4-2015

A Robust Fringe-Adjusted Joint Transform Correlator for Efficient Object Detection

Paheding Sidike
University of Dayton

Vijayan K. Asari
University of Dayton, vasari1@udayton.edu

Mohammad S. Alam
University of South Alabama

A robust fringe-adjusted joint transform correlator for efficient object detection

Paheding Sidike[a], Vijayan K. Asari[a] and Mohammad S. Alam[b]
[a]Dept. of Electrical and Computer Engineering, University of Dayton, OH, USA 45469;
[b]Dept. of Electrical and Computer Engineering, University of South Alabama, AL, USA 36688

ABSTRACT

The fringe-adjusted joint transform correlation (FJTC) technique has been widely used for real-time optical pattern recognition applications. However, the classical FJTC technique suffers from target distortions due to noise, scale, rotation and illumination variations of the targets in input scenes. Several improvements of the FJTC have been proposed in the literature to accommodate these problems. Some popular techniques such as synthetic discriminant function (SDF) based FJTC was designed to alleviate the problems of scale and rotation variations of the target, whereas wavelet based FJTC has been found to yield better performance for noisy targets in the input scenes. While these techniques integrated with specific features to improve performance of the FJTC, a unified and synergistic approach to equip the FJTC with robust features is yet to be done. Thus, in this paper, a robust FJTC technique based on sequential filtering approach is proposed. The proposed method is developed in such a way that it is insensitive to rotation, scale, noise and illumination variations of the targets. Specifically, local phase (LP) features from monogenic signal is utilized to reduce the effect of background illumination thereby achieving illumination invariance. The SDF is implemented to achieve rotation and scale invariance, whereas the logarithmic fringe-adjusted filter (LFAF) is employed to reduce the noise effect. The proposed technique can be used as a real-time region-of-interest detector in wide-area surveillance for automatic object detection. The feasibility of the proposed technique has been tested on aerial imagery and has observed promising performance in detection accuracy.

Keywords: object detection, joint transform correlation, fringe-adjusted filter, correlation, local phase, synthetic discriminant function, logarithmic fringe-adjusted filter

1. INTRODUCTION

In real-time optical pattern recognition, fringe-adjusted joint transform correlation (FJTC) [1] has shown promising results compared to alternate joint transform correlation (JTC) [2] techniques. However, it has been found that the performance of the FJTC is degraded when there is distortion, due to noise, scale, rotation and illumination variations, in the input scene.

Several enhanced versions of the FJTC have been proposed in literatures, such as synthetic discriminant function (SDF) based FJTC [3, 4] which was designed to alleviate the problems of scale and rotation variations of the target. In this technique, a training step is required to obtain equal correlation peak intensity for all training images by finding the difference between the maximum and the minimum correlation peak intensities in each iteration stage. Our previous work in [5] suggests that using histogram representation and spectral FJTC can achieve rotation invariant FJTC. However, these techniques are not robust to handle noise and illumination changes.

Illumination invariant JTCs [6, 7] have been introduced to accommodate the JTC to lighting sensitiveness when a target appears in various lighting conditions. Recently, local phase based FJTC (LPFJTC) [8] and logarithmic FJTC (LFJTC) [9] have been found to yield better performance for targets under varying illumination and noise, respectively. However, these techniques fails if there is certain scale or rotation change of the reference target occurs in the input plane.

Optical Pattern Recognition XXVI, edited by David Casasent, Mohammad S. Alam,
Proc. of SPIE Vol. 9477, 947707 · © 2015 SPIE · CCC code: 0277-786X/15/$18
doi: 10.1117/12.2086848

While abovementioned techniques usually combine specific algorithm with JTC to tackle certain target distortion, a unified and synergistic approach to equip the FJTC with robust features is yet to be done. Therefore, in this paper, a new FJTC based object detection scheme is proposed. The proposed method is designed in such a way that it is insensitive to rotation, scale, noise and illumination variations of the targets. To achieve this, we utilize local phase (LP) features from monogenic signal [10] to accommodate illumination changes of the background, and the SDF is implemented to achieve rotation and scale invariance. Finally the logarithmic fringe-adjusted filter (LFAF) is employed to reduce the noise effect. The proposed technique can be used as a real-time region-of-interest detector in wide-area surveillance for automatic object detection.

The rest of the paper is organized as follows. In section 2, we describe the steps in the proposed method that employs LP, SDF, and LFJTC to achieve robust FJTC. In section 3, we test the proposed algorithm for object detection including objects under different scale, rotation, illumination and noisy conditions in the input scene. Finally, in Section 4, we conclude our findings.

2. METHODOLOGY

Objects in an input scene could be mainly distorted by four factors: rotation, scale, noise and illumination as shown in Fig. 1. This introduces difficulty for FJTC-based techniques for efficient target discrimination. Therefore, our goal is to reduce the sensitivity of the FJTC to object distortions so that improve the detection capability in terms of sharper correlation peak intensity, narrow correlation width and higher pattern discriminability.

(a) Rotation (c) Noise

(b) Scale (b) Illumination

Fig. 1. Illustration of object distortions. First images in (a)-(d) are the reference target image (e.g. letter 'M'), and the others are the distorted targets.

A block diagram of the proposed method is briefly illustrated in Fig. 2. At first, the LP feature of the reference image is extracted using monogenic signal analysis. Then, SDF is applied for the obtained local phase information. Finally, the result of SDF undergoes the LFJTC technique. The final correlation output produces high correlation peaks for a matched target and negligible correlation peaks for a mismatch. The following part of this section will introduce the mathematical formulation of the proposed algorithm.

2.1 Local Phase (LP)

The LP is a contextual feature that is computed from the monogenic signal [10], expressed by

$$\varphi(\boldsymbol{x}) = arctan\left(\frac{\sqrt{f_1^2(\boldsymbol{x})+f_2^2(\boldsymbol{x})}}{f(\boldsymbol{x})}\right), \varphi \in [0,\pi) \quad (1)$$

where $\varphi(\boldsymbol{x})$ is the local phase and $f(\boldsymbol{x})$, $f_1(\boldsymbol{x})$ and $f_2(\boldsymbol{x})$ are the components of the monogenic signal.

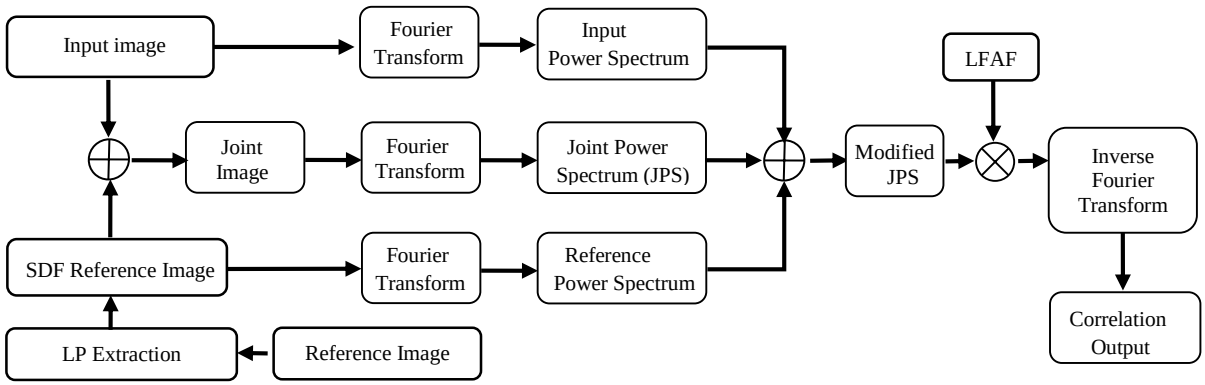

Fig. 2. Block diagram of the proposed scheme. LP: Local phase. JPS: Joint power spectrum. LFAF: Logarithmic fringe-adjusted filter.

To obtain $f(\boldsymbol{x})$, $f_1(\boldsymbol{x})$ and $f_2(\boldsymbol{x})$, first let $I(\boldsymbol{x})$ be an image, represented by

$$I(\boldsymbol{x}) = A(\boldsymbol{x})\cos(\varphi) \tag{2}$$

where $A(\boldsymbol{x})$ represents the local amplitude, and $\boldsymbol{x} = (x, y)$ is the spatial coordinates of the signal I. Then we can obtain the components of the monogenic signal representation $(f(\boldsymbol{x}), f_1(\boldsymbol{x}), f_2(\boldsymbol{x}))$ by convolving I with the transform function of even and odd pairs of spherical quadrature filters (SQFs), computed respectively as

$$f(\boldsymbol{x}) = S(\boldsymbol{x}) * g_e(\boldsymbol{x}) \tag{3}$$

$$f_1(\boldsymbol{x}) = S(\boldsymbol{x}) * g_{o1}(\boldsymbol{x}) \tag{4}$$

$$f_2(\boldsymbol{x}) = S(\boldsymbol{x}) * g_{o2}(\boldsymbol{x}) \tag{5}$$

where '$*$' represents the 2D convolution, $g_e(\boldsymbol{x})$ is the spatial domain representations of log Gabor filter, and $g_{o1}(\boldsymbol{x})$ and $g_{o2}(\boldsymbol{x})$ are the odd set of SQFs, respectively. In terms of physical interpretation, the local phase contains the structure information of the objects.

2.2 LP-based FJTC

To enable FJTC to accommodate target distortion due to illumination changes, we incorporate LP concept into the FJTC. In our previous work [8], we applied LP on both the reference image and the input image to achieve illumination invariance, however, it would computational expensive to apply LP on input images with big size, and furthermore it only can be processed online. To avoid this problem, we propose to extract LP feature only on the reference image and which can be pre-calculated and stored, thus it does not deteriorate the processing speed in real time. The mathematical formation of this proposed method is described as follows.

Assume that $\varphi_r(x, y + y_0)$ and $s(x, y - y_0)$ represents the local phase of the reference image and the unknown input scene, respectively. And they are separated by a distance $2y_0$ along the y axis. Accordingly, the input joint $f(x, y)$ is expressed as

$$f(x, y) = \varphi_r(x, y + y_0) + s(x, y - y_0). \tag{6}$$

Applying Fourier Transform to Eq. (6), yields,

$$F(u, v) = \Phi_R(u, v)\exp(jvy_0) + S(u, v)\exp(-jvy_0), \tag{7}$$

where $\Phi_R(u,v)$ and $S(u,v)$ are the Fourier transforms of $\varphi_r(x,y)$ and $s(x,y)$, respectively; u and v are mutually independent frequency domain variables. Then joint power spectrum (JPS) can be obtained by

$$|F(u,v)|^2 = |\Phi_R(u,v)|^2 + |S(u,v)|^2 + \Phi_R(u,v)S^*(u,v)\exp(j2vy_0) + \Phi_R^*(u,v)S(u,v)\exp(-j2vy_0) \quad (8)$$

where * denotes a conjugate.

To eliminate the undesired strong zero-order peak that usually present in the output plane, the Fourier plane image subtraction technique [11] is used. In this technique, both the input-only image power spectrum and the reference-only power spectrum are subtracted from the JPS. The modified JPS is thus given by

$$\begin{aligned} P(u,v) &= |F(u,v)|^2 - |\Phi_R(u,v)|^2 - |S(u,v)|^2 \\ &= \Phi_R(u,v)S^*(u,v)\exp(j2vy_0) + S(u,v)\Phi_R^*(u,v)\exp(-j2vy_0). \end{aligned} \quad (9)$$

The modified JPS of Eq. (8) is then multiplied by LP- based generalized FAF (LPGFAF) [8], given by

$$\begin{aligned} G(u,v) &= P(u,v) \times H_{lpgfaf}(u,v) \\ &= P(u,v) \times \{B(u,v)[A(u,v) + |\Phi_R(u,v)|^m]^{-1}\} \end{aligned} \quad (10)$$

where $H_{lpgfaf}(u,v)$ represents the LPGFAF, $A(u,v)$ and $B(u,v)$ are either constants or functions, and m is a constant, usually m is set to 0, 1, or 2. An inverse transform of Eq. (10) yields the correlation output.

2.3 SDF-based LPFJTC

Though successful in obtaining the illumination-invariant pattern recognition property from LPFJTC, FJTC technique suffers from scale and rotations variations of the target in the input scene. To alleviate this problem, we introduce the SDF concept on the LPFJTC. Accordingly, the LP reference image is first synthesized by using a linear combination of the distorted images from a training set, given by

$$\varphi_r(x,y) = \sum_{i=1}^{N} a_i \varphi_{r_i}(x,y) \quad (11)$$

where a_n represents the weights, and $\varphi_{r_1}, \varphi_{r_2}, \ldots, \varphi_{r_i}, \ldots, \varphi_{r_N}$ are N training images represents possible distortions of the reference image $\varphi_r(x,y)$. Consequently the joint image in Eq. (6) rewritten as

$$f(x,y) = \sum_{i=1}^{N} a_i \varphi_{r_i}(x, y+y_0) + s(x, y-y_0) \quad (12)$$

.

Accordingly, the corresponding JPS, using the Fourier Plane image subtraction technique, is described by

$$\begin{aligned} \hat{P}(u,v) &= |F(u,v)|^2 - |\Phi_R(u,v)|^2 - |S(u,v)|^2 \\ &= \sum_{i=1}^{N} a_i \Phi_{R_i}(u,v)\, S^*(u,v)\exp(j2vy_0) + \sum_{i=1}^{N} a^*{}_i\, \Phi^*{}_{R_i}(u,v)\, S(u,v)\exp(-j2vy_0) \end{aligned} \quad (13)$$

When the spatial SDF $\varphi_r(x,y)$ is used in a fractional power FJTC, the $H_{lpgfaf}(u,v)$ as in Eq. (10) is reformulated to

$$H(u,v) = \frac{B(u,v)}{A(u,v) + \left|\sum_{i=1}^{N} a_i \Phi_{R_i}(u,v)\right|^m} \quad (14)$$

Then the SDF-based fractional power FJTC is obtained by multiplying $H(u,v)$ with $\widehat{\Phi}_P(u,v)$, thus $G(u,v)$ in Eq. (10) becomes

$$\hat{G}(u,v) = \frac{\hat{P}(u,v)B(u,v)}{A(u,v) + \left|\sum_{i=1}^{N} a_i \Phi_{R_i}(u,v)\right|^m} \tag{15}$$

To obtain a constant correlation peak intensity for each training image. $\varphi_{r_i}(x,y), i = 1,2,\ldots,N$, an iterative procedure [12] based on the Newton-Raphson algorithm is used. Then the synthesis coefficients a_i $(n = 1,2,\ldots,N)$ is limited to be real and obtain the trial solutions of the coefficients by using the iterative formula:

$$a_i^j = a_i^j + \delta\,(C_{max}^j - C_i^j) \tag{16}$$

where j is the iteration number, C_{max}^j is the maximum value of the correlation peaks for each iteration, C_i^j is the correlation peak for the ith image of training set obtained in the iteration j. δ is the relaxation factor to determine the rate of changing the coefficients from one iteration to next, and $a_i^0 = 1$.

2.4 Noise reduction using LFAF

To attenuate noise such as sensor noise, we employ the LFAF during the generation of SDF composite image and final JPS processes. Consequently, $H(u,v)$ in Eq. (14) and $\hat{G}(u,v)$ in Eq. (15) are recomputed accordingly as

$$\tilde{H}(u,v) = log\ [H(u,v)] = log\left[\frac{B(u,v)}{A(u,v) + \left|\sum_{i=1}^{N} a_i \Phi_{R_i}(u,v)\right|^m}\right] \tag{17}$$

and

$$\tilde{G}(u,v) = \tilde{H}(u,v) + log\left[\hat{P}(u,v)\right] \tag{18}$$

Finally, the resultant JPS in Eq. (18) is inversed Fourier transformed to produce the correlation output in spatial domain.

3. TEST RESULTS

We consider experimental results for the FJTC, LFJTC, and the proposed robust FJTC (RFJTC) using the CLIF (Columbus Large Image Format) 2007 [13] dataset. In the experiment, we select a specific intersection in a frame with a high amount of cars for object detection. The size of the selected region of the frame is 220 × 220 pixels as shown in Fig. 3(a), and an object, manually marked by a red-color circle in Fig.3 (a), is chosen to be the target. For testing purposes, $A(u,v)$ was set to 0.001 to overcome the pole problem, $B(u,v)$ was set to unity and m is 1 for all filters used in FJTC, LFJTC, and RFJTC.

To verify the robustness of the RFJTC, twelve test images are created from the selected frame by the following three steps: 1) rotate the frame in increments of 30 degrees from 0 degree to 360 degrees and crop the intersection regions with a high amount of cars of the frame and maintain the size of 220 × 220 pixels, considering each rotation as a new image which leads to twelve images; 2) randomly choose five images among the twelve images to change their intensity values by a factor of 0.3, 0.5, 0.6 1.5 and 2, respectively; 3) randomly select two images among the twelve images to change their image size by a factor of 0.6 and 1.5. These resultant twelve test images now contains global illumination changes, object scale and rotational variations, plus sensor noise if possible. Figure 3 (b), for example, shows an image with illumination change by a factor of 0.3 and with a rotation, while Fig.3 (c) shows a rotated image at an angle of 60 degrees, and Fig.3 (d) shows a scaled image (scale factor of 0.6) with a rotation.

In RFJTC, the training set for composing the SDF reference image by rotating the reference image (size: 30 × 30 pixels), obtained from Fig. 3(a) circled region, in increments of 20 degrees from 0 degree to 360 degrees, which form the 18 training images. And then we select two sets from these 18 images to change their scales, where one set (6 images) is

scaled with a factor of 0.5 and another set (2 images) is scaled to 1.2. Lastly we linearly combine these 18 images as a composite reference image to feed into the rest of RFJTC process.

Figure 4 illustrates the detection results of FJTC, LFJTC, and RFJTC for the aforementioned twelve challenging scenarios. In Fig. 4(a), the target presents in the scene as same as the reference image, and all the three techniques detect the target without any ambiguity. From Figs. 4(b) to 4 (l), the target has been distorted in the scene due to illumination, rotation and scale changes. In Fig. 4(d), it is clear that FJTC fails to detect the target and LFJTC nearly loses the target. This is because of the changes of illumination and rotation in the scene has detrimental effect on the performance of FJTC and LFJTC, however, there is no harm to RFJTC. Figure 4(g) shows the darker scene with a rotation compared to the original image in Fig. 4(a), it can be seen that FJTC and RFJTC found the target with high discrimination ability, while LFJTC produces false alarm. Comparing the detection in the rest test images, it is evident that the proposed RFJTC is able to detect the distorted and undistorted targets while rejecting the non-target in the presence of illumination, rotation and scale variations, and the target always remain inside the red box (RFJTC detected target region) for the entire testing images, indicating the effectiveness of the proposed algorithm. Note that if the detected object is located outside of the bounding box indicates a lower accuracy or false alarm.

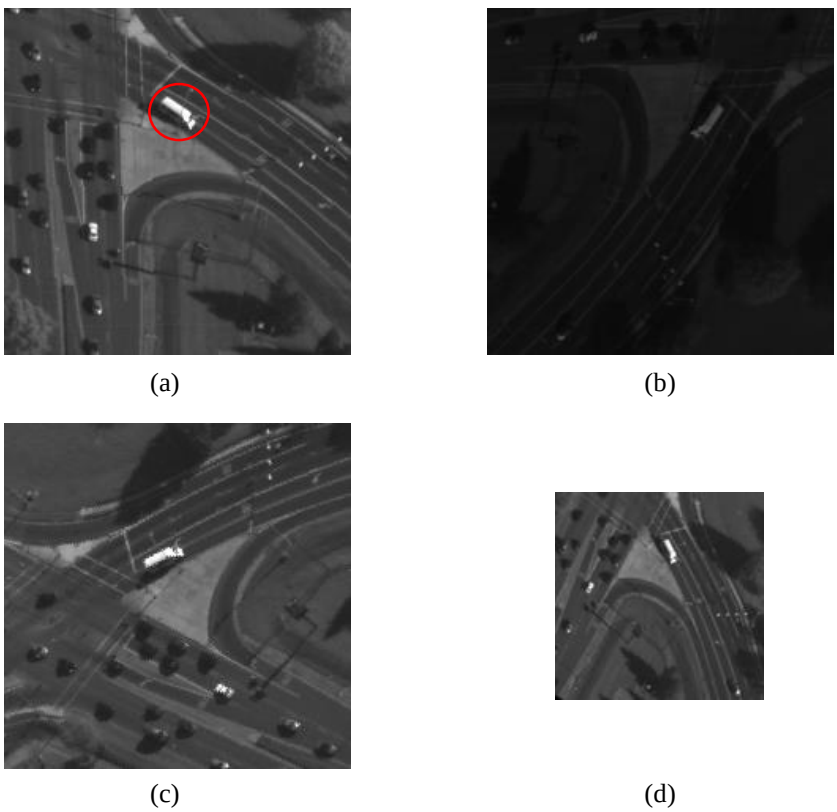

Fig. 3. Sample images used in the experiment. (a) Raw image (red-circle: target), (b) illumination and rotation changes, (c) rotation change, and (d) scale and rotation changes.

4. CONCLUSION

In this paper, we presented a new algorithm for detecting a target under varying illumination, rotation, and scaling. By utilizing LP feature from the monogenic signal and SDF-LFAF-based filter design, we achieved a robust pattern recognition technique which offers a high target discriminability. Test results show that the existing JTCs such as the FJTC and LFJTC fail to detect a target when there is certain factors of target distortion, whereas the proposed algorithm successfully detects the target without any false alarm. Future research work will focus on enhancing and implementing the proposed algorithm for object tracking in video frames with complex background.

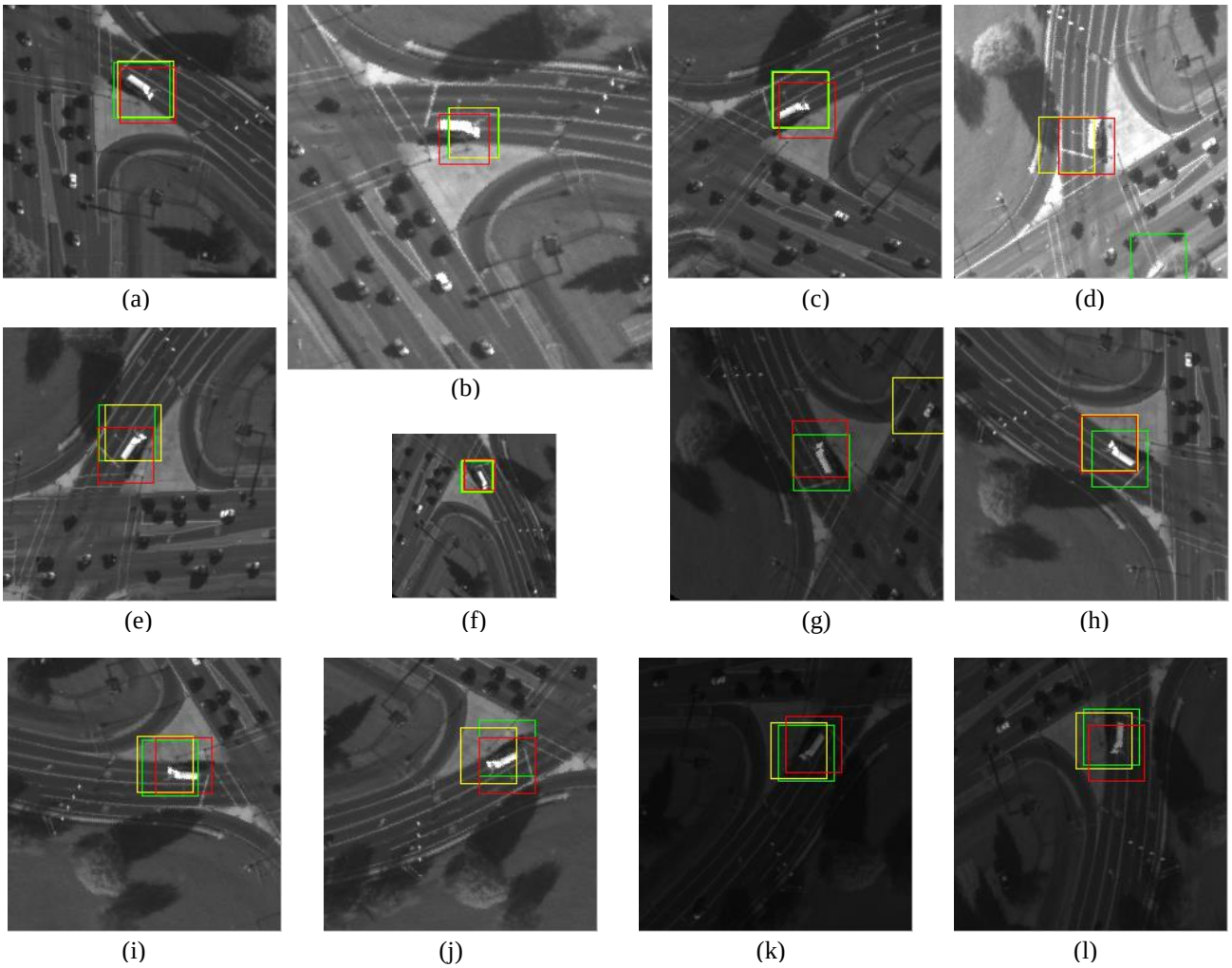

Fig. 4. Object detection results obtained using FJTC (green square), LFJTC (yellow square) and the proposed RFJTC (red square). If the detected object is located outside of the bounding box indicates a lower accuracy or false alarm.

REFERENCES

[1] Alam, M. S. and Karim, M. A., "Fringe-adjusted joint transform correlator," *Appl. Opt.* 32, 4344–4350 (1993).

[2] Weaver, C. & Goodman, J., "A Technique for Optically Convolving two Functions", *Applied Optics*, vol. 5, no. 7, pp. 1248-1249, 1966.

[3] Wang, R., Chatwin, C. R. and Shang, L., "Synthetic discriminant function fringe-adjusted joint transform correlator," *Optical Engineering*, vol. 34, pp. 2935–2944, 1995.

[4] Alam, M., "Distortion invariant Fractional Power Fringe-Adjusted Joint Transform Correlation", *Optical Engineering*, vol. 37, pp. 138-143, 1998.

[5] Sidike, P., Aspiras, T., Asari, V. K. and Alam, M. S., "A Rotation-Invariant Pattern Recognition Using Spectral Fringe-Adjusted Joint Transform Correlator and Histogram Representation", *Proc. SPIE, Optical Pattern Recognition XXV*, vol. 9094, pp. 90940F, 2014.

[6] Zhang, S. and Karim, M., "Illumination invariant pattern recognition with joint transform correlator-based morphological correlation", *Applied Optics*, vol. 38, no. 35, pp. 7228-7237, 1999.
[7] Javidi, B., Horner, J. L., Fazlollahi, A. H., Li, L., "Illumination-invariant pattern recognition with a binary nonlinear joint transform correlator using spatial frequency dependent threshold function," *Proc. SPIE* 2026, 1993.
[8] Sidike, P., Asari, V. K. and Alam, M. S., "Illumination invariant pattern recognition using fringe-adjusted joint transform correlator and monogenic signal", *Proc. SPIE, Image Processing: Machine Vision Applications Vii*, vol. 9024, pp. 90240C, 2014.
[9] Sidike, P. and Alam, M. S., "logarithmic fringe-adjusted joint transform correlation", *Optical Engineering*, vol. 52, pp.103108:1-9, 2013.
[10] Felsberg, M., Sommer, G., "The monogenic signal," *IEEE Transactions on Signal Processing* 49(12), pp. 3136–3144, 2001.
[11] Alam, M. S. and Karim, M. A., "Multiple target detection using a modified fringe-adjusted joint transform correlator," *Optical Engineering*, 33(5), pp. 1610-1617, (1994).
[12] Schils, G. F. and Sweeney, D. W., "Iterative technique for the synthesis of optical correlator filters," *J. Opt. Soc. Am.* A 3, pp. 1433–1442, 1986.
[13] AFRL, "Columbus Large Image Format Dataset," 2007. Available at: https://www.sdms.afrl.af.mil/index.php?collection=clif2007.