

Multiple-input joint transform correlator for wavelet feature extraction

Samuel P. Kozaitis, MEMBER SPIE
Florida Institute of Technology
Division of Electrical, Computer Science and
Engineering
150 W. University Boulevard
Melbourne, Florida 32901
E-mail: spk@ee.fit.edu

Mark A. Getbehead
Rome Laboratory
Photonics Center
25 Electronics Parkway
Rome, New York 13441

Abstract. We describe a joint transform correlator (JTC) that uses multiple input images encoded in the spatial domain for multiwavelet feature extraction. We extend the theory of a JTC to multiple inputs, which enables various combinations of cross-correlations between input images to be performed. Furthermore, we provide experimental results for four inputs with an optically addressed spatial light modulator in the Fourier plane. In addition, it is possible that the space-bandwidth product for multiwavelet feature extraction can be made the same as for a two-input JTC. © 1998 Society of Photo-Optical Instrumentation Engineers. [S0091-3286(98)03304-2]

Subject terms: correlator; feature extraction; joint transform correlator; optical pattern recognition; wavelet transform.

Paper 30067 received June 21, 1997; revised manuscript received Oct. 27, 1997; accepted for publication Oct. 28, 1997.

1 Introduction

The main difference between a joint transform correlator (JTC) and a conventional VanderLugt ($4f$) correlator is that the images in the JTC are encoded in the spatial domain. In terms of their output, both types of correlators perform the same operation, the cross-correlation between two images. In a $4f$ correlator, the Fourier transform of an input image is imaged onto the Fourier transform of a reference image. In a JTC, both input and reference images are input simultaneously in the spatial domain, then the Fourier transform is performed. Compared to the $4f$ correlator, the JTC has advantages such as ease of alignment and avoidance of spatial filter synthesis.^{1,2}

Although most configurations of the JTC use two input images, the use of multiple input images may enable additional functions to be performed. For example, an image and two different wavelets were used in a JTC for multispectral (multiwavelet) wavelet feature extraction of the image.³ In this approach, different versions of an input image corresponding to different wavelet scales appeared at different locations in the output plane. Another JTC configuration used three inputs, but required an additional electronic optical processing step to implement the wavelet transform.⁴ Another approach used⁵ two SLMs with a holographic mask to produce a wavelet-based JTC. In addition, a JTC was used with a single reference image and multiple inputs,⁶ and color pattern recognition was achieved with a multichannel approach.⁷ Although the wavelet transform has provided an application for a JTC with multiple inputs,⁸ a generalized JTC with multiple inputs may find more applications, including those in other areas, if its output function is properly described.

We describe a JTC that uses multiple input images encoded in the spatial domain. In the next section we briefly describe the general theory of a conventional JTC, and then extend the discussion to multiple inputs. Next, we provide some experimental results to verify the results of the

theory. We also describe how a multiple-input JTC can be used for multiwavelet analysis and propose an architecture that has the same space-bandwidth product as a two-input JTC.

2 Conventional JTC

To perform the correlation operation with a JTC, functions are encoded in the input plane. A schematic diagram of a conventional JTC is shown in Fig. 1. To perform the cross-correlation between the images $b(x,y)$ and $d(x,y)$, they are centered at $x = \pm \alpha$. A lens produces the Fourier transform when the input plane is illuminated with coherent light. In the Fourier plane, the complex light field is

$$U = B(p,q)\exp(j\alpha p) + D(p,q)\exp(-j\alpha p), \quad (1)$$

where $B(p,q)$ is the Fourier transform of $b(x,y)$, and similarly for $d(x,y)$. A square-law device such as a liquid crystal light valve is placed in the Fourier plane before another lens performs the Fourier transform. Alternatively, a camera can be used to record the power spectrum and display it on a spatial light modulator (SLM) before the Fourier transform is performed. The output intensity distribution from a square-law detector can be written as

$$\begin{aligned} |U|^2 = UU^* = & [B(p,q)\exp(j\alpha p) + D(p,q)\exp(-j\alpha p)] \\ & \times [B(p,q)\exp(j\alpha p) + D(p,q)\exp(-j\alpha p)]^*. \end{aligned} \quad (2)$$

Multiplying terms, taking the Fourier transform, and grouping terms results in

$$\begin{aligned} I = & b(x,y) \otimes b(x,y) + d(x,y) \otimes d(x,y) + b(x,y) \\ & \otimes d(x+2\alpha, y) + d(x,y) \otimes b(x-2\alpha, y), \end{aligned} \quad (3)$$

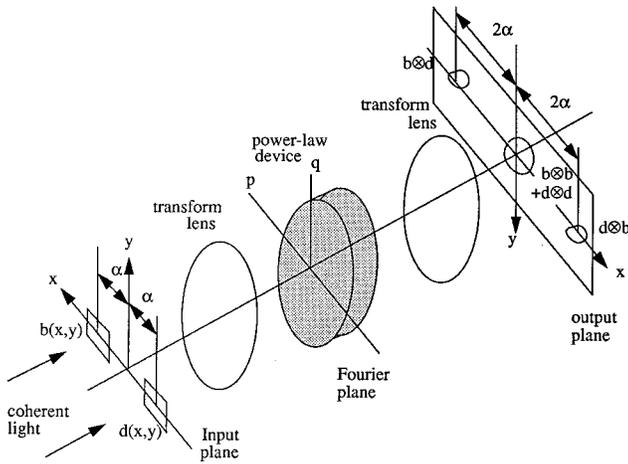


Fig. 1 Schematic diagram of a conventional JTC.

in the output plane, where \otimes indicates the correlation operation. The first two terms are the autocorrelations of the input functions and appear on the optical axis. The third and fourth terms are the cross-correlation between the two input functions and appear at $x = \pm 2\alpha$, as shown in Fig. 1.

3 Multiple-Input JTC

We consider a JTC with an arbitrary number of input images arranged along a line. We use a JTC as shown in Fig. 1 but consider n inputs separated by α in the input plane arranged along the x axis, as shown in Fig. 2. The images are labeled $a_1(x,y)$ to $a_n(x,y)$, with the center image labeled as $a_{(n+1)/2}(x,y)$. Using this configuration, the complex light field in the Fourier plane is

$$\begin{aligned}
 U = & A_1(p,q) \exp \left[j \left(\frac{n-1}{2} \right) \alpha p \right] \\
 & + A_2(p,q) \exp \left[j \left(\frac{n-3}{2} \right) \alpha p \right] \cdots A_{(n+1)/2}(p,q) \\
 & + \cdots A_n(p,q) \exp \left[-j \left(\frac{n-1}{2} \right) \alpha p \right].
 \end{aligned} \tag{4}$$

The output intensity distribution from a square-law detector is written as $|U|^2 = U \times U^*$. Multiplying $n \times n$ terms,

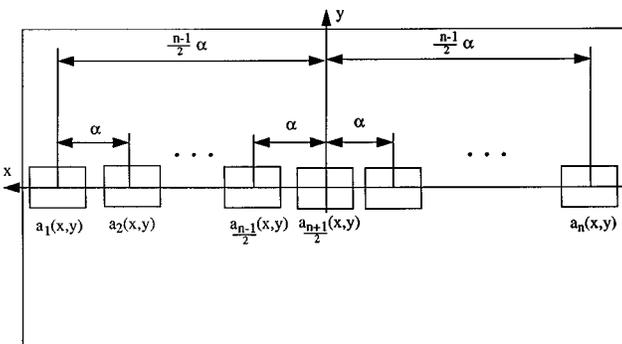


Fig. 2 Input plane of a multiple-input JTC.

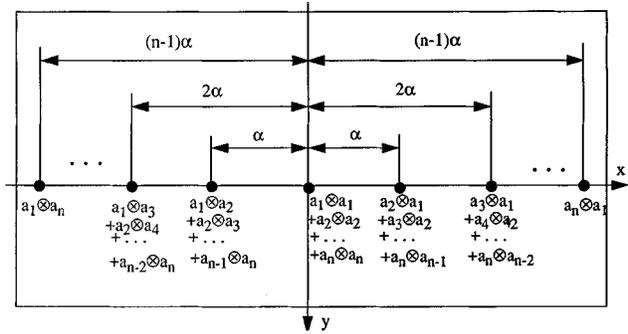


Fig. 3 Output plane of a multiple-input JTC.

taking the Fourier transform, and grouping terms resulted in $2n - 1$ locations in the output plane where a correlation response would occur. The output plane was described as

$$\begin{aligned}
 I = & \left[\sum_{i=1}^n a_i(x,y) \otimes a_i(x,y) \right] + \left\{ \sum_{i=1}^{n-1} a_i(x,y) \otimes a_{i+1}(x + \alpha, y) \right. \\
 & + \sum_{i=1}^{n-2} a_i(x,y) \otimes a_{i+2}(x + 2\alpha, y) \\
 & + \cdots \left. \sum_{i=1}^{n-(n-1)} a_i(x,y) \otimes a_{i+(n-1)}[x + (n-1)\alpha, y] \right\} \\
 & + \left\{ \sum_{i=1}^{n-1} a_{i+1}(x,y) \otimes a_i(x - \alpha, y) + \sum_{i=1}^{n-2} a_{i+2}(x,y) \right. \\
 & \otimes a_i(x - 2\alpha, y) \cdots \sum_{i=1}^{n-(n-1)} a_{n-1}(x,y) \\
 & \left. \otimes a_{i+(n-1)}[x - (n-1)\alpha, y] \right\}.
 \end{aligned} \tag{5}$$

The first term in brackets in Eq. (5) contains the dark current (DC) terms, which is the sum of the autocorrelations of all the input images. The second term in braces corresponds to $n - 1$ terms to the left of the DC term, which are shown in Fig. 3; each of these terms is separated by a distance α . The third term is similar to the second but corresponds to the right of the DC term in Fig. 3. In addition, the terms on each side of the DC term are the same but rotated by π rad as in a conventional JTC.

4 Experimental Results with the Multiple-Input JTC

We obtained experimental results to verify the operation of a multiple-input JTC. We used a JTC with an input displayed on film and an optically addressed SLM in the Fourier plane. We used a model SPT-25 optically addressed SLM operating in a transmissive mode, manufactured by Micro-Optics Technologies, Inc. To operate the device, light of 400 to 500 nm is used as write light that is absorbed by a photosensor layer. Light of 600 to 1100 nm is used as

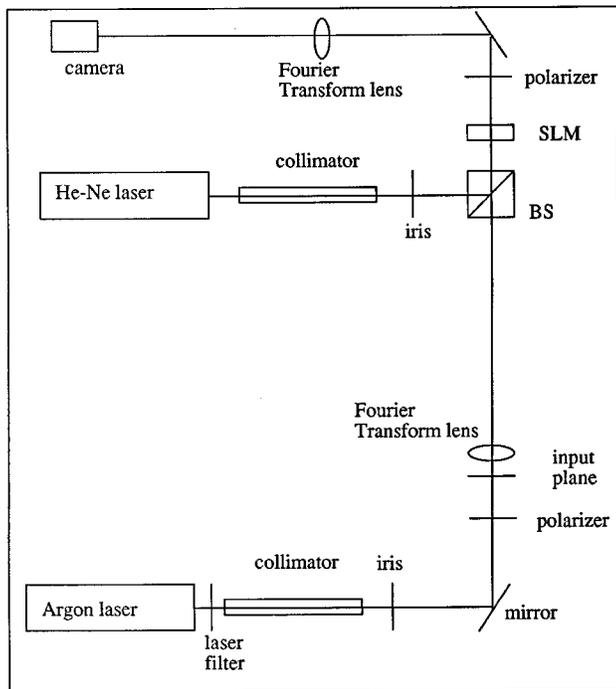


Fig. 4 Experimental setup of a multiple-input JTC.

the read light that carries the write-light information. A dc bias of 2 to 15 V is placed across the device for proper operation.

Our experimental setup is shown schematically in Fig. 4. The argon laser uses a $10\times$ microscope objective, a $25\text{-}\mu\text{m}$ pinhole, and an $f=250$ nm lens in the spatial filter assembly. The filter is a narrow-band laser filter centered at 515 nm. Note that this wavelength is outside of the specification of the SLM, but we obtained better results with this line than with the 488-nm line of the laser. The illuminating light just before the input plane has a power of 4.0 mW and both Fourier transform lenses have a focal length of $f=500$ mm. The He-Ne laser used as the read light of the SLM uses a $10\times$ microscope objective, a $10\text{-}\mu\text{m}$ pinhole, and an $f=250$ mm lens in the spatial filter assembly. The voltage across the SLM is 12.5 V.

The input plane consists of four of the same images as shown in Fig. 5 where $\alpha=1.5$ mm. The images are each labeled as $f(x,y)$ and substituted into Eq. (5); the expected output plane is described schematically as in Fig. 6. The output plane consists of a bright spot on the optical axis and four responses on either side. The responses located at distances α and 3α from the optical axis have the same intensity, as do the responses located at 2α and 4α . The responses at α and 3α differ from those at 2α and 4α by a factor of 2, which would correspond to a factor of 4 if the power spectrum was detected.

The experimental result corresponding to the left half of Fig. 6 is shown in Fig. 7(a). The optical axis corresponding to the center of Fig. 6 is shown to the far right of Fig. 7(a). Two bright spots along the x axis correspond to the autocorrelation responses located at distances α and 3α from the optical axis. Two additional autocorrelations appear at distances of 2α and 4α from the optical axis. A profile plot



Fig. 5 Input image used in experiment.

along the x axis through the autocorrelation peaks is shown in Fig. 7(b) where the peaks can be seen more clearly.

5 Multiwavelet Feature Extraction

5.1 Wavelet Transform

The wavelet transform has been shown to be useful in many areas including pattern recognition.⁹ The wavelet transform can be thought of as a correlation between a wavelet and an input signal. The 2-D wavelet transform of an image $f(x,y)$ can be written as

$$W_f(a,b) = \iint f(x,y)w_{ab}^*(x,y) dx dy, \quad (6)$$

where $w_{ab}(x,y)$ is a set of wavelets that are dilated and shifted versions of a wavelet $w(x,y)$, where

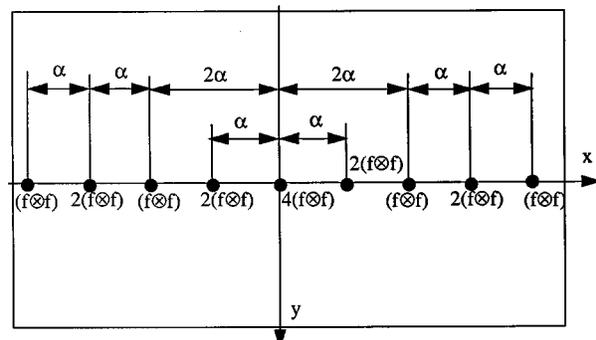


Fig. 6 Location and relative value of correlation response of a multiple-input JTC when Fig. 5 is used as the input.

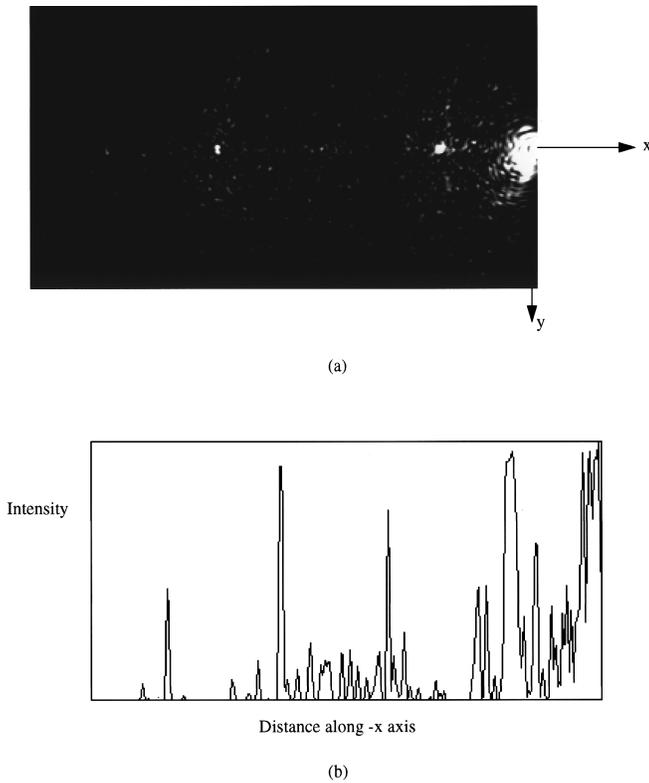


Fig. 7 Left-hand side of output plane obtained experimentally when Fig. 5 is used as the input plane (a) intensity distribution and (b) intensity plot of the x axis.

$$w_{ab}^*(x,y) = \frac{1}{a} w\left(\frac{x-b_x}{a}, \frac{x-b_y}{a}\right), \quad (7)$$

where a is a scale factor, and b_x and b_y are translations in the x and y directions, respectively. Because a wavelet can be thought of as a bandpass filter, an image can be viewed in a particular scale or frequency band after cross-correlation with a wavelet. The cross-correlation result is the image corresponding to the frequency band of the wavelet; it is not necessary to perform an additional inverse transform.

To display a version of an input image that corresponds to one wavelet scale using a JTC, one input is the image of interest and the other is the wavelet function.¹⁰ The resulting image then appears at the output of the JTC. Because wavelets have zero mean, this may present a difficulty in the implementation. To remove the DC component from the wavelet function, it was experimentally shown that the wavelet function could be encoded in phase.³

To view an image corresponding to more than one frequency band, the image should be cross-correlated with differently scaled wavelets separately. The results of the cross-correlations are summed to produce an image corresponding to multiple frequency bands. If the sum is coherent, then positive and negative contributions to the output image are possible, which is necessary for the general case.

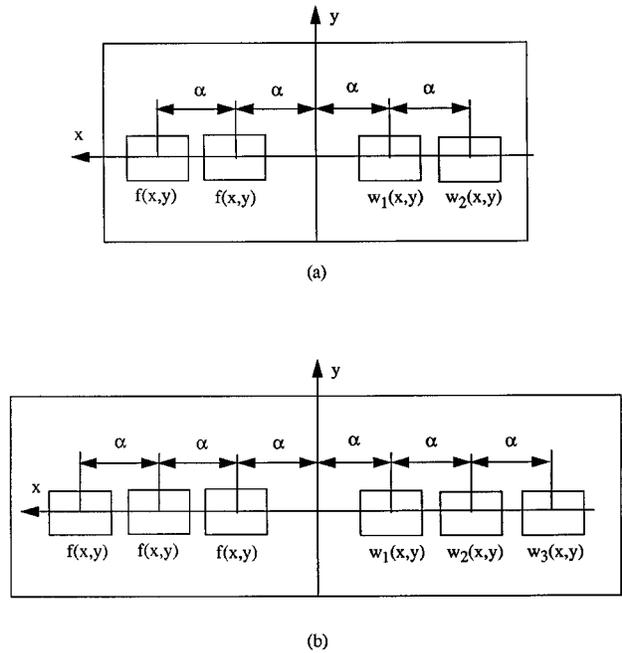


Fig. 8 Input plane for multiwavelet processing (a) two and (b) three wavelet scales.

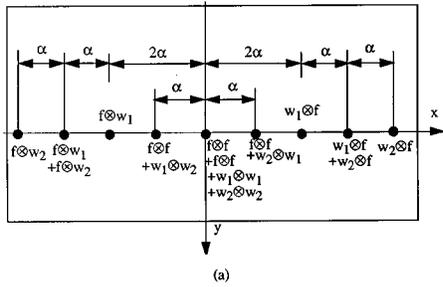
5.2 Multiwavelet Analysis

Under certain conditions the input arrangement in Fig. 2 could produce a version of an input image corresponding to multiple wavelet scales. We considered an example that produced a version of an input image that corresponded to two wavelet scales. In other words, the output image corresponded to the sum of the cross-correlations of the input image and two different wavelets. To perform this type of operation we referred to Fig. 2 and set $n=5$ and $a_3(x,y) = 0$. We considered the input image $f(x,y)$ at locations $a_1(x,y)$ and $a_2(x,y)$, so $f(x,y) = a_1(x,y) = a_2(x,y)$. Finally, we set a wavelet corresponding to one scale $w_1(x,y) = a_4(x,y)$, and the wavelet at another scale $w_2(x,y) = a_5(x,y)$, as shown in Fig. 8(a).

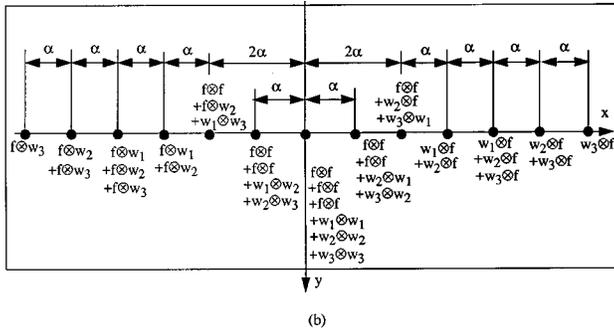
The response in the output plane was obtained by substituting the appropriate variables in Eq. (5), and is represented schematically as shown in Fig. 9(a). The third term from the DC is the term of interest here. It is the coherent sum of correlations of the input image and two different wavelets at two different scales. A diagram of the input and output planes for three wavelet scales are shown in Figs. 8(b) and 9(b).

In general, a version of an input image corresponding to m wavelet scales could be produced. A total of $2m$ inputs are needed, m replicated input images on one side of the optical axis and m wavelet images on the other. There would be $(4m - 2)/2$ correlation results appearing on each side of the central autocorrelation peak. The m 'th response from the far end of either side would contain the sum of the images at the different scales.

The separation of the input and wavelet images must be far enough apart so that the output images do not overlap. We considered the maximum widths of the input image and wavelet functions to be D_i and D_w , respectively, where



(a)



(b)

Fig. 9 Output plane for multiwavelet processing (a) two and (b) three wavelet scales.

$D_i > D_w$. Assuming the system has unit magnification, the maximum sizes in the output plane for the input image autocorrelation and input image-wavelet cross-correlations are $2D_i$ and $D_i + D_w$, respectively. In general, it can be seen [from Fig. 9(a), for example] that to separate all the output images, the equality

$$\alpha > 1/2(4D_i) \tag{8}$$

must hold. However, we are primarily interested in the off-axis terms, which do not include autocorrelations of the input scenes. In addition, the input images are the same and if an object within them moves, they will move together. Therefore, the condition in Eq. (8) will correspond to a separation between images in the input plane of

$$\alpha > D_i + 2D_w \tag{9}$$

to avoid overlap in the output plane.

5.3 Simulation

To simulate the result we used a Bessel-Gaussian wavelet³ described as

$$\frac{1}{\sqrt{a}} w\left(\frac{r}{a}\right) = \frac{1}{\sqrt{a}} J_0\left(\frac{r}{a}\right) \exp\left[-\left(\frac{r}{2a\sigma}\right)^2\right], \tag{10}$$

where σ is related to the width of the bandpass response, a is the scale factor, and r is the distance from the origin.

We performed simulations with an image of an airplane $f(x,y)$ placed into the configuration of Fig. 8(a) using a total of 512×512 pixels. The image of the plane contained 50×50 pixels and α was set to 50 pixels. For $w_1(x,y)$ and



Fig. 10 Center portion of input plane used in simulation experiments.

$w_2(x,y)$, values of $a_1 = 0.7$ and $a_2 = 1.2$ were used, respectively. The center portion of the input plane is shown in Fig. 10 before subtraction of the DC of the wavelet images. Although the input images are placed a distance D_i apart in violation of Eq. (9), the airplanes are separated by a distance of about $D_i + D_{w1} + D_{w2}$, where D_{w1} and D_{w2} are the widths of the wavelet functions.

The power spectrum is shown in Fig. 11(a) and the output plane is shown in Fig. 11(b) and consists of nine responses. The three responses, consisting of that along the optical axis and those immediately to the right and left of the optical axis at $\pm\alpha$, were autocorrelation responses. They all have large values and overlap with each other. In addition, they overlap the responses located $\pm 2\alpha$ from the optical axis. Therefore, the large spot in Fig. 11(b) consists of five correlation responses and was shown for different values of thresholds in Fig. 12.

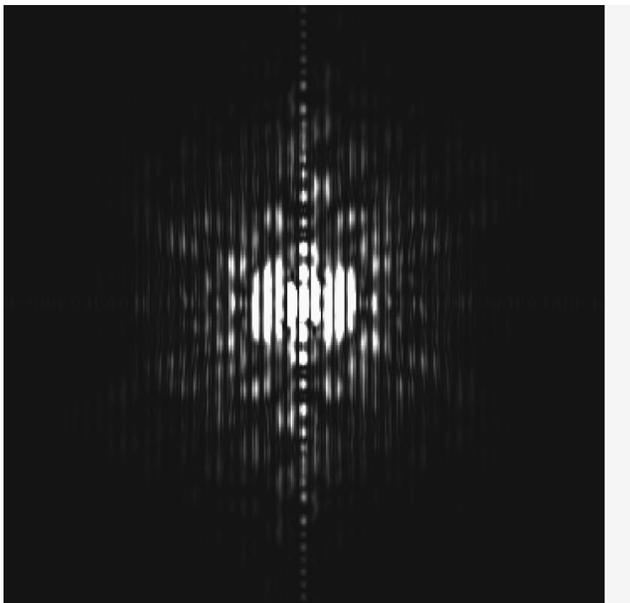
The remaining correlations are those of interest and consist of two on each side of the large central spot. The two remaining responses on the left-hand side of Fig. 11(b) are expanded and shown in Fig. 13(a). The response on the left corresponds to the wavelet with $a_2 = 1.2$ and that on the right corresponds to the sum of the wavelets with a_1 and a_2 . Changing a_1 to $a_1 = 1.0$ changes the response on the right but not on the left, as shown in Fig. 13(b). A small part of the correlation between the image and the wavelet associated with a_1 can be seen in the left side of the image.

5.4 Implementation

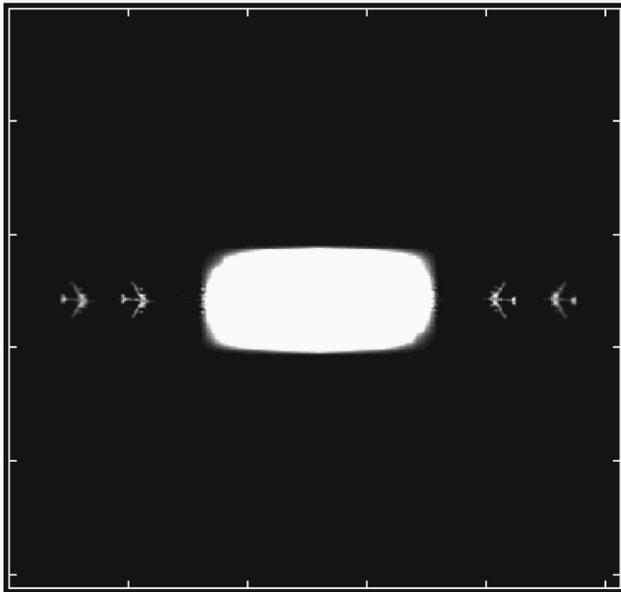
To implement the multiple-input JTC optically a few points must be considered. One is the removal of the DC from the wavelet image. Wavelets have zero mean, so the wavelet image cannot be used directly. One solution is to display the wavelet in phase³ or use a DC block arrangement.

Because more than one scale of a wavelet is used in our arrangement, both multiple wavelet and multiple input images, one each for every scale of the wavelet used. The most straightforward approach would be to display all images on an SLM and use a single lens in the JTC. Although wavelets and their scales could be changed dynamically, the large space-bandwidth product and the size of Fourier transform lens would limit the number of wavelet scales that could be used. To reduce the restriction on the lens, the input and wavelet portions of the input plane can be split and focused separately each with a Fourier transform lens.¹¹

For the input image side, the image needs to be split into multiple images. This could be performed by conventional optics using beamsplitters and mirrors. Alternatively, a holographic element could be used to split the input images and possibly include the lens as schematically shown in Fig. 14. A similar arrangement could be used on the wavelet side. The wavelet itself is fixed and does not move;



(a)



(b)

Fig. 11 Output of multiwavelet simulation (a) power spectrum and (b) output plane.

however, the selection of different wavelet scales, or different wavelets, that could be changed dynamically would be desirable. It is possible that a spatially multiplexed hologram could be used. This might be accomplished by a deflection system on the beam that is incident on the holographic element.

6 Conclusion

We showed how a JTC can be used with multiple input images and how a multiple-input JTC can be used for multiwavelet analysis of an input image. Using an image and wavelet as inputs, for m wavelet scales, m versions of the wavelet and m copies of the input image must be generated.

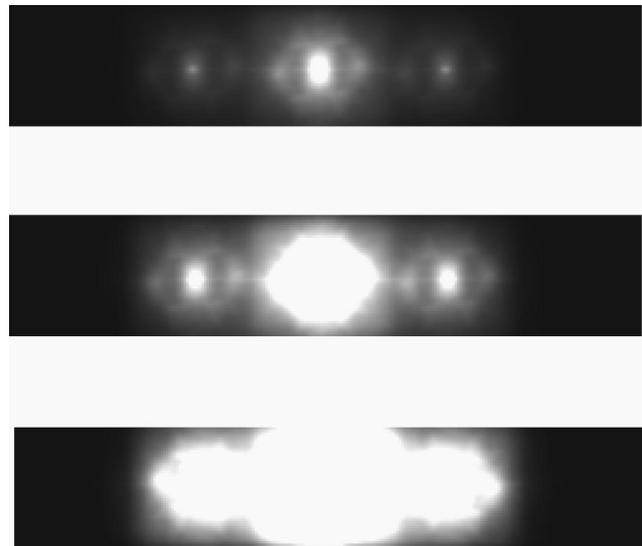
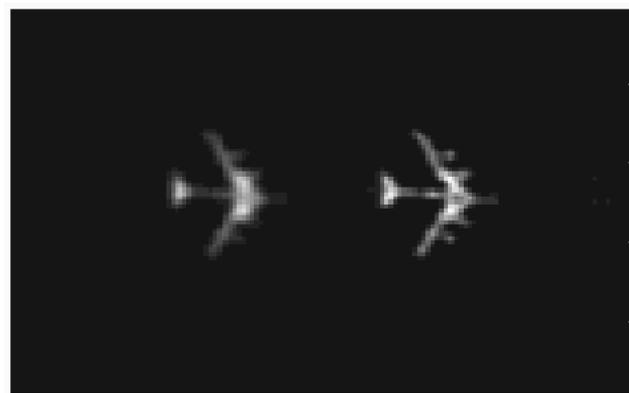


Fig. 12 Center portion of output plane thresholded at three different values showing the overlap of correlation responses.



(a)



(b)

Fig. 13 Close-up view of left-hand side of Fig. 11. Responses on the left in both cases are for $a_2=1.2$: (a) response on the right for a_1+a_2 , where $a_1=0.7$, and (b) response on the right for a_1+a_2 , where $a_1=1.0$.

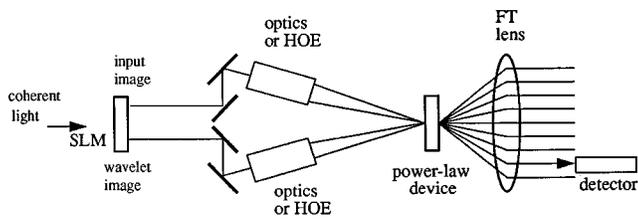


Fig. 14 Schematic diagram of multiple-input JTC used for multi-wavelet analysis.

The output consisted of $4m - 1$ correlation results, one of which is the desired output. Using holographic elements, the space-bandwidth product of the system can be made the same as for a two-input JTC. In addition, the relationship between the wavelet scales can be made independently by increasing the space-bandwidth of the system. Because an output image can be produced that corresponds to multiple wavelet scales, the system is potentially more useful than one that corresponds to a single scale.

References

1. C. S. Weaver and J. W. Goodman, "A technique for optically convolving two functions," *Appl. Opt.* **5**, 1248–1249 (1966).
2. B. Javidi and J. L. Horner, Eds., *Real-Time Optical Information Processing*, Academic Press, San Diego, CA (1994).
3. P. S. Erbach, D. A. Gregory, and X. Yang, "Optical wavelet transform by the phase-only joint-transform correlator," *Appl. Opt.* **35**(17), 3117–3126 (1996).
4. X. Yang, N. P. Cavaris, and M. Wen, "Optical wavelet correlators for cluttered target identification," in *Optical Pattern Recognition V, Proc. SPIE 2237*, 402–418 (1994).
5. J. Li, Y. Zhang, and J. Hu, "Object recognition with a wavelet-transform-based joint transform correlator," *Opt. Eng.* **35**(3), 775–777 (1996).
6. J. Wang and B. Javidi, "Multiobject detection using the binary joint transform correlator with different types of thresholding methods," *Opt. Eng.* **33**, 1793–1804 (1994).
7. M. Deutsch, J. Garcia, and D. Mendlovic, "Multichannel single-output color pattern recognition by use of a joint-transform correlator," *Appl. Opt.* **35**(35), 6976–6982 (1996).
8. S. P. Kozaitis, M. A. Getbehead, and W. E. Foor, "Multispectral image feature extraction by the joint wavelet-transform correlator," in *Optical Pattern Recognition VIII*, D. P. Casasent, Ed., *Proc. SPIE 3073*, 322–331 (1997).
9. M. B. Ruskai, B. Beylkin, R. Coifman, I. Daubechies, S. Mallat, Y. Meyer, and L. Raphael, *Wavelets and Their Applications*, Jones & Bartlett, Boston, MA (1992).
10. W. Wang, G. Jin, Y. Yan, and M. Wu, "Joint wavelet-transform correlator for image feature extraction," *Appl. Opt.* **34**(2), 370–376 (1995).
11. T. C. Lee et al., "Dual-axis joint-Fourier-transform correlator," *Opt. Lett.* **4**, 121–123 (1979).



Samuel P. Kozaitis received his PhD in electrical engineering in 1986 from Wayne State University, where he was an assistant research professor. He has also been with General Motors Research Laboratories, was a research fellow at the Photonics Center of the Rome Laboratory from 1988 to 1997, spent two summers at the National Aeronautics and Space Administration Kennedy Space Center, and has performed consulting for government and industry. He is currently an associate professor in the Division of Electrical, Computer Science, and Engineering at the Florida Institute of Technology. His major research interests include optical pattern recognition, wavelet transforms, image processing, and higher order statistics.



Mark A. Getbehead received his BSEE degree from Valparaiso University in 1983 and his MS in computer science from the State University of New York Institute of Technology in 1994. He has been a research engineer with Rome Laboratory since 1983, and since 1991, in the Photonics Center, where he is a project leader in an optical wavelet processing effort. His interests include wavelet transform applications, optical pattern recognition, and image compression.