PROCEEDINGS OF SPIE

SPIEDigitalLibrary.org/conference-proceedings-of-spie

Feature-based correlation filters for object recognition

Samuel Peter Kozaitis
Wesley E. Foor

SPIE.

Feature-based correlation filters for object recognition

S. P. Kozaitis
Florida Institute of Technology
Department of Electrical and Computer Engineering
Melbourne, FL 32901-6988

W. E. Foor
Rome Laboratory
Photonics Center
Griffiss AFB, NY 13441

ABSTRACT

Using an optical correlator, we experimentally evaluated a binary phase-only filter (BPOF) designed to recognize objects not in the training set used to design the filter.Such a filter is essential for recognizing objects from actual sensors. We used an approach is that is as descriptive as a BPOF yet robust to object and background variations of an unknown or nonrepeatable type. We generated our filter by comparing the values of spatial frequencies of a training set. Our filter was easily calculated and offered potentially superior performance to other correlation filters.

1.0 INTRODUCTION

The major difficulty encountered when using a binary phase-only filter (BPOF) in an optical correlator is the filter's sensitivity to changes in the object's appearance. In reality, imagery from sensors is often nonrepeatable; the distortion cannot be repeated exactly. Therefore, the creation of BPOFs to recognize objects that vary in a nonrepeatable way is essential for recognizing objects from actual sensors. Recently, an approach based on preprocessing an image and using a binary filter was applied to aircraft identification from actual sensors. Digital image processing techniques were used to confine images to vary in a limited but unknown manner, then filters were used to identify objects.[1] An approach is required that is as descriptive as a BPOF yet robust to object or background variations of an unknown or nonrepeatable type.

We investigated the use of a BPOF that was calculated from an object's features. In this way, we attempted to make a BPOF more robust to unknown variations than other designs. We have previously shown by computer simulation that our filter offered superior performance to SDF filters[2] for our problem; here, we provide experimental results for a version of our filter. We considered a one-class problem; our results were generated using a training set from only one class of objects. To help evaluate the potential of our approach, we used imagery from actual sensors that were not from the original training set. The next section briefly discusses the filter formation. Next, our optical system and experimental results are discussed followed by conclusions.

0-8194-0969-3/93/$4.00

2.0 FILTER FORMULATION

In actual sensor imagery, the global shape of an object is frequently too distorted to generate a specific version of the object. Therefore, an input object may not correlate well with a filter even though the input and filter are from the same class.[1-3] In contrast to SDF filter formulation, we developed a filter whose values were determined by features of the objects in a training set. We attempted to find a filter that represented the critical characteristics of a class of objects so that objects outside the training set but in the same class could be identified. Therefore, we examined features that were invariant with respect to the training set.

Generally, the cross-correlation between two images is maximized when the mean squared error (MSE) between the images is minimized.[4] The correlation operation measures the similarity of images; therefore, images will correlate well if their Euclidean distance in signal space is small. In signal space, an image is represented as a point and each axis may represent a spatial frequency. A point along an axis represents the value of that spatial frequency. The region in signal space that represents images that correlate well with a given image is generally a multidimensional sphere centered on the given image as shown schematically in Fig. 1. The radius of the sphere is determined by a threshold in the correlation plane where any correlation response above the threshold is considered to be a match with the given image. As the threshold decreases, the radius of the sphere will increase.

We attempted to form n training images into a cluster in signal space by retaining only spatial frequencies with a small spread of values. We examined the Discrete Fourier Transforms (DFTs) of the training images at each spatial frequency. The DFT of the kth training image was represented by $S_k[u,v]$ where u,v are discrete spatial frequencies. The values of the spatial frequencies were examined across the entire training set in terms of their similarity. We considered the distance between their values in the complex plane as a measure of their similarity. The smaller the distance, the more similar the values.

2.1 Binary phase-only filter

One way BPOFs are different than both POF and matched filters is in how they are quantized. Only two possible values exist to describe a BPOF at a spatial frequency. If two BPOFs had the same value at a particular spatial frequency, then the distance for that spatial frequency was zero, if the values were different then the distance was one. This rule was described by the exclusive-OR function as

$$d_{BPOF}(x_1, x_2) = \left(\frac{(\angle x_1) - \theta}{\pi}\right) \oplus \left(\frac{(\angle x_2) - \theta}{\pi}\right), \tag{1}$$

where x_1 and x_2 represent spatial frequencies of a BPOF that had been thresholded along an axis of θ and had a magnitude of 1 with a phase value of either θ or $\pi + \theta$. We used a criterion function to measure the quality of a cluster of values as

$$J_{u,v} = \frac{1}{n^2} \sum_{k=1}^{n} \sum_{l=1}^{n} d_{BPOF}(x_k, x_l) \tag{2}$$

which determines the average distance between all training images for a particular spatial frequency. Equivalently, the criterion function was based on the number of occurrences of a +1 or -1 at a particular spatial frequency across a training set and written as

$$J'_{u,v} = \sum_{k=1}^{n} H_k[u,v] \tag{3}$$

where $H_k\,[u, v]$ is the BPOF of the kth training image. For the function J'_{ij}, a tightly grouped cluster is indicated by a large value.

Sometimes features or spatial frequencies may offer conflicting information. For example, large values of $J_{u,v}$ or small values of J'_{ij} may not be useful for identification. Because some spatial light modulators (SLMs) can implement a zero state, we avoided the use of spatial frequencies with small values of J'_{ij} by setting its value to zero. Therefore, we didn't use spatial frequencies that offered conflicting information and generated a ternary-valued filter. We described the filter as

$$H_p[u,v] = \left\{ \begin{array}{c} 1 if\ p \le J'_{u,v} \\ -1\ if(-p) \ge J'_{u,v} \\ 0\ otherwise \end{array} \right\} \tag{4}$$

where p is the threshold that determines the order of the filter.

3.0 EXPERIMENT

3.1 Optical system

We used an binary optical correlator with two magneto-optic SLMs. We used a 256 x 256 pixel device in the input plane and a 128 x 128 pixel device in the Fourier plane. Although such a configuration may have its advantages,[5] we used only the central 128 x 128 pixels of the input SLM so we could evaluate the use of our filters without affects due to the difference in the number of pixels between the SLMs. Because the pixel sizes of the two SLMs were the same, our system operated as if two 128 x 128 devices were used; both SLMs had square pixels of 56μm on a side with 76μm center-to-center spacing. We used a He-Ne laser with λ = 633nm and the correlation result was recorded by a CCD camera that used 256 x 256 pixels and an 8-bit digitizer.

Infrared (IR) imagery (8-14 μm) of ground scenes from actual sensors were used to evaluate and generate filters. Images were digitized with 128 x 128 pixels with 8 bits/pixel. Because the application was to binary SLMs, the imagery ultimately had to be thresholded. We used objects of the same scale and rotation to examine the effects of unknown variations of an inputs. We thresholded the images as shown in Fig. 2.

We used a filter generated as in Eq. (4) from 5 training images. The BPOF for each training image was calculated; hen, the BPOFs of all the training images are added together pixel by pixel. If the value at a pixel of the array was greater than or equal to 1 then that pixel was set to +1. If the value at a pixel of the array was less than or equal to - 1, then the pixel was set to -1.

3.2 Results

We first compared the results of the different filters by computer simulation with images from the training set. The results are normalized by setting the maximum correlation height for each filter equal to 100. The maximum value for the feature-based filters for $p = 1$, 3, and 5 were 68.9%, 57.8% and 34.2% the maximum of the fSDF filter response. The results for images within the training set are shown in Table 1. The average correlation height produced by the fSDF was higher than the feature-based filter for $p = 1$, but lower when $p = 3$ or 5. As p increased to 5, the average correlation height increased. The feature-based filter produced more consistent results than the fSDF filter at the expense of decreased and broader correlation peaks.

Using the optical system described earlier, we performed some experiments using the same training set as in the simulation. We did not generate ternary filters, so we only compared the fSDF filter to the feature-based filter for $p = 1$. The maximum correlation height of the feature-based filter was 90% that of the fSDF filter. The remaining results are shown in Table 1. The average PRMSR was 11.0 db for the fSDF filter and 10.9 db for the feature-based filter. A plot of the simulated and experimental results for the an image in the training set and the feature-based filter are shown in Fig. 3.

TABLE 1. Results of experiments on training set

	Filter					
	Simulation				Experiment	
	fSDF	Feature-based p = 1	Feature-based p = 3	Feature-based p = 5	fSDF	Feature-based p = 1
Avg. normalized correlation height	81.6	78.2	85.0	92.7	88.3	83.2
Percentage of transmitting pixels	100	100	46.2	13.6	100	100
Avg. SNR (% of fSDF response)	1.0	0.66	1.06	2.3	1.0	0.85

Because we were interested in identifying objects based on their features, we tested the filter with objects that were not in the training set but in the same class. We chose five images that were visually similar to the training set. Using the fSDF filter, they produced an average normalized correlation height of 61.4% of the maximum for that filter. Using the feature-based filter, the average correlation height was 69.8%. We also chose four images that were similar to the training set but were in the same class. The normalized correlation heights were 39.0 and 47.1 for the fSDF and feature-based filter.

4.0 DISCUSSION AND CONCLUSION

We experimentally showed that in the case where all pixels are transmitting, an fSDF and feature-based filter offered similar performance. The feature-based filter was slightly more consistent for objects within the training set than the fSDF filter. Furthermore, the fSDF filter offer better discrimination. Because our experimental results generally agreed with our computer simulations, we expect feature-based filters where some pixels are set to 0 to give superior performance than SDF filter for recognizing objects with similar features of a training set.

A feature-based filter was more easily calculated and trained than an fSDF filter. The feature-based filter required about an order of magnitude fewer calculations than the fSDF filter. In addition, new images added to the training set required the fSDF filter to be completely recalculated which was not the case in the feature-based filter.

5.0 ACKNOWLEDGMENT

This work was performed in the Photonics Center of Rome Laboratory at Griffiss AFB, NY.

6.0 REFERENCES

[1] S. P. Kozaitis and W. E. Foor, "Performance of synthetic discriminant functions for binary phase-only filtering of thresholded imagery," *Optical Eng*ineering 31(4) 830-837(1992)

[2] S. P. Kozaitis and W. E. Foor, "Distortion-invariant correlation using nonlinear feature-based binary phase-only filters," *Optical Information Processing Systems and Architectures IV*, B. Javidi, Ed., Proc. SPIE (1992)

[3] S. P. Kozaitis, S. Halby, and W. Foor, "Experimental performance of a binary phase-only optical correlator using visual and infrared imagery," *Optical Information Processing Systems and Architectures*, B. Javidi, Ed., Proc. SPIE 1296, 140-151 (1990)

[4] P. Maragos, "Morphological correlation and mean absolute error criteria," *Proc. of ICASSP* '89, IEEE, 1568-1571 (1989)

[5] S. P. Kozaitis and W. E. Foor, "Optical correlation using a filter reduced in resolution," *Optical Engineering*, 31(9), (Sept. 1992)

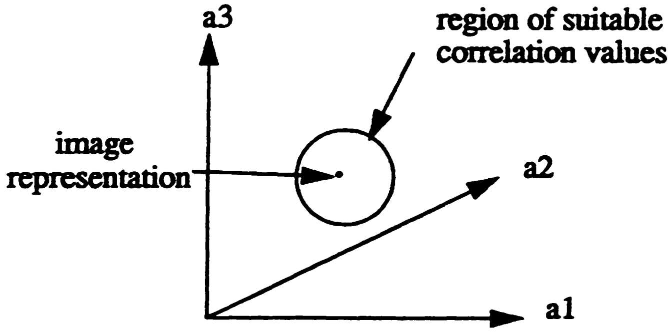


FIGURE 1. Signal space representation of cross-correlation with a image

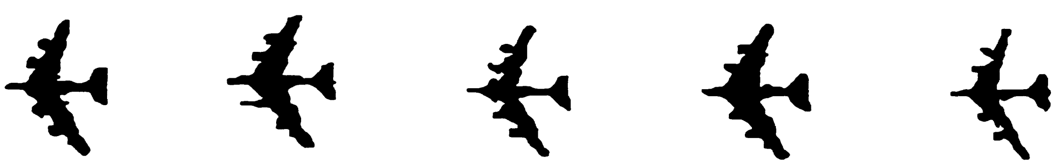
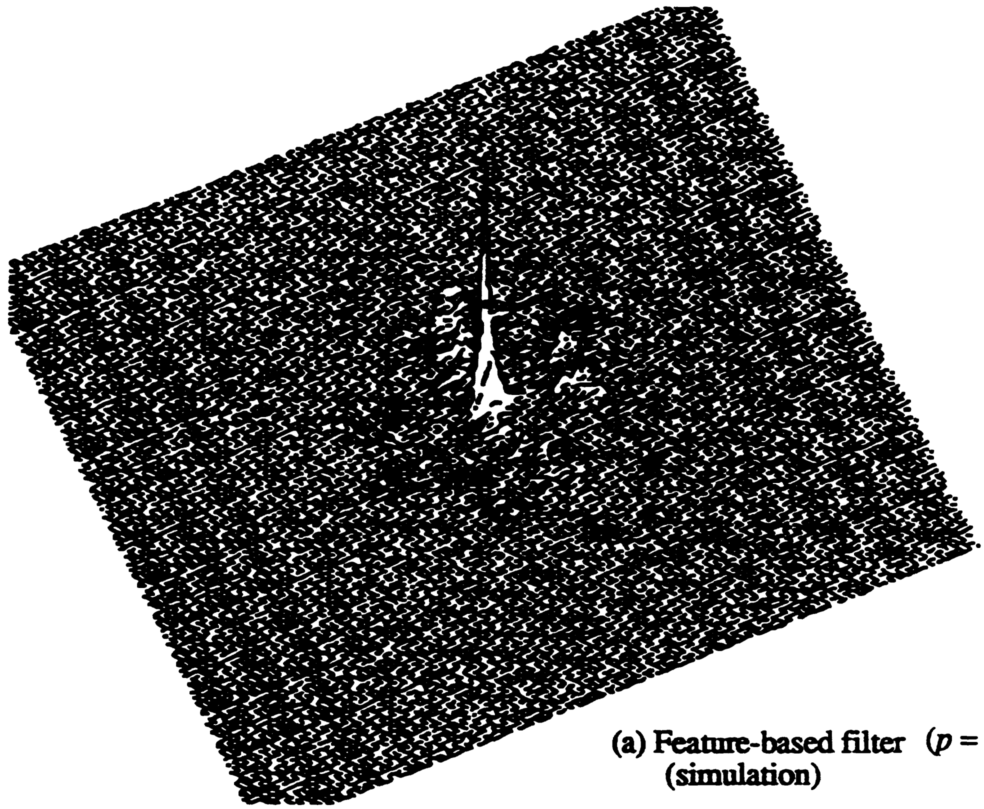
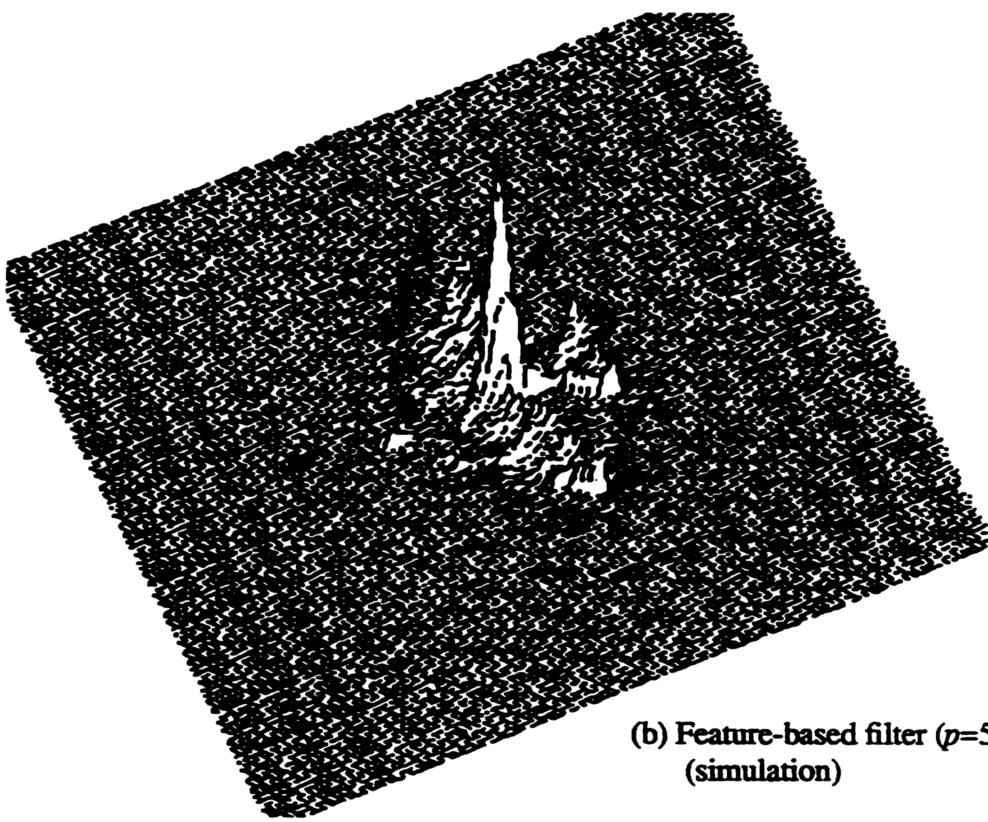


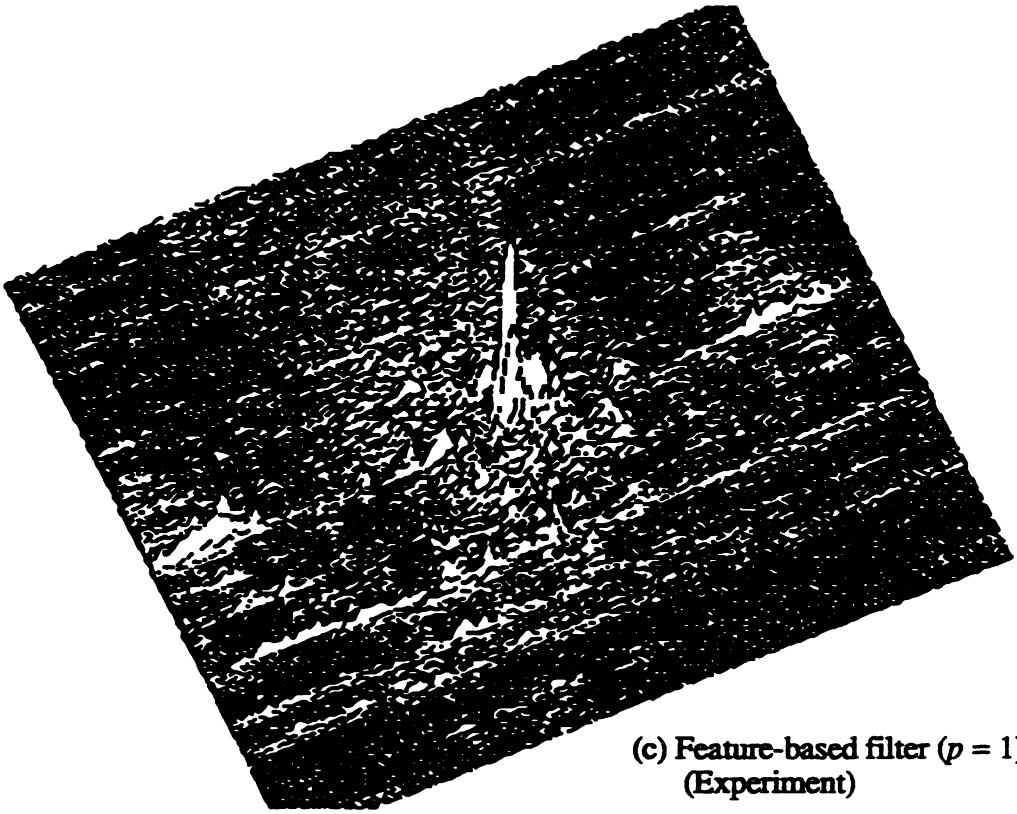
FIGURE 2. Images used in training set obtained by thresholding actual sensor data



(a) Feature-based filter ($p = 1$)
(simulation)



(b) Feature-based filter (p=5)
(simulation)



(c) Feature-based filter ($p = 1$)
(Experiment)

FIGURE 3. Examples of cross-correlation responses for an image in the training set and a feature-based filter