ISSN 2278-3083
Volume 6, No.1, January – February 2017
International Journal of Science and Applied Information Technology
Available Online at http://www.warse.org/ijsait/static/pdf/file/ijsait01612017.pdf

Object Distance Determination Using a Joint Transform Correlator

Ronald Marsh[1], Tyler Schmitt[2]
[1]University of North Dakota, USA, rmarsh@cs.und.edu
[2]Le Moyne College, USA, schmittj@stu.lemoyne.edu

ABSTRACT

Using two cameras separated by some distance, along with the concepts of parallax, we can determine the distance to an object. The joint transform correlator is a natural tool to consider for this task. Since the Fourier transform, on which optical correlators are based, is lossless and retains shift information, we suppose that distance information encoded in a stereo image pair is preserved through the correlation process. We then recover that distance by investigating the location of the correlation peak(s). Data obtained indicates that it is possible to extract distance information from a joint transform correlation result.

Key words : Fourier optics, signal processing, machine vision, correlation.

1. INTRODUCTION

Due to their low cost and ease of launch vehicle integration, U-class spacecraft (CubeSats), comprising multiples of 10×10×11.35 cm cubic units, have generated a lot of interest [1]. At the University of North Dakota an interdisciplinary, student-run CubeSat program, the Open Prototype for Educational Nanosats (OPEN) [2], has been established with a primary goal of creating a CubeSat design that is very affordable (< $5000.00). The first CubeSat (OpenOrbiter [3]) is scheduled for launch in 2017 and will be an Earth observing satellite. However, we have proposed other versions that will require multiple CubeSats and will require these spacecraft to remain at specific distances from each other. Given the limited power budget we have, due to our low-cost design, we are very interested in a passive approach to maintaining this separation. Hence, our interest in optical correlators**.**

1.1 Optical Correlators

An optical correlator is a device for comparing two images using Fourier transforms and is generally used for pattern matching purposes. More specifically, an optical correlator takes two input images and outputs another image showing the cross-correlation. Or, in other words, in what relative position do the two images best match. If the two input images are sufficiently similar and similarly oriented (scale, aspect, and rotate), one will see a bright spot in the correlation result, commonly referred to as a peak. The location and brightness of the peak indicates the location of one image within the other and the level of similarity.

The first successful optical correlator, the matched filter correlator (MFC), was developed by VanderLugt in 1963 [4]. The MFC was a pioneering development in the optical correlator field and is still in use today. The significance of VanderLugt's design was that the required Fourier transforms would be computed optically rather than digitally, which, in 1963, provided a serious performance advantage. However, a drawback to VanderLugt's design was that performance was highly sensitive to the alignment of the optical train.

The MFC was designed to locate a particular image, known as the "reference", inside one or many other images, referred to as the "target". With the advent of modern spatial light modulators, it is typically the case that the "reference" image be digitally processed into a "filter". This processing usually involves performing a Fourier transform on the "reference" image followed by another step that may or may not include the manipulation of the real and imaginary components of the Fourier transform. Figure 1 depicts the MFC process.

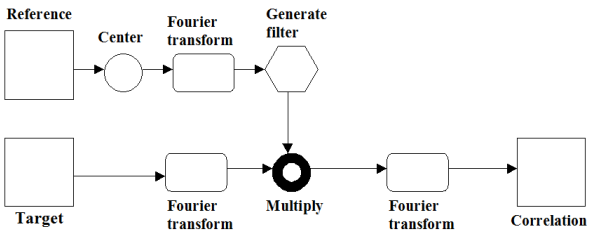

Figure 1: Matched filter correlator schematic.

As such, the reference and target image are treated differently within the correlator, and the MFC is best suited to such asymmetrical applications and not the application of interest here.

The joint transform correlator (JTC) was invented in 1966 [5] by Weaver and Goodman and was intended to overcome the limitations of the matched filter correlator. The JTC is much

less sensitive to instrument alignment but is also less space-efficient. In addition, both input images in a JTC undergo the same transformations; there is no concept of a "target" or "reference" image and there is no digital preprocessing of either image. Figure 2 depicts the JTC process. Thus, the JTC is better suited to applications that require no preferential treatment of either input image, such as the application of interest here.

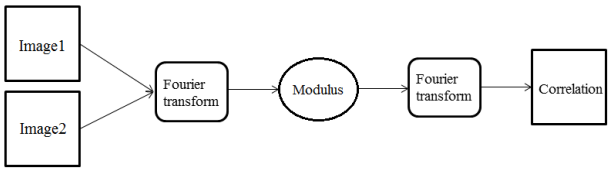

Figure 2: Joint transform correlator schematic.

As expected, correlation (MFC or JTC) need not be done optically anymore; the same process can be performed digitally [6] affording easier post-analysis of the correlation result and affording easier implementation on a CubeSat.

1.1 Application to Distance Detection

As noted in the abstract, many animals use two eyes to determine distance to an object and we can we can do something very similar using two cameras separated by some distance along with the concepts of parallax [7,8,9]. The idea behind parallax is to compare two images of the same scene taken from different vantage points leading to an indicator of distance. Optical correlators, in particular, the joint transform correlator, is a natural tool to consider for this task. If implemented using traditional computational means, a computer would detect the distance to the object by measuring the shift in the object's 2D location from one vantage point to another [10].

Since the Fourier transform, on which optical correlators are based, is lossless, retains shift information, and has a fixed computational complexity regardless of image content, we postulate that any distance information encoded in the original images will be preserved through the correlation process. In effect, the correlator automates the process of finding a common pixel in the algorithm outlined in [11]. Since we use a new pair of images, acquired from a pair of cameras, for each distance measurement, the joint transform correlator is the logical choice. To the best of our knowledge, the use of a JTC to determine distance to an object is a novel approach.

2. EXPERIMENT

Images were acquire by two Microsoft LifeCam Cinema webcams aligned horizontally. One experiment used a 19.05 cm separation between cameras, while the other used a 24.13 cm separation between cameras (figure 3). LifeCam Cinema webcams have a 73° field of view and an autofocusing lens. To best replicate the conditions expected in orbit, where there would be little background information, the target was a 2.5625"x4.375" cell phone on a wall in a dark room.

We measured the distance to the wall with a tape measure, then took a picture with each camera. Since the LifeCam Cinema webcams acquire 640x360 color images, we used the convert utility from the ImageMagick [12] package to expand the image size, of all images acquired, to 1024x1024 facilitating the use of the fast Fourier transform. The ImageMagick script used is:

```
Convert input_file.jpg –gravity center –background black -extent
1024x1024 output_file.pgm
```

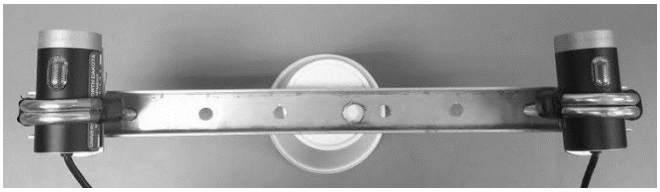

Figure 3: Microsoft LifeCam webcam used for image collection.

The ImageMagick script takes the original sized (640x360) color jpeg image, centers the original image in the expanded space (1024x1024), fills the extra space with black pixels, and saves the result to a1024x1024 sized gray scale pgm image. Figure 4 depicts the result (note that this image was NOT used in the experiment and is included purely for demonstrative purposes).

Figure 4: Example of ImageMagick preprocessing.

2.1 JTC Implementation

There are a variety of approaches that can be used to generate a joint transform correlation, the method shown in figure 2 (probably the simplest and most common), a method proposed by Jin and Lee that uses a fractional Fourier transform [13], a method proposed by Wang, He, and Sheng that uses power spectrum subtracting and exponential filtering [14]. Nomura, in 1998, proposed that Phase-encoded JTC [15]. Finally, Cherri and Alam proposed the reference phase-encoded fringe-adjusted joint transform correlation [16].

We chose the method shown in figure 2 and implemented it with some changes, as follows. As noted above the images are

padded to form 1024x1024 images, using our variation of the JTC shown in figure 2 requires that we perform the JTC on a 4096x1024 array/image. Initially, the 4096x1024 array contains the two images acquired from the left and right cameras (left image and right image). However, originally we intentionally positioned the images such that the left image is located on the left side of the array and the right image is shifted by N pixels to the right, where N is determined by equation 1.

$$N = 2 * width + \left(\frac{width}{2}\right) \quad (1)$$

Where width is the width of the input image (1024 in our case).

The intent is to add a certain amount of phase into the JTC such that the resultant correlation peak will be located in the center of the leftmost image if no difference exists in the input images. Using the same example image as used in figure 4, we obtain the 4096x1024 image shown in figure 6. Please note that the background was changed from black to a dark gray to better depict the positioning of the two images in the 4096x1024 image. In the actual tests the background was left black.

Once the input array is populated a fast Fourier transform (FFT) is performed along the horizontal dimension (4096 pixels), followed by a fast Fourier transform (FFT) performed along the vertical dimension (1024 pixels). We then calculate the modulus of the result and populate the input array with that data. The process is then repeated creating the correlation result.

Figure 5: Example input to JTC.

Figure 6 shows the JTC result using the input shown in figure 5. Three peaks are visible. The peak enclosed by the box is a peak that is always present with this implementation of the JTC, sometimes referred to as the DC term or 0 Hz term. Since there is no positional difference between the left and right images shown in figure 5, the actual correlation peak(s) (circled) are located in the center of the image (leftmost image in figure 5), as expected. We have two peaks (circled) due to the mirroring nature of the Fourier transform.

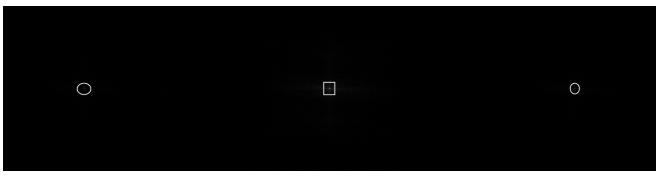

Figure 6: Example correlation output.

However, as we found in an earlier study [17], when we positioned the images using the offset described by equation 1, we would get a "flipping" of the axis as the object gets closer. During this study we found that by changing the offset we eliminate that "flipping" effect. The offset used is now show in equation 2.

$$N = 2 * width \quad (2)$$

Where width is the width of the input image (1024 in our case).

Figures 7 and 8 show the input and correlation output using the offset of equation 2. In figure 8, instead of using circles, we used arrows to point out the location of the peaks.

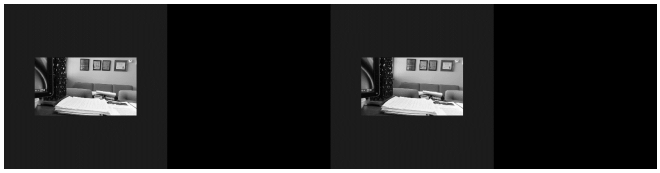

Figure 7: Example input to JTC.

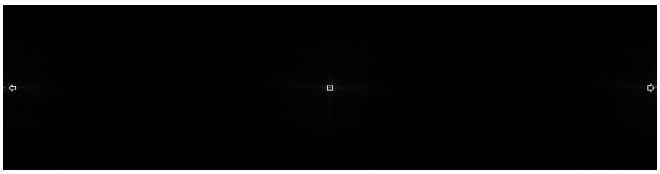

Figure 8: Example correlation output.

3. RESULTS

For the 24.13 cm (9.5 inch) camera separation, images were acquired at different distances ranging from 0.6096 m to 6.096 m. For the 19.05 cm (7.5 inch) camera separation, images were acquired at different distances ranging from 0.24384 m to 6.096 m. We plot the location of the peak versus the distance to the wall to find a clear relation. Coordinates of the peak are measured from center, since distance to the wall should not depend on the size of the image. Note that we take the absolute value of the coordinates since swapping the input images negates the peak's coordinates.

Since we collected data at a limited number of distances, we then applied a curve fit to produce a method applicable to any distance. For the 19.05 cm camera separation we found that the data was best represented by equation 3.

$$Y = \frac{1}{(0.009238918 + 0.002446541\ X^{0.95})} \quad (3)$$

Where X is the correlation peak location and Y is the resulting distance value.

Table 1 provides the data used. The first column is the measured object distance (distance to cameras in feet), the second is the correlation peak location, the third column is the

distance derived from equation 3, the fourth is the difference between the actual and curve fit distance, and the fifth and sixth columns are the error, in inches and in cm, of the distance calculation using equation 3. Figure 9 provides the same data only plotted. The measured points are represented as dots with the curve fit data represented as the line.

Table 1: Data for 19.05 cm camera separation.

Object Distance	Peak Location	Object Distance	Location Difference	Error (cm)
0.708	834	0.682	0.026	0.80
0.792	761	0.743	0.048	1.48
0.875	663	0.846	0.029	0.87
0.958	577	0.965	-0.007	0.21
1	500	1.104	-0.104	3.17
2	255	2.074	-0.074	2.26
3	167	3.072	-0.072	2.18
4	122	4.099	-0.099	3.01
5	97	5.050	-0.050	1.51
10	47	9.607	0.393	11.98
15	28	14.875	0.125	3.82
20	19	20.259	-0.259	7.91

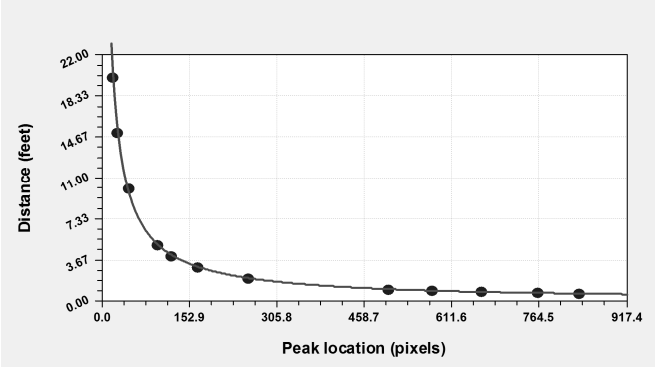

Figure 9: Distance vs peak location for 19.05 cm data using equation 3.

As with the 19.05 cm camera separation, we only collected data at a limited number of distances for the 24.13 cm camera separation. Therefore, we again applied a curve fit to produce a method applicable to any distance. For the 24.13 cm camera separation we found that the data was best represented by equation 4.

$$Y = \frac{1}{(-0.0005970295 + 0.0022771661\,X^{0.8954256116})} \quad (4)$$

Where X is the correlation peak location and Y is the resulting distance value.

Table 2 provides the data used. The first column is the measured object distance (distance to cameras in feet), the second is the correlation peak location, the third column is the distance derived from equation 4, the fourth is the difference between the actual and curve fit distance, and the fifth and sixth columns are the error, in inches and in cm, of the distance calculation using equation 4. Figure 10 provides the same data only plotted. The measured points are represented as dots with the curve fit data represented as the line.

Table 2: Data for 24.13 cm camera separation.

Object Distance	Peak Location	Object Distance	Location Difference	Error (cm)
0.958	890	1.004	-0.046	1.41
1.042	852	1.044	-0.002	0.07
1.67	741	1.184	0.486	14.83
2	374	2.184	-0.184	5.62
3	248	3.158	-0.158	4.80
4	186	4.088	-0.088	2.67
5	148	5.018	-0.018	0.56
10	70	9.840	0.160	4.89
15	44	14.958	0.042	1.29
20	31	20.534	-0.534	16.28

An obvious question, since both camera separations can be modeled by a Harris curve, is what would be the error if one were to simply average the two sets of parameters into a single Harris curve? Creating a curve represented by equation 5.

$$Y = \frac{1}{(0.0016234312 + 0.0012301854\,X^{0.9227181056})} \quad (5)$$

Where X is the correlation peak location and Y is the resulting distance value.

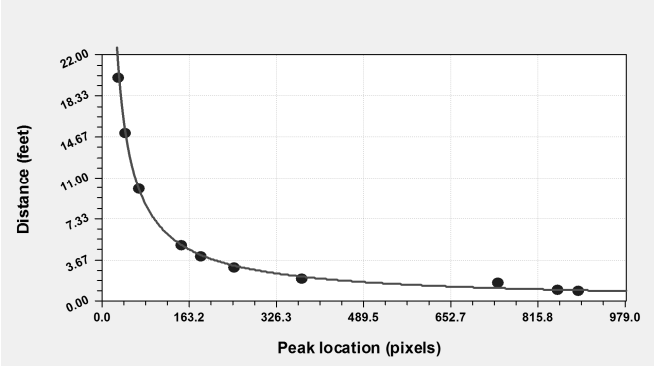

Figure 10: Distance vs peak location for 24.13 cm data using equation 4.

The resulting errors are much larger than when using a curve specific to the camera separation. For the 19.05 cm camera separation the minimum error increases from 0.316 inch (0.80 cm) to 1.735 inch (4.41 cm) while the maximum error increases from 4.716 inch (11.98 cm) to 81.057 inch (205.88 cm). For the 24.13 cm camera separation the minimum error increases from 0.028 inch (0.07 cm) to 1.86 inch (4.72 cm) while the maximum error increases from 6.411 inch (16.28 cm) to 32.336 inch (82.13 cm). The resulting errors are much larger than when using a curve specific to the camera separation. We also provide a plot of the 24.13 cm camera separation data using equation 5 in figure 11.

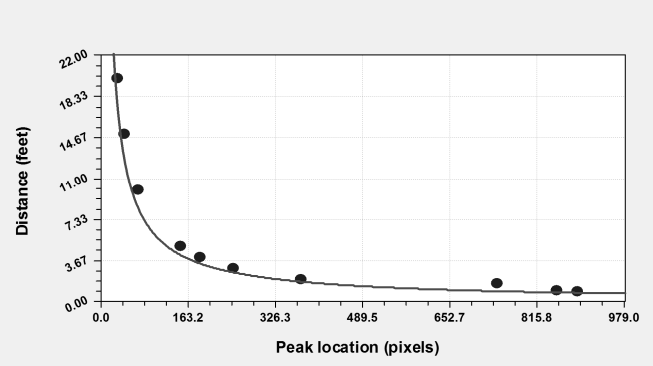

Figure 11: Distance vs peak location for 24.13 cm data using equation 5.

4. CONCLUSION

Work to date has achieved its goal of proving that distance can be recovered from a joint transform correlation peak. We have also shown that a relatively simply Harris curve can be used to extrapolate distances between known measured points with very good accuracy (15 cm or less). It is important to note that this formula was generated under the specific conditions, including a 73° field of view and the two separations used. As we have shown, the specific parameters making up the interpolation formula is very dependent on the camera separations; otherwise, as shown, the error becomes much larger, probably too large for most applications. Finally, at this time the method has not been tested in/onboard an application such as the CubeSat platform (for which this research was originally conceived). The work to date and near-future work is entirely proof of concept, and we may expect that applications follow once a strong foundation is laid and such satellite(s) can be fabricated and launched.

ACKNOWLEDGEMENT

This material is based upon work supported by the National Science Foundation Research Experiences for Undergraduates under Grant No. (NSF 1359244).

REFERENCES

1. R. Deepak and R. Twiggs, **Thinking out of the box: Space science beyond the CubeSat**, *Journal of Small Satellites*, vol. 1, no. 1, pp. 3–7, 2012.
2. J. Straub. **An open prototype for educational NanoSats: Increasing national space engineering productivity via a low-cost platform**, in *Proc. 2nd National Academy of Inventors Conference*, Tampa, FL., February 2013.
3. Openorbiter, http://www.openorbiter.und.edu, 2016.
4. A. Vander Lugt, **Signal detection by complex spatial filtering**, *IEEE Transactions on Information Theory*, vol. 10, pp. 139-145, 1964.
5. C. S. Weaver and J. W. Goodman, **A Technique for Optically Convolving Two Functions**, *Applied Optics*, vol. 5, pp. 1248-1249, 1966.
6. J. Barry and R. Tedrake, **Pushbroom stereo for high-speed navigation in cluttered environments**, in *Proc. IEEE International Conference on Robotics and Automation (ICRA)*, Seattle, WA, 2015, pp. 3046-3052.
7. H. Walcher, 1994. *Position sensing – Angle and distance measurement for engineers*, 2nd ed. Butterworth-Heinemann Ltd, Ch. 3.
8. M.A. Mahammed, A.I. Melhum, and F.A. Kochery, **Object Distance Measurement by Stereo Vision**, *International Journal of Science and Applied Information Technology (IJSAIT)*, vol. 2, no. 2, pp. 5-8, 2013.
9. M.Y. Kim, S.M. Ayaz, J. Park, and Y. Roh, **Adaptive 3D sensing system based on variable magnification using stereo vision and structured light**, *Optics and Lasers in Engineering*, vol. 55, pp. 113-127, 2014.
10. J. Mrovlje and D. Vrančić, **Distance measuring based on stereoscopic pictures**, in *Proc. 9th International PhD Workshop on Systems and Control: Young Generation Viewpoint*. 2008.
11. E. Tjandranegara, **Distance Estimation Algorithm for Stereo Pair Images**, Purdue ECE Tech. Rep. 64, West Lafayette, IN, 2005.
12. ImageMagick, http://www.imagemagick.org/, 2016.
13. S. Jin, S. Lee, **Joint transform correlator with fractional Fourier transform**, *Optics Communications*, vol. 207, pp. 161-168, 2002.
14. H. Wang, J. He, Z. Sheng, **Optoelectronic hybrid joint transform correlator based on power spectrum subtracting and exponential filtering**, in *Proc. SPIE Information Optics and Photonics Technology*, vol. 5642, 2005, pp. 445-450.
15. T. Nomura, **Phase-encoded joint transform correlator to reduce the influence of extraneous signals,** *Applied Optics***,** vol. 37, issue 17, pp. 3651-3655, 1998.
16. A.K Cherri, M.S. Alam, **Reference Phase-Encoded Fringe-Adjusted Joint Transform Correlation**, *Applied Optics*, vol. 40, issue 8, pp. 1216-1225, April 2001.
17. A. Layton and R. Marsh, **Object Distance Detection using a Joint Transform Correlator**, in *Proc. IEEE International Conference on Electro/Information Technology*, Grand Forks, ND, May 2016.