal for each pixel. The resulting numbers are then converted from the rectilinear coordinates  $\{x, y\}$  to a polar coordinate  $\{r, \theta\}$ , by using the relation  $\theta = \tan^{-1}(y/x)$  and  $r = (x^2 + y^2)^{1/2}$ . The values of the signal are then represented in a two-dimensional rectilinear array, where  $r$  and  $\theta$  form the two orthogonal coordinates. In order to carry out this mapping, it is necessary to exclude the information in a small circle around the center of the amplitude of the FT. The radius of this circle,  $r_0$ , is chosen to be small enough to ensure that important features in the image are not lost. Finally, we map the signals

from the  $\{r, \theta\}$  array to a  $\{\rho, \theta\}$  array, where  $\rho = \ln(r/r_0)$ . This is the array that is interfered with the auxiliary beam in each arm. More details of this process can be found in [19].

## 2.2 Experimental setup

For this demonstration we have used a simplified version of the architecture proposed in [12]. This is illustrated schematically in Fig. 1. A continuous-wave diode-pumped solid-state laser (Verdi V2) at 532 nm is used as the light source. The laser beam starts with a diameter of 1 mm, which is spatially filtered and expanded to 1" (25.4 mm). This beam is passed through a 50/50 Beam Splitter (BS) into two arms; the Image Arm and the PSS Arm. The latter leads to a mirror mounted on a Piezo-electric Transducer (PZT-1a) which redirects the beam through a shutter (S1) to a Mach-Zehnder Interferometer (MZI). The MZI, along with PZT-2, a pair of photo-detectors (MZI PD) that are separated to detect two different fringes in the MZI interference pattern, and a Proportional-Integral-Differential (PID) controller, forms a phase-stabilization system. This MZI has two BS's inserted in one path. These redirect two plane waves ( $C_1, C_2$ ) towards the image arms, with  $C_1$  passing through PZT-1b. The phase-stabilization system allows us to lock the phase difference between  $C_1$  and  $C_2$  according to a bias voltage applied to the output of the PID controller. This is discussed in greater detail in section 2.4. The image arm also passes through a shutter (S2) and is then split into the reference and query arms. Each of these two beams reflects off an amplitude modulated (AM) SLM to produce the image beams ( $H_1, H_2$ ), each of which is then directed towards a biconvex lens. The lens produces the two dimensional FT of the image at its focal plane. Each of the FT'd image beams ( $M_1, M_2$ ) then interferes with the corresponding plane wave prior to being detected by an FPA placed at the focal distance of the biconvex lens. For this setup we used the Thorlabs USB2.0 CMOS camera (DCC1545M), which has a resolution of 1280x1024 pixels, to perform the function of the FPA.

The use of shutters allows us to choose what we detect. We can detect just the FT'd image beams ( $B = |M|^2$ ) by closing S1 and opening S2; just the plane waves ( $|C|^2$ ) by opening S1 and closing S2; or the interference patterns ( $A = |M + C|^2$ ) by opening both shutters.

The SLM's used for this demonstration are custom-made using Texas Instrument's DLP3000 modules. These work using Digital Micro-mirror Devices (DMD's) which rapidly move to reflect light towards and then away from a target, effectively functioning as AM SLM's. The DLP3000 modules have a physical resolution of 684 x 608 pixels, but operate in a wide aspect ratio of 854 x 480. The active area of the SLM is 0.3" (7.62 mm) and each individual micro-mirror measures 7.6  $\mu$ m across.

## 2.3 Mathematical Model of the HOC

In this version of the HOC, each set of measurements ( $A_j, B_j$ , and  $C_j$ ; where  $j=1,2$ ) is taken by opening and closing the shutters as described in the previous section, using the subscript '1' to denote the reference image, and the subscript '2' for the query.

The FT of each image and each plane wave can be expressed as follows:

$$M_j = |M_j| \exp(i\phi_j) \quad C_j = |C_j| \exp(i\Psi_j) \quad (1)$$

where  $\phi_j(x, y)$  is the phase of the FT'd image beam at the FPA plane, and  $\Psi_j$  is the phase of the interfering plane wave at the same point. Here,  $|M_j|$  and  $\phi_j$  are functions of  $(x, y)$ ,

but  $|C_j|$  and  $\Psi_j$  are assumed to be constant on the FPA plane. The detected interference pattern between the FT of the image ( $M_j$ ) and the plane wave ( $C_j$ ) is given by:

$$\begin{aligned} A_j &= |M_j + C_j|^2 \\ &= |M_j|^2 + |C_j|^2 + |M_j||C_j| \left( \exp(i[\phi_j - \Psi_j]) + \exp(-i[\phi_j - \Psi_j]) \right) \end{aligned} \quad (2)$$

This digital signal array can be stored on an FPGA along with the signal arrays  $B_j$  and  $|C_j|^2$ . The FPGA can then perform a subtraction to obtain:

$$\begin{aligned} S_j &= A_j - B_j - C_j \\ &= |M_j||C_j| \left( \exp(i[\phi_j - \Psi_j]) + \exp(-i[\phi_j - \Psi_j]) \right) \\ &= M_j C_j^* + M_j^* C_j \end{aligned} \quad (3)$$

This signal can be stored for both the reference ( $S_1$ ) and the query ( $S_2$ ) image in the same FPGA and later multiplied together using four-quadrant multiplication to find the signal array  $S$ :

$$\begin{aligned} S &= S_1 \times S_2 \\ &= (M_1 C_1^* + M_1^* C_1) \times (M_2 C_2^* + M_2^* C_2) \\ &= \alpha^* M_1 M_2 + \alpha M_1^* M_2^* + \beta^* M_1 M_2^* + \beta M_1^* M_2 \end{aligned} \quad (4)$$

where  $\alpha = C_1 C_2$ ,  $\beta = C_1 C_2^*$ .

The resulting signal can be sent to an SLM to be transferred into the optical domain using a laser. Here, the signal beam can be FT'd by passing through a biconvex lens, presenting the final output signal  $S_f$  at the focal plane:

$$\begin{aligned} S_f &= \mathcal{F}\{S\} \\ &= \alpha^* \mathcal{F}\{M_1 M_2\} + \alpha \mathcal{F}\{M_1^* M_2^*\} + \beta^* \mathcal{F}\{M_1 M_2^*\} + \beta \mathcal{F}\{M_1^* M_2\} \end{aligned} \quad (5)$$

Here  $\mathcal{F}$  stands for the FT. Because  $M_j$  is the FT of an image  $H_j$ , we can now use the well-known relationship between the FT of products of functions and convolutions and cross-correlations to express more explicitly the four terms in  $S_f$ :

$$\begin{aligned} S_f &= \alpha^* T_1 + \alpha T_2 + \beta^* T_3 + \beta T_4 \\ T_1 &= H_1(x, y) \otimes H_2(x, y) \\ T_2 &= H_1(-x, -y) \otimes H_2(-x, -y) \\ T_3 &= H_2(x, y) \odot H_1(x, y) \\ T_4 &= H_1(x, y) \odot H_2(x, y) \end{aligned} \quad (6)$$

where  $\otimes$  indicates two-dimensional convolution, and  $\odot$  indicates two-dimensional cross-correlation. This shows that using three intensity signals ( $A$ ,  $B$ , and  $C$ ) from each arm we can find the correlation between the two images. In Eqs. (4) to (6) we have grouped together the factors corresponding to the plane waves  $C_1$  and  $C_2$  into constants ( $\alpha$  and  $\beta$ ). A more explicit expression of these terms reveals the following:

$$\begin{aligned}\alpha &= C_1 C_2 & \beta &= C_1 C_2^* \\ \alpha &= |C_1| |C_2| \exp(i(\Psi_1 + \Psi_2))\end{aligned}\quad (7)$$

$$\beta = |C_1| |C_2| \exp(i(\Psi_1 - \Psi_2)) \quad (8)$$

It is clear that the output of the HOC depends nontrivially on the phases of the plane waves at their respective FPA's. We are also only interested in the cross-correlation terms of our output signal ( $T_3$  and  $T_4$ ); as such it is our goal to maximize  $|\beta|$  and minimize  $\alpha$  while maintaining both values stable. For this we have designed and implemented a PSS that is explained in the next section.

#### 2.4 Phase Stabilization and Scanning

The PSS can be considered to be a specific type of optical phase-locked loop (OPLL) with the added phase scan. Currently there are very few ways to implement a stable OPLL [22–24], and integrated circuits that perform this task are still at the research stage. To overcome this problem, we designed a discrete OPLL that can maintain lock for some time, along with a method of quickly reestablishing optimum lock values. The HOC requires us to control the phase difference between our Reference and Query auxiliary plane waves.

From Eq. (8) it is clear that  $|\beta|$  will reach its maximum value when  $\Psi_1 - \Psi_2 = 2\pi m$ , where ' $m$ ' is an integer. In order to achieve such a value, the HOC architecture incorporates an MZI with an adjustable mirror (PZT-2) and two coupled detectors (MZI PD), as shown in Fig. 2, which is a subset of the complete apparatus shown in Fig. 1. These detectors are separated a short distance on the plane normal to the direction of propagation of the laser, which allows them to detect different fringes of the interference pattern generated in the MZI. An electronic circuit finds the difference in intensity between these detectors and converts it into a voltage that is then fed into a low noise pre-amp and then a PID controller. The output of the PID is then added to a bias voltage that allows us to control the locking point before being connected to PZT-2. This system operates under the assumption that the mirrors and the optical path lengths are very stable. For this reason, the optical table is floated and the experiment is enclosed so as to minimize air turbulence.

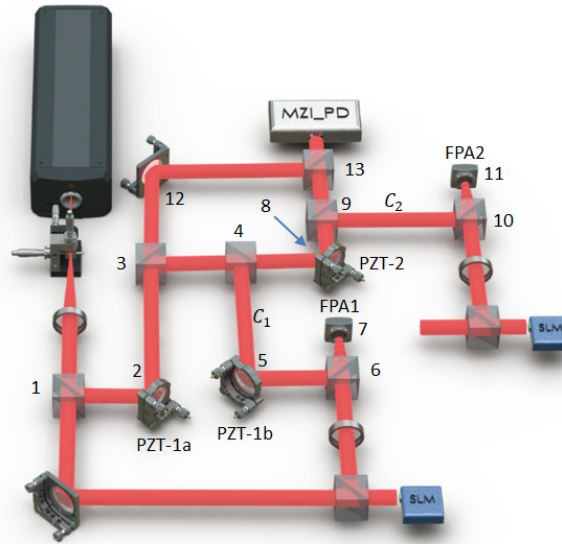


Fig. 2. Simplified HOC diagram. The numbers represent vertices used to describe path lengths.



The first plane wave ( $C_1$ ) is extracted from the MZI prior to the PZT, having travelled a distance  $L_{c1}$  from the first BS to FPA-1a, given by:

$$L_{c1} = l_{1,2} + l_{2,3} + l_{3,4} + l_{4,5} + l_{5,6} + l_{6,7} \quad (9)$$

where  $l_{m,n}$  indicates the path from element  $m$  to element  $n$ . The second plane wave ( $C_2$ ) is extracted after PZT-2. The total path length for this plane wave from the first BS to FPA-2 is  $L_{c2}$ , given by:

$$L_{c2} = l_{1,2} + l_{2,3} + l_{3,4} + l_{4,8} + l_{8,9} + l_{9,10} + l_{10,11} \quad (10)$$

PZT-2 allows us to change  $[l_{4,8} + l_{8,9}]$  via the bias voltage, thus extending or shortening  $L_{c2}$ . Without considering the effects of the optical components (BS's and mirrors) which produce constant phase shifts, the phase of each plane wave can be written as:

$$\Psi_j = k \times L_{cj}; \quad k = 2\pi / \lambda \quad (11)$$

Using this expression we can now find the phase difference to be:

$$\begin{aligned} \Delta\Psi &= \Psi_1 - \Psi_2 = k \times (L_{c1} - L_{c2}) + \Delta\phi_{OE} \\ &= k \times [l_{4,5} + l_{5,6} + l_{6,7} - (l_{4,8} + l_{8,9} + l_{9,10} + l_{10,11})] + \Delta\phi_{OE} \end{aligned} \quad (12)$$

where  $\Delta\phi_{OE}$  represents the constant difference in phase shift produced by the optical element in each path. We can also find the sum of the phases:

$$\begin{aligned} \Sigma\Psi &= \Psi_1 + \Psi_2 = k \times (L_{c1} + L_{c2}) + 2\phi_{init} \\ &= k \times [2 \times (l_{1,2} + l_{2,3} + l_{3,4}) + l_{4,5} + l_{5,6} + l_{6,7} + l_{4,8} + l_{8,9} + l_{9,10} + l_{10,11}] + 2\phi_{init} \end{aligned} \quad (13)$$

where  $\phi_{init}$  is the phase of the laser prior to hitting the first BS. By adjusting PZT-2 we are able to control  $L_{c2}$  directly, allowing us to adjust the value of  $\Delta\Psi$ . This phase difference is independent of  $[l_{1,2} + l_{2,3}]$ , which is controlled by PZT-1a. However, it is clear that  $\Sigma\Psi$  does depend on these lengths (i.e.  $[l_{1,2} + l_{2,3}]$ ). In this way we can scan this phase (i.e.  $\Sigma\Psi$ ), thus varying  $\alpha$ , while separately adjusting  $\Delta\Psi$  to maximize  $|\beta|$ . Nevertheless, it is clear that varying  $\Delta\Psi$  will also change  $\Sigma\Psi$ , as they both depend on  $[l_{4,8} + l_{8,9}]$ . For this reason, to vary  $\beta$  without affecting  $\alpha$  it is necessary to have another PZT, denoted PZT-1b (shown in Fig. 2), as explained below.

We define  $\Delta l_{pzt}$  as the matching change in  $l_{4,8}$  and  $l_{8,9}$  produced by the displacement of PZT-2 away from its static point. Similarly, we can also define  $\Delta l_{ref}$  as the matching change in  $l_{4,5}$  and  $l_{5,6}$  due to the displacement of PZT-1b. This gives us:

$$\begin{aligned} l_{4,8} + l_{8,9} &= l'_{4,8} + l'_{8,9} + \Delta l_{pzt} \\ l_{4,5} + l_{5,6} &= l'_{4,5} + l'_{5,6} + \Delta l_{ref} \end{aligned} \quad (14)$$

where  $l'_{n,m}$  represents the path length when the relevant PZT is at its static point. We can now write:

$$\begin{aligned}\Delta\Psi &= k \times [l'_{4,5} + l'_{5,6} + l_{6,7} - (l'_{4,8} + l'_{8,9} + l_{9,10} + l_{10,11}) + (\Delta l_{ref} - \Delta l_{pzt})] + \Delta\phi_{OE} \\ \Sigma\Psi &= k \times [2 * (l_{1,2} + l_{2,3} + l_{3,4}) + l'_{4,5} + l'_{5,6} + l_{6,7} + l'_{4,8} + l'_{8,9} + l_{9,10} + l_{10,11} + (\Delta l_{ref} + \Delta l_{pzt})] + 2\phi_{init}\end{aligned}\quad (15)$$

From Eq. (15) it is clear that by setting

$$\Delta l_{ref} = -\Delta l_{pzt} \quad (16)$$

we can get rid of the PZT-2 dependence of  $\Sigma\Psi$  while doubling it in  $\Delta\Psi$  :

$$\begin{aligned}\Delta\Psi &= k \times [l'_{4,5} + l'_{5,6} + l_{6,7} - (l'_{4,8} + l'_{8,9} + l_{9,10} + l_{10,11}) - (2\Delta l_{pzt})] + \Delta\phi_{OE} \\ \Sigma\Psi &= k \times [2 * (l_{1,2} + l_{2,3} + l_{3,4}) + l'_{4,5} + l'_{5,6} + l_{6,7} + l'_{4,8} + l'_{8,9} + l_{9,10} + l_{10,11}] + 2\phi_{init}\end{aligned}\quad (17)$$

Mechanically this means that PZT-1b has to be programmed to move the exact same distance as PZT-2, but in the opposite direction. This can be achieved using a feed-forward system where an inverted version of the bias signal applied to PZT-2 is sent to PZT-1b.

The PID system that controls PZT-2 receives its feedback from MZI\_PD. The phase difference between the two path lengths in the MZI can be written as:

$$\Delta\Psi_{MZI} = \Psi_{control} - \Psi_{static} \quad (18)$$

where

$$\begin{aligned}\Psi_{control} &= k \times (l_{3,4} + l'_{4,8} + l'_{8,9} + l_{9,13} + \Delta l_{pzt}) \\ \Psi_{static} &= k \times (l_{3,12} + l_{12,13})\end{aligned}\quad (19)$$

so that

$$\Delta\Psi_{MZI} = k \times [l_{3,4} + l'_{4,8} + l'_{8,9} + l_{9,13} + \Delta l_{pzt} - (l_{3,12} + l_{12,13})] \quad (20)$$

This means that to lock the PID to a specific phase at MZI\_PD ( $\Delta\Psi_{MZI}$ ) we will have a set value of  $\Delta l_{pzt}$ , which will also lock  $\Delta\Psi$ . We can vary this value by use of a bias voltage that is added to the output of the PID controller [25].

As was previously shown, PZT-1a allows us to adjust the value of  $\Psi_1$  and  $\Psi_2$  simultaneously without changing  $\Delta\Psi$ . By continuously running a ramp signal at some frequency  $\omega_s$  on this PZT, we can scan over a wide range of phases. By applying a Low Pass Filter (LPF) to the detected signal with a cutoff frequency  $\omega_c \ll \omega_s$  we can get rid of the  $\alpha$  term in Eq. (6), leaving only the cross-correlation signals in our final HOC output:

$$\begin{aligned}S_f &= \beta^* T_3 + \beta T_4 \\ &= \beta^* [H_2(x, y) \odot H_1(x, y)] + \beta [H_1(x, y) \odot H_2(x, y)]\end{aligned}\quad (21)$$

This is the ideal way to operate the HOC. However, because the phase scan operates in the time domain, this method requires that all six signals ( $A_j, B_j$ , and  $C_j$ ) be detected simultaneously with six FPA's, and without shutters, which greatly increases the complexity of the system. As such, we did not implement the scanning segment of the PSS for the demonstration reported here. It should be noted that it is still possible to see the results of a correlation without washing out the  $\alpha$  term, but one must be careful to distinguish between the correlation and convolution terms.



One way to reach the maximum value of  $|\beta|$  for an unknown  $\alpha$  is to run a series of known matched images through the HOC at varying bias voltages. This works as follows. One image is set as both the Reference and Query inputs. The HOC then runs a correlation, for a particular bias voltage. This will yield a match at the output of the HOC. The bias voltage is then changed within the range of operation of the PZT, repeating the correlation. The result will again be a match, but the overall output intensity will have either increased or decreased. The bias voltage is changed so as to look for the maximum intensity. This process is repeated, changing the bias in progressively smaller steps until the maximum output intensity is found.

### 3. Polar Mellin transform in the HOC

Due to the properties of the FT and lenses, the detection of a FT'd optical signal will be shift invariant. However, changes to the scale and rotation of the images will alter the scale and rotation of the FT, thus preventing the HOC from achieving a match. To counteract this we can instead compare images that have been pre-processed via the use of the Polar Mellin Transform (PMT).

Because the PMT is, by definition, in log-polar coordinates; two identical images with different rotations will present the same PMT with a shift in the  $\theta$  coordinate corresponding to the relative rotation angle between them. Similarly, any change in scale will manifest as a shift in the log-radial coordinate  $\rho$ . By performing the PMT we are essentially converting any rotation and scale changes into translational shifts. Given that the established HOC architecture is inherently shift invariant and that the PMT is very closely related to the FT, it is thus well suited for adding rotation and scale invariance into the HOC architecture, as explained in detail in [19].

The steps to obtain the PMT in an optoelectronic system are as follows: 1- Find the FT of the image. 2- Determine the amplitude of the FT. (2a- Determine the intensity of the FT. 2b- Find the square root of the intensity). 3- Perform circular DC blocking. 4- Map polar coordinates into a rectilinear plane where  $x$  and  $y$  correspond to the  $r$  and  $\theta$  axes. 5- Transform radial coordinate to the logarithm of the ratio of the radial coordinate and a reference length.

Steps 1 and 2a can be performed using a laser, an SLM, a FT lens, and an FPA. In this setup we used a single arm of our existing HOC architecture with the PSS shutter (S1) closed. Steps 2b-5 are then performed by a computer. The resulting PMT image is then used as an input to the HOC.

By using a PMT image as a reference and converting a query image into its PMT, the HOC is able to find the correlation of the two original images in a shift, scale, and rotation invariant manner.

Given that all real digital images are composed of positive integer values, their FT will always contain a high value at the center (DC). The transformation from  $\{x, y\}$  to  $\{\rho, \theta\}$  of such an image will produce an output that has a non-zero value for  $\rho = 0$ . It is impossible to transform this point to the log-polar domain. To avoid this, we cut a small hole in the intensity profile of the FT at DC prior to performing the polar coordinate transformation. This is called circular DC blocking [19]. It is important that the hole be small enough not to erase important information from the non-DC area of the FT. However, making the hole very small requires high pixel density. A convenient compromise is to use a small hole of a constant size for all images.

If a constant-size circular DC block is chosen, the PMT conversion process can be achieved without any complex computations. The final three steps of the PMT process are independent of the detected image and can be achieved by physically connecting an  $\{x, y\}$  coordinate input to a rectilinear-mapped  $\{\rho, \theta\}$  coordinate output (neglecting the

connections corresponding to the circular DC block hole). In this way a single Application Specific Integrated Circuit (ASIC) could perform the PMT with the help of a FT lens. If an FPA and an SLM are built into this ASIC, the HOC would be able to achieve shift, scale, and rotation invariance using regular non-PMT images by inserting the ASIC at each image arm as shown in Fig. 3.

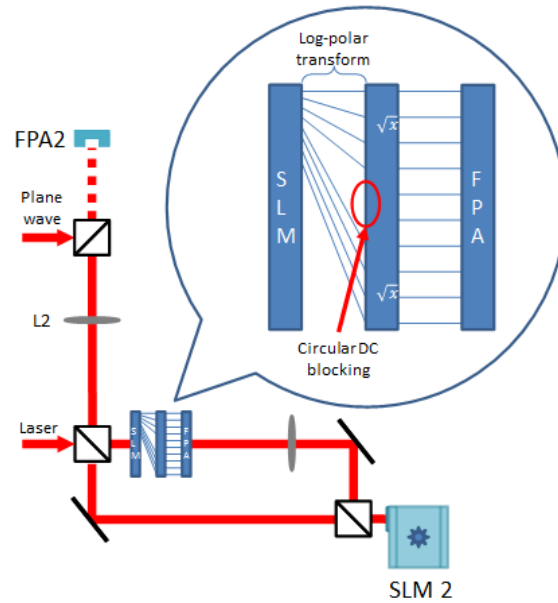


Fig. 3. PMT ASIC design and incorporation into the query arm of the simplified HOC architecture.

Ideally we would expect the external SLM to be connected to either a camera or a computer to provide the non-PMT images. It would also be beneficial to incorporate such a system only at the query arm as shown in Fig. 3, with the reference arm using a holographic memory disk instead of an SLM to store a large database of PMT reference images.

#### 4. Experimental results

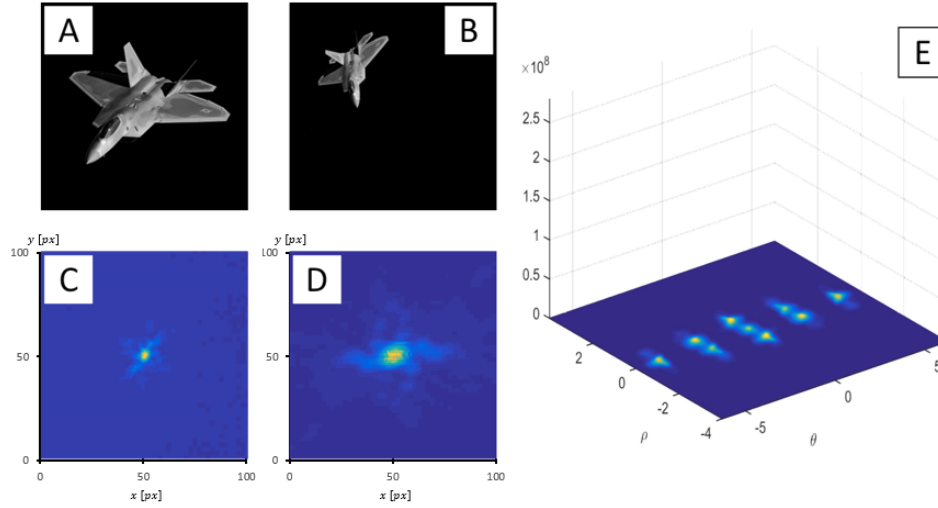


Fig. 4. HOC results without the PMT. (A): Reference Input. (B): Query Input. (C): Measured FT of A. (D): Measured FT of B. (E): Output scaled to the intensity of a known match.

For this experiment, a grayscale image of an F-22 Raptor fighter jet was chosen for its excellent contrast, unique shape, and real-world value. Prior to running the experiment, the HOC was calibrated to its optimum bias voltage by using the method described in section 2.4 of this document.

The original reference image is shown in Fig. 4(a). The query image shown in Fig. 4(b) has been shifted and is scaled by a factor of 0.5 with a rotation of  $48.25^\circ$  counterclockwise with respect to the reference. The detected FT's of these two images are shown in Figs. 4(c) and 4(d) respectively. Because the query image is scaled, its FT is larger than the reference while also presenting a rotation. Because of these two factors, the HOC was unable to detect a match, producing an almost flat output signal  $|S_f|^2$  in Fig. 4(e).

Figures 4.C and 4.D were then used as FT intensities in the PMT conversion process described in section 3. The PMT'd images are shown in Figs. 5(a) and 5(b), where the vertical axis represents  $\theta$  and the horizontal axis represents  $\rho$ . Using these PMT images as new inputs to the HOC, their FT's (Figs. 5(c) and 5(d)) were detected. In these new FT's, the scale and rotation of the query image with respect to the reference is no longer visible. This is corroborated by the output  $|S_f|^2$  shown in Fig. 5(e) which shows a clear peak that is  $\sim 15$  times larger than that of Fig. 4(e), indicating a successful correlation. On Fig. 5(b) we have added a red horizontal line that marks the value of  $\theta$  that corresponds to  $\theta = 0$  in Fig. 5(a). This line shows the translational shift of the PMT caused by the rotation of the original query image. The section of the PMT that corresponds to the top of Fig. 5(a) has looped around to be under this red line.

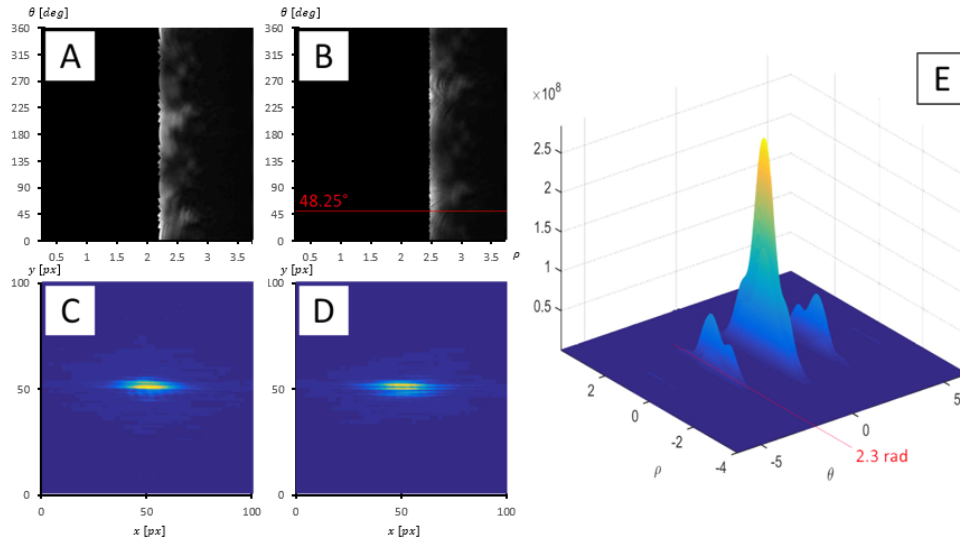


Fig. 5. HOC results with the PMT. (A): Reference PMT Input (converted from 4.C). (B): Query PMT Input (converted from 4.D). (C): Detected FT of A. (D): Detected FT of B. (E): Output scaled to the intensity of a known match.

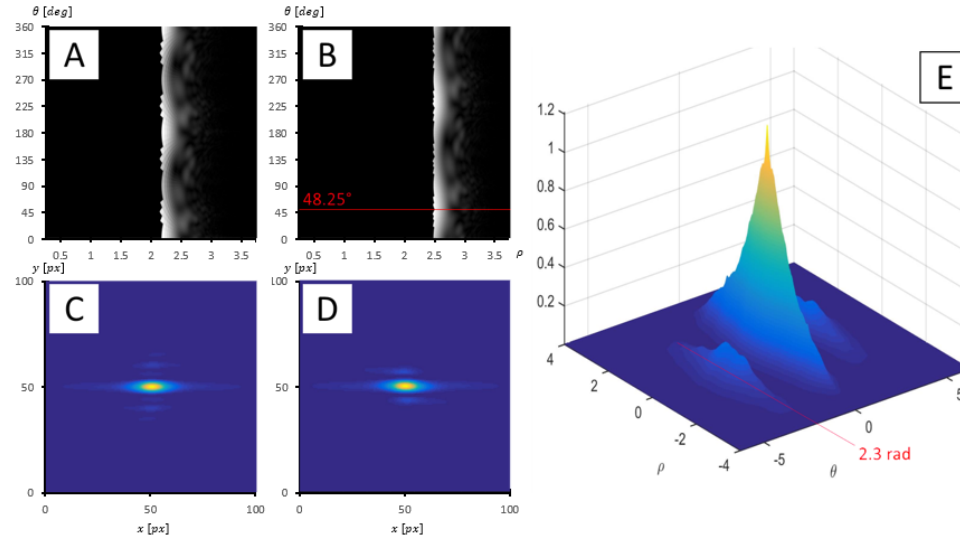


Fig. 6. HOC simulation with the PMT. All images are simulated. (A): Reference PMT Input (from 4.A). (B): Query PMT Input (from 4.D). (C): 2-D Fast Fourier Transform (FFT) of A. (D): FFT of B. (E): Output normalized to 1.

To complement these results, a simulation using the same input images was run. This is shown in Fig. 6, corresponding to the ideal reference PMT, ideal query PMT, their ideal FT's, and the simulated HOC output  $|S_f|^2$ . In Fig. 6(b) we have added a similar red line to the one in Fig. 5(B), this time corresponding to  $\theta = 0$  in Fig. 6(a).

By measuring the distance in pixels between the bottom of the PMT and the red line, recalling that the full vertical axis represents  $360^\circ$ , we can estimate the rotation of the query image to be  $\approx 48^\circ$ , which is close to the real rotation of  $48.25^\circ$ .

Similarly, the distance between the central peak of the output signal and the two lateral peaks in Figs. 5(e) and 6(e) has been marked with a red line. This is located at  $\theta = 2.3 \text{ rad}$ , which is equivalent to a rotation of  $48.22^\circ$ .

## 5. Conclusions and outlook

We have demonstrated that an HOC built using commercially available components and incorporating the PMT is able to find a match in a shift, scale, and rotation invariant manner, yielding an output that is  $\sim 15$  times larger when a match is found vs when it is not found (without the PMT). Furthermore, the relative rotation of the query image with respect to the reference image in a match can be found in the output signal by measuring the distance from the central peak to one of the two lateral peaks. We have also shown that the behavior of the PMT-augmented HOC aligns with the theory by presenting simulated results that correspond to our experiment.

The development of the PMT-HOC can be categorized in three stages. In stage 1, we have demonstrated the functionality of the system by manually using a computer to perform the electronic processing. In stage 2, the PMT's of images and the mathematical processes required can be performed by an FPGA, thus fully automating the system. In stage 3, all of the signal processing can be done by using specially designed integrated circuits that can be incorporated into the FPA's and SLM's, forming an IGPU. This stage would allow for high-speed automation of the system, performing correlations in a time scale as short as a few microseconds.

## Funding

Air Force Office of Scientific Research (AFOSR) (FA9550-18-01-0359).

## References

1. A. Vander Lugt, "Signal detection by complex spatial filtering," *IEEE Trans. Inf. Theory* **10**(2), 139–145 (1964).
2. A. Heifetz, J. T. Shen, J.-K. Lee, and M. S. Shahriar, "Translation-invariant object recognition system using an optical correlator and a super-parallel holographic random access memory," *Opt. Eng.* **45**, 025201 (2006).
3. A. Heifetz, G. S. Pati, J. T. Shen, J. K. Lee, M. S. Shahriar, C. Phan, and M. Yamamoto, "Shift-invariant real-time edge-enhanced VanderLugt correlator using video-rate compatible photorefractive polymer," *Appl. Opt.* **45**(24), 6148–6153 (2006).
4. Q. Tang and B. Javidi, "Multiple-object detection with a chirp-encoded joint transform correlator," *Appl. Opt.* **32**(26), 5079–5088 (1993).
5. J. Khoury, M. Cronin-golomb, P. Gianino, and C. Woods, "Photorefractive two-beam-coupling nonlinear joint-transform correlator," *J. Opt. Soc. Am. B* **11**(11), 2167–2174 (1994).
6. B. Javidi, J. Li, and Q. Tang, "Optical implementation of neural networks for face recognition by the use of nonlinear joint transform correlators," *Appl. Opt.* **34**(20), 3950–3962 (1995).
7. F. T. S. Yu and X. J. Lu, "A real-time programmable joint transform correlator," *Opt. Commun.* **52**(1), 10–16 (1984).
8. M. S. Shahriar, R. Tripathi, M. Kleinschmit, J. Donoghue, W. Weathers, M. Huq, and J. T. Shen, "Superparallel holographic correlator for ultrafast database searches," *Opt. Lett.* **28**(7), 525–527 (2003).
9. F. Lei, M. Iton, and T. Yatagai, "Adaptive binary joint transform correlator for image recognition," *Appl. Opt.* **41**(35), 7416–7421 (2002).
10. D. A. Gregory, J. A. Loudin, and H.-K. Liu, "Joint transform correlator limitations," *Proc. SPIE* **1053**, 198–207 (1989).
11. B. Javidi and C.-J. Kuo, "Joint transform image correlation using a binary spatial light modulator at the Fourier plane," *Appl. Opt.* **27**(4), 663–665 (1988).
12. M. S. Monjur, S. Tseng, R. Tripathi, J. J. Donoghue, and M. S. Shahriar, "Hybrid optoelectronic correlator architecture for shift-invariant target recognition," *J. Opt. Soc. Am. A* **31**(1), 41–47 (2014).
13. M. S. Monjur, S. Tseng, M. F. Fouda, and S. M. Shahriar, "Experimental demonstration of the hybrid optoelectronic correlator for target recognition," *Appl. Opt.* **56**(10), 2754–2759 (2017).
14. D. Casasent and D. Psaltis, "Position, rotation, and scale invariant optical correlation," *Appl. Opt.* **15**(7), 1795–1799 (1976).
15. D. Casasent and D. Psaltis, "Scale invariant optical correlation using Mellin transforms," *Opt. Commun.* **17**(1), 59–63 (1976).
16. D. Casasent and D. Psaltis, "New optical transforms for pattern recognition," *Proc. IEEE* **65**(1), 77–84 (1977).
17. D. Asselin and H. H. Arsenault, "Rotation and scale invariance with polar and log-polar coordinate

- transformations,” *Opt. Commun.* **104**(4-6), 391–404 (1994).
18. D. Sazbon, Z. Zalevsky, E. Rivlin, and D. Mendlovic, “Using Fourier/Mellin-based correlators and their fractional versions in navigational tasks,” *Pattern Recognit.* **35**(12), 2993–2999 (2002).
  19. M. S. Monjur, S. Tseng, R. Tripathi, and M. S. Shahriar, “Incorporation of polar Mellin transform in a hybrid optoelectronic correlator for scale and rotation invariant target recognition,” *J. Opt. Soc. Am. A* **31**(6), 1259–1272 (2014).
  20. W. Shi, J. Caballero, F. Huszár, J. Totz, A. Aitken, R. Bishop, D. Rueckert, and Z. Wang, “Real-Time Single Image and Video Super-Resolution Using an Efficient Sub-Pixel Convolutional Neural Network,” in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (2016), 1874–1883.
  21. M. Noskov, V. Tutatchikov, M. Lapchik, M. Ragulina, and T. Yamskikh, “Application of parallel version two-dimensional fast Fourier transform algorithm, analog of the Cooley-Tukey algorithm, for digital image processing of satellite data,” in *E3S Web of Conferences (EDP Sciences, 2019)*, paper 01012.
  22. G. W. Li, S. J. Huang, H. S. Wu, S. Fang, D. S. Hong, T. Mohamed, and D. J. Han, “A Michelson interferometer for relative phase locking of optical beams,” *J. Phys. Soc. Japan* **77**, 024301 (2008).
  23. B. W. Shiau, T. P. Ku, and D. J. Han, “Real-time phase difference control of optical beams using a mach-zehnder interferometer,” *J. Phys. Soc. Japan* **79**, 034302 (2010).
  24. M. Lu, H. C. Park, E. Bloch, L. A. Johansson, M. J. Rodwell, and L. A. Coldren, “An integrated heterodyne optical phase-locked loop with record offset locking frequency,” in *Optical Fiber Communication Conference OSA Technical Digest Series (Optical Society of America, 2014)*, paper Tu2H.4.
  25. Because the PZT is an electro-mechanical device, it requires a voltage source and a control circuit that introduce electrical noise into the system. A PID controller allows the PZT to maintain a more stable position. However, PID systems require a feedback loop. An MZI was constructed to provide the feedback for the PID via interferometry. This constitutes a mechanically controlled OPLL. This would not be required in an integrated system where the functionality of the PZT may be replaced by other means of phase control.