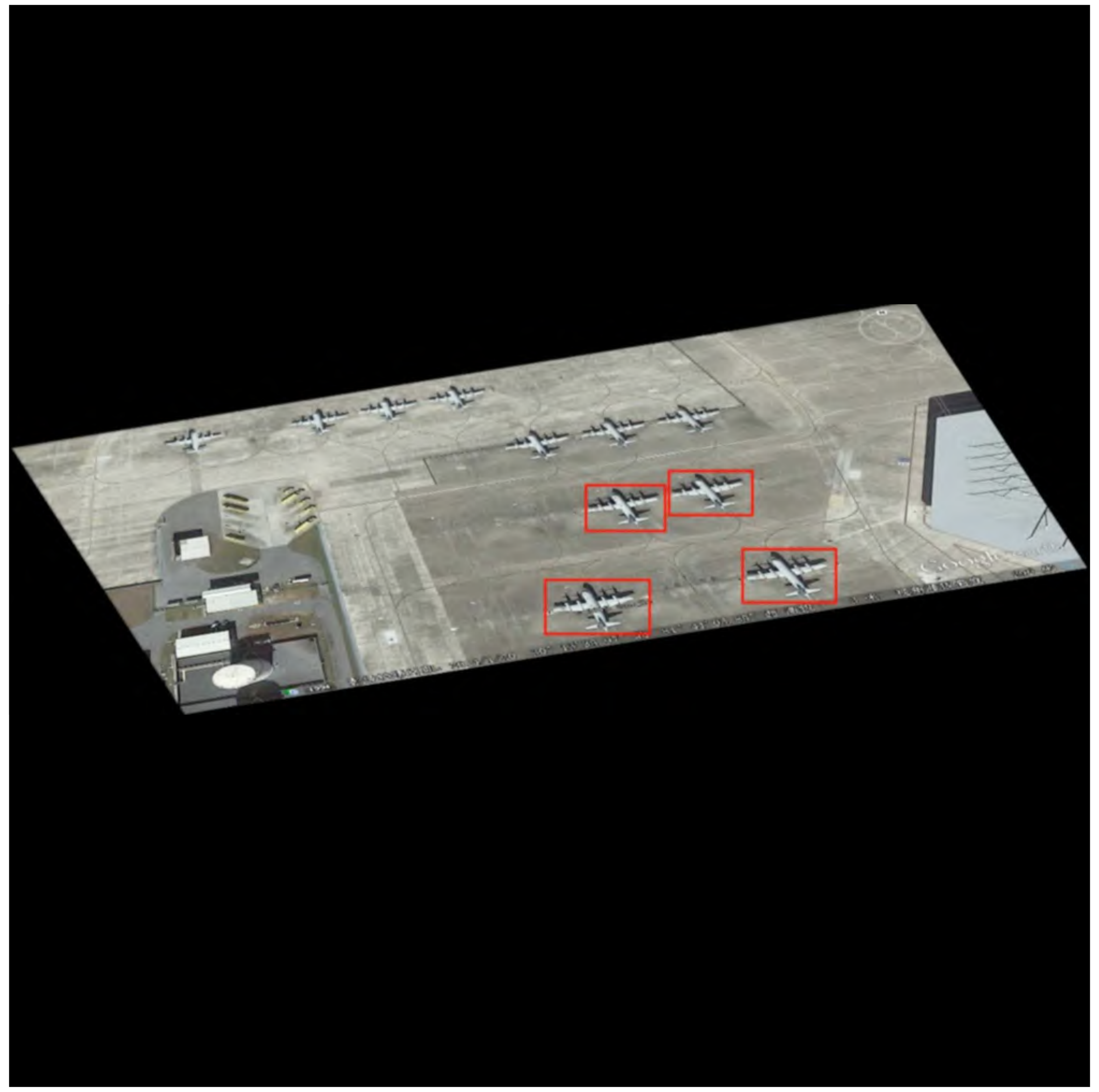
View-angle Invariant Object Monitoring Without Image Registration

Xin Zhang[1,2], Chunlei Huo[1,2], Chunhong Pan[1,2]

1. NLPR, Institute of Automation, Chinese Academy of Sciences
2. School of Artificial Intelligence, University of Chinese Academy of Sciences .

ICASSP2020 Barcelona

Introduction



Input： Multi-temporal remote sensing images of different times in the same area.

Ouput： The result of object monitoring.

Introduction

Limitations of the traditional change detection methods：

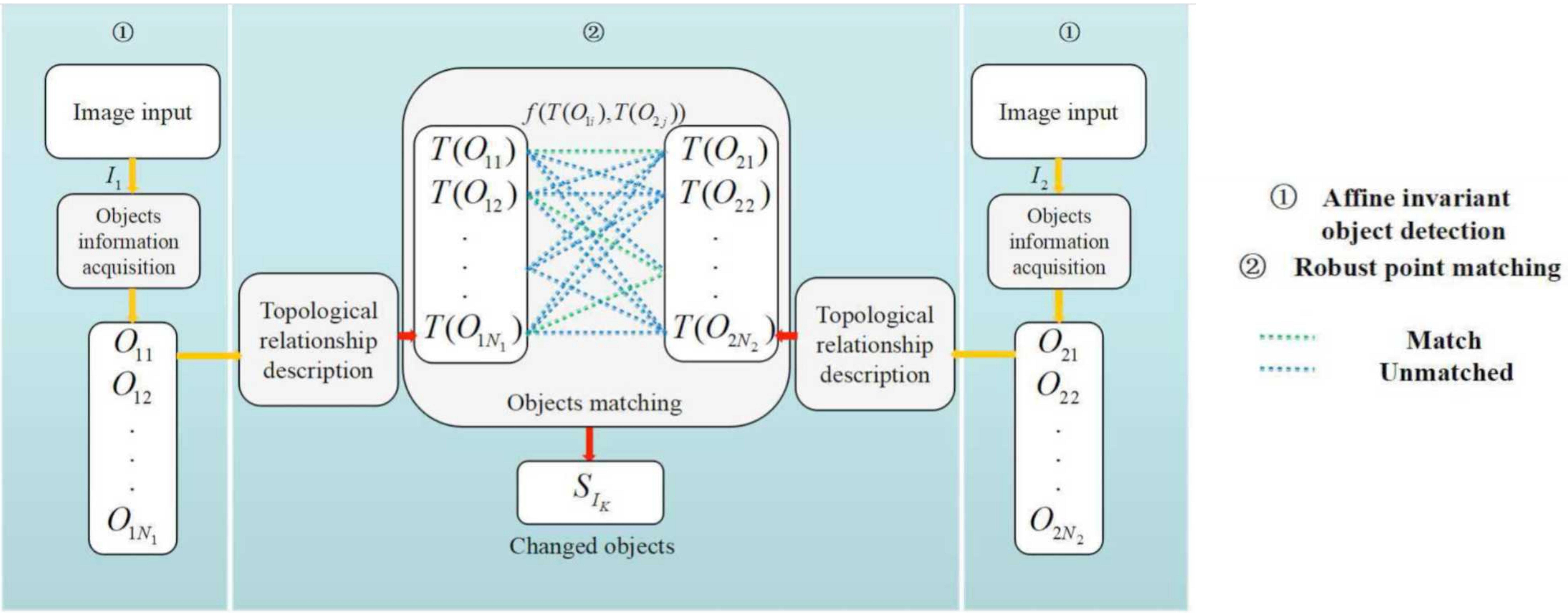
$$I_{change}(p_i) = S(\tilde{I}_1(p_i), \tilde{I}_2(p_i)) \in \{0, 1\}$$

- The traditional change detection performance depends highly on the accuracy of **image registration**.
- The results obtained by the current change detection algorithm are pixel-level changes, it is lack of rich **semantic information** of the changed objects.

The Approach

1. The framework of object monitoring

Decoupling the object-level change detection task into two sub-tasks: **objects information acquisition** and **matching**.



The Approach

1. The framework of object monitoring

Different from traditional change detection methods, we innovatively decouple the object-level change detection task into two sub-tasks: **objects information acquisition** and **matching**.

$$C_{I_1}(i) = \sum_{j=1}^{N_2} f(T(O_{1i}), T(O_{2j})) \in \{0,1\}, i = 1,2,\ldots,N_1$$

$$C_{I_2}(j) = \sum_{i=1}^{N_1} f(T(O_{1i}), T(O_{2j})) \in \{0,1\}, j = 1,2,\ldots,N_2$$

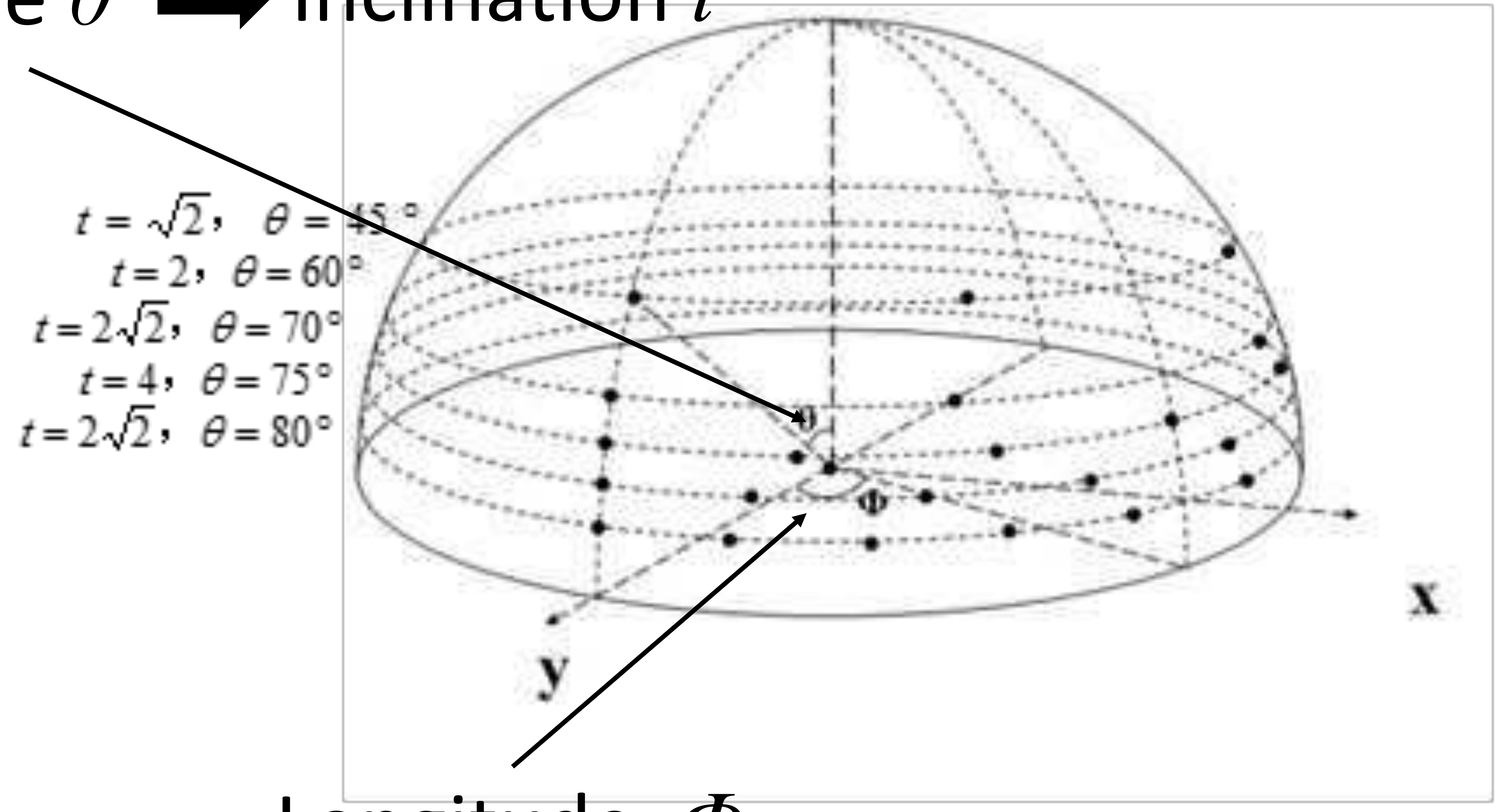
$$S_{I_K} = \{O_{Kl} \mid C_{I_K}(l) = 0, l = 1,2,\ldots,N_K\}, K = 1,2$$

The Approach

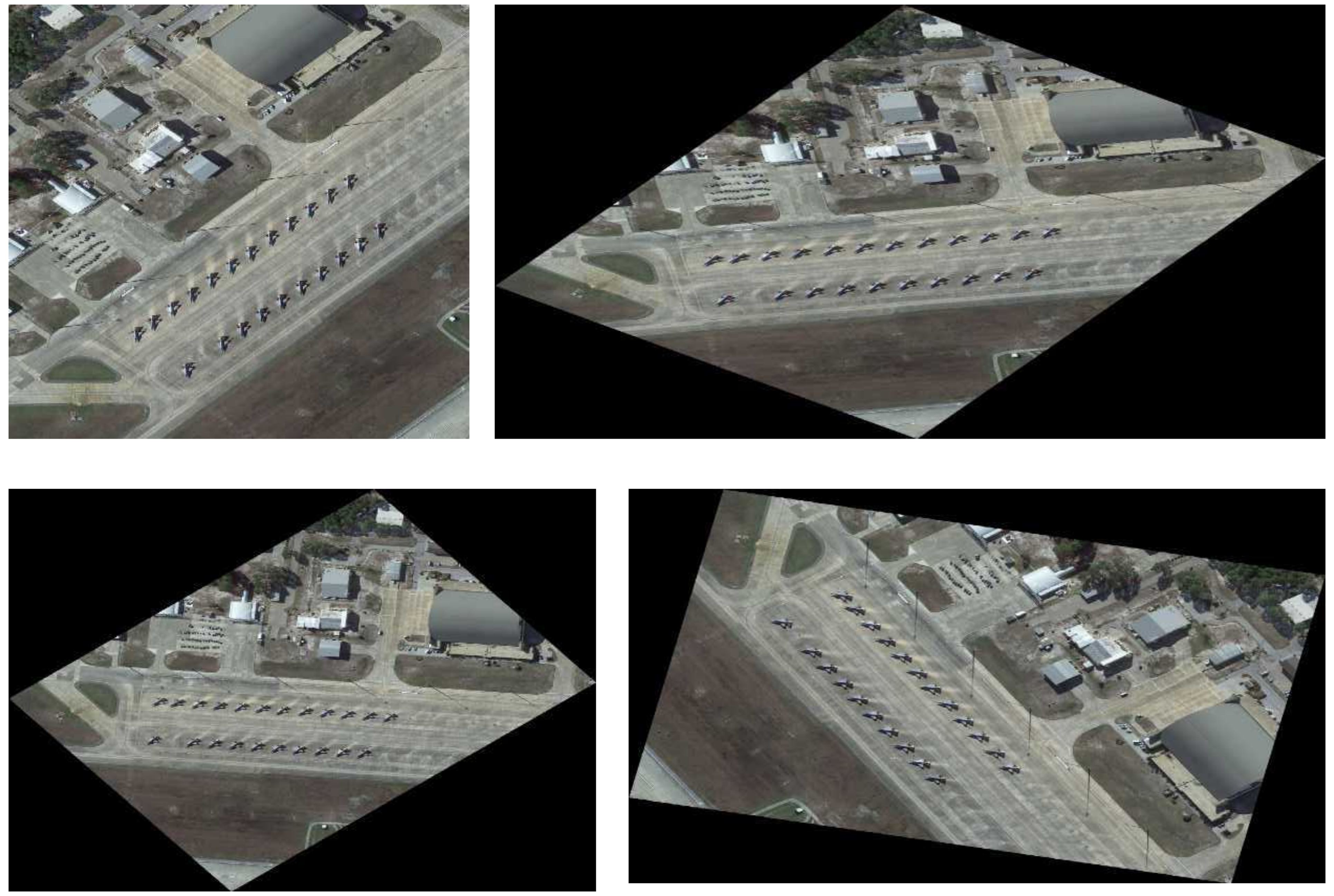
2. Affine Invariant Object Detection

- Multi-view images are generated by simulating the position of the camera, including **longitude** and **latitude**.
- Each training image is rotated by the longitude parameter Φ and tilted by the inclination parameter t.

Latitude θ ➡ Inclination t

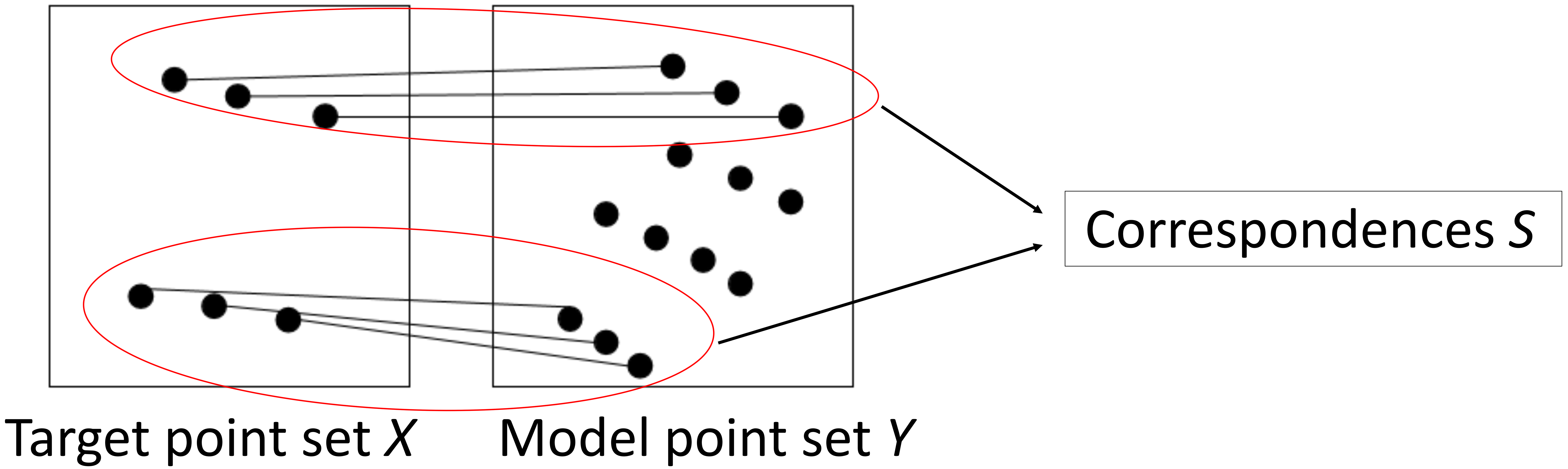


Longitude Φ



The Approach

3. Robust Point Matching



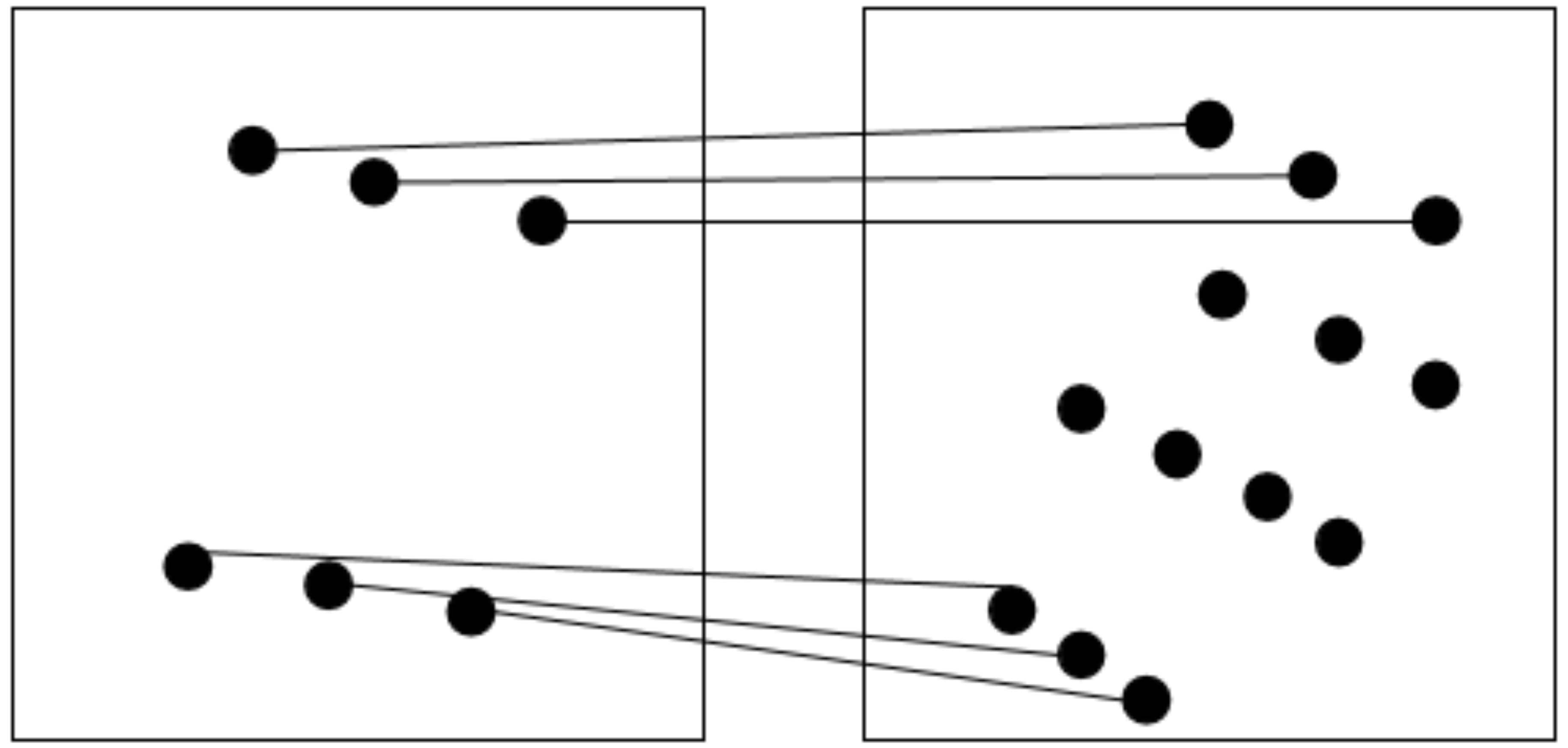
$$X = \{x_1, x_2, \ldots, x_N\}$$
$$Y = \{y_1, y_2, \ldots, y_M\}$$

$$\Rightarrow \quad S = \{(x_{n_1}, y_{n_1}), (x_{n_2}, y_{n_2}), \ldots, (x_{n_K}, y_{n_K})\}$$

The Approach

3. Robust Point Matching

- MR-RPM[1] consider **shape context** [2] algorithm as topological relationship descriptor. By calculating the neighborhood distribution of each point, the point can be encoded. If there is a group of point pair in the two point sets whose context encoding distance is very close, we can assume that this two points have a high probability of matching.
- Since objects contained in the images may be fewer and the topological relationship between **small point sets** is **weak**!

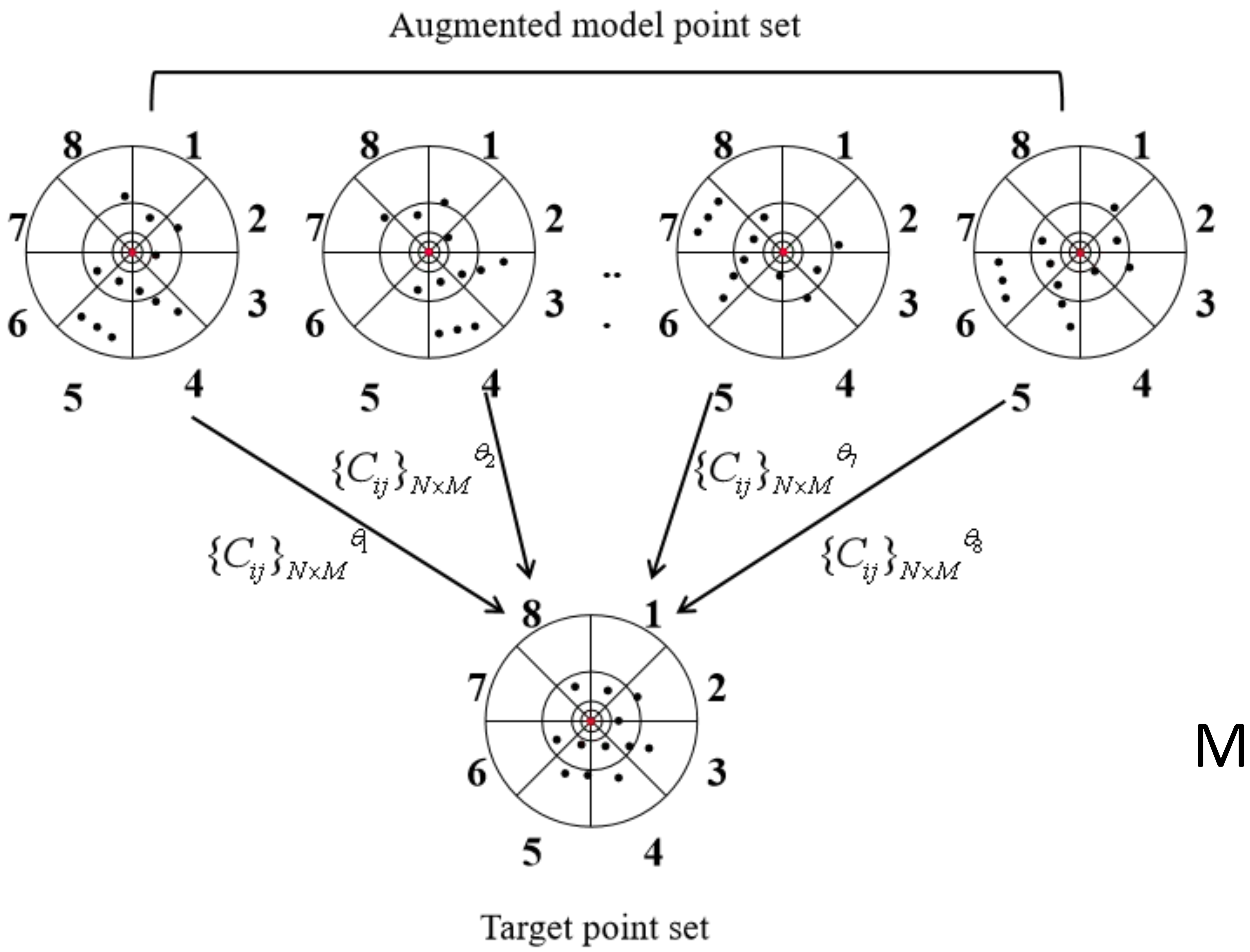


[1] Jiayi Ma, Jia Wu, Ji Zhao, Junjun Jiang, Huabing Zhou, and Quan Z Sheng, "Nonrigid point set registration with robust transformation learning under manifold regularization," *IEEE Transactions on Neural Networks*, pp. 1–14, 2018.
[2] Serge Belongie, Jitendra Malik, and J Puzicha, "Shape context: A new descriptor for shape matching and object recognition," pp. 831–837, 2000.

The Approach

3. Robust Point Matching

- An improved shape context algorithm is presented to deal with **small point sets matching**.



$$\theta_{opt} = \arg\min_{\theta}\{\overline{\{C_{ij}\}_{N\times M}{}^{\theta_l}}, l = 1, \ldots, L\}$$

$$C_{ij}^{\theta_l} = \frac{1}{2}\sum_{k=1}^{K}\frac{[h_i(k)^{\theta_l} - h_j(k)]^2}{h_i(k)^{\theta_l} + h_j(k)}$$

MR-RPM- SPS

Experiments

1. Experiment on Object detection

- **700** high resolution remote sensing images containing planes are used to generate multi-view images. Each image generates **43** images with different view angles.
- Each of the models is trained for **602000** iterations using 1 Nvidia Titan Xp (12GB) GPU.
- **3000** test images are generated by 300 remote sensing images through the affine transformation.

Experiments

1. Experiment on Object detection

Method	AP^{50}	AP^{60}	AP^{70}
FRCNN	97.43	96.59	92.25
AiFRCNN	**99.48**	**98.69**	**92.59**

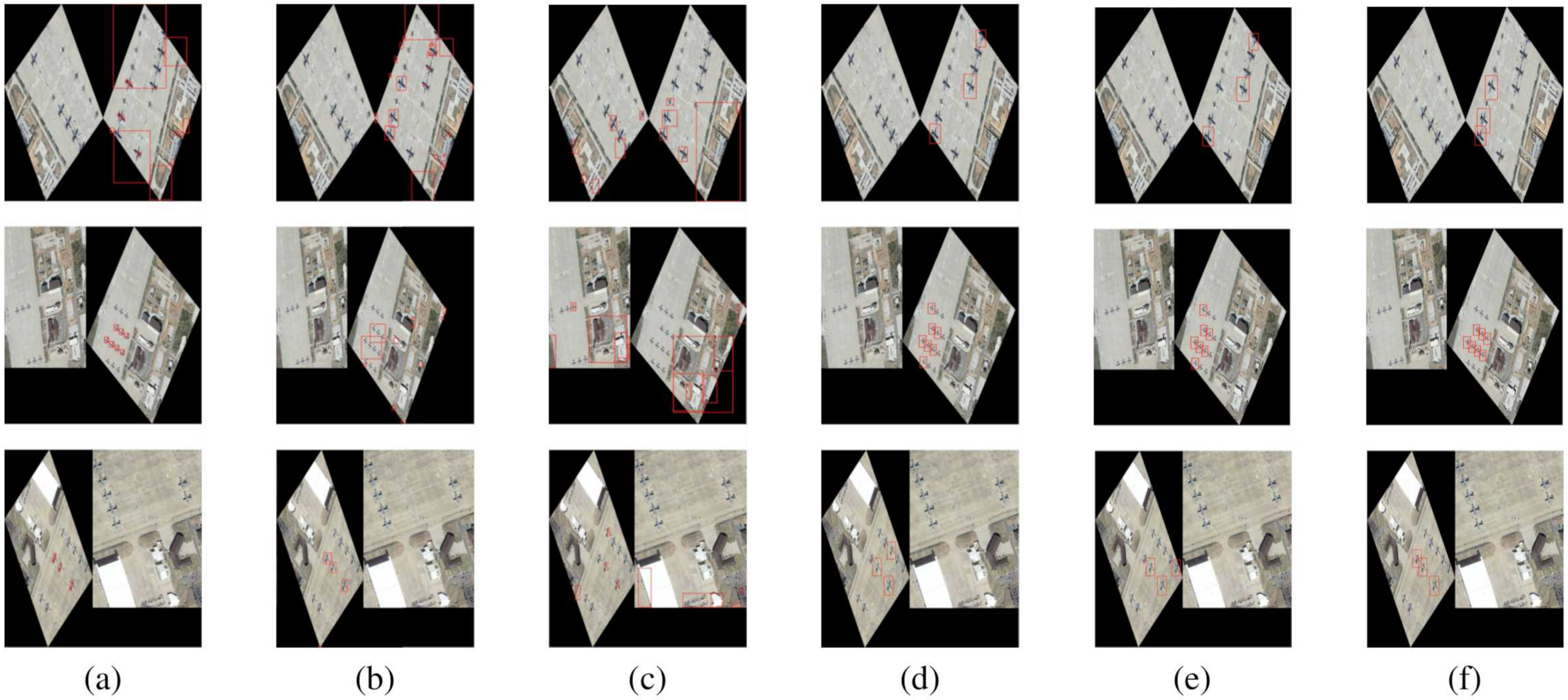
Experiments

2. Experiments on Change Detection

Approach	$FDR(\%)$	$MDR(\%)$	$PCC(\%)$	$Kappa$
Registration-based approach	26.15	39.55	69.5	0.33
DNN-based approach	37.3	23.02	66.30	0.31
MSER+ASIFT	43.87	59.2	51.27	-0.028
AiFRCNN+CPD	23.22	48.82	68.53	0.28
AiFRCNN+MR-RPM	27.71	58.27	62.44	0.14
AiFRCNN+MR-RPM-SPS	**6.74**	**14.17**	**90.86**	**0.80**

Experiments

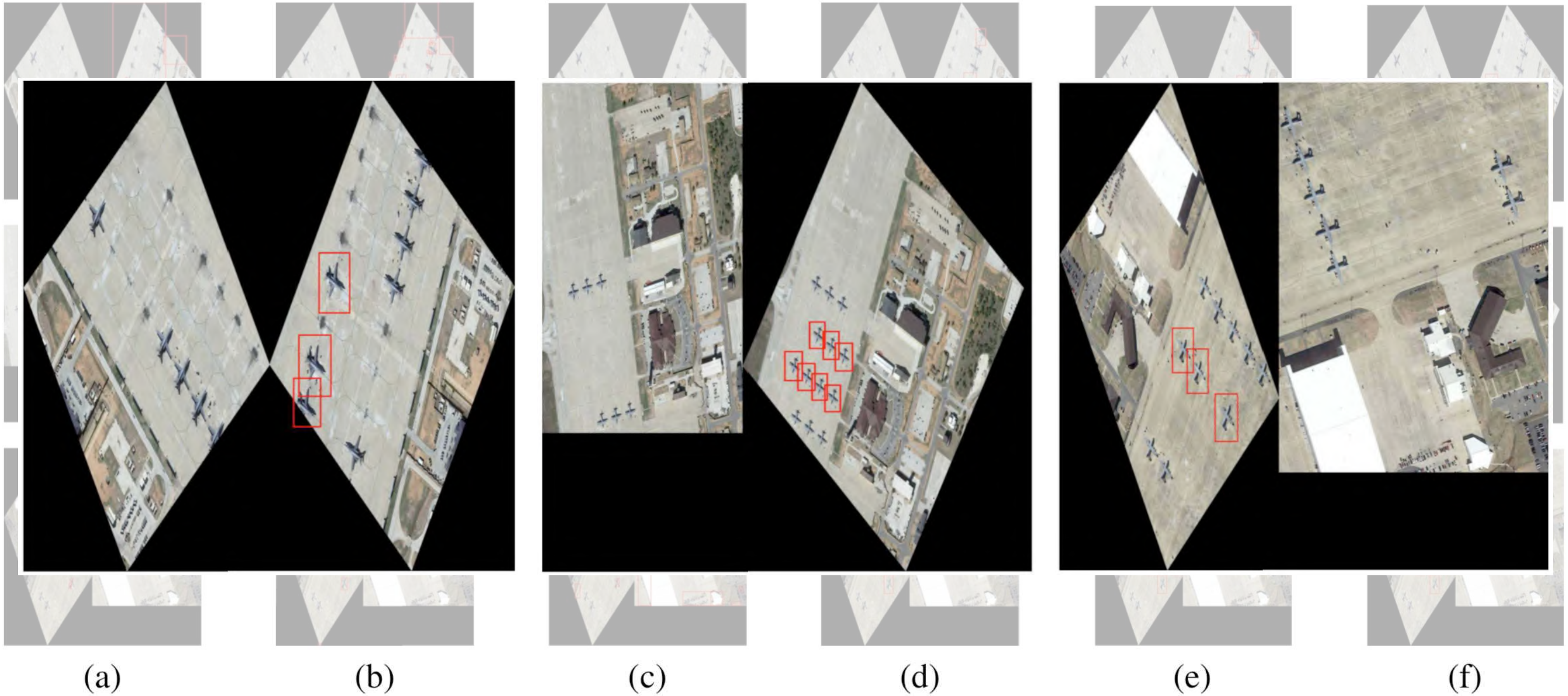
2. Experiments on Change Detection



(a) Registration-based approach, (b) DNN-based approach, (c) MSER+ASIFT, (d) AiFRCNN+CPD, (e) AiFRCNN+MR-RPM, (f) AiFRCNN+MR-RPM-SPS.

Experiments

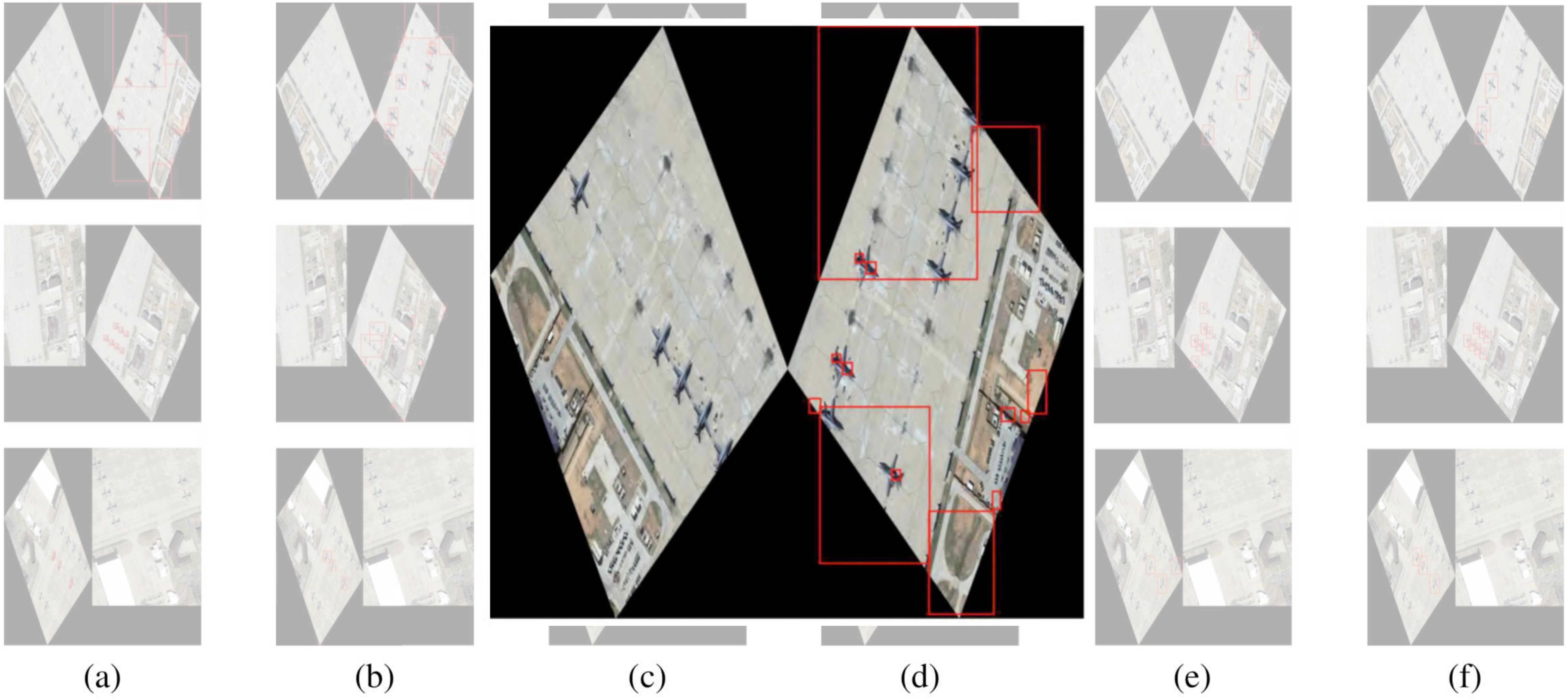
2. Experiments on Change Detection



(a) Registration-based approach, (b) DNN-based approach, (c) MSER+ASIFT, (d) AiFRCNN+CPD, (e) AiFRCNN+MR-RPM, (f) **AiFRCNN+MR-RPM-SPS**.

Experiments

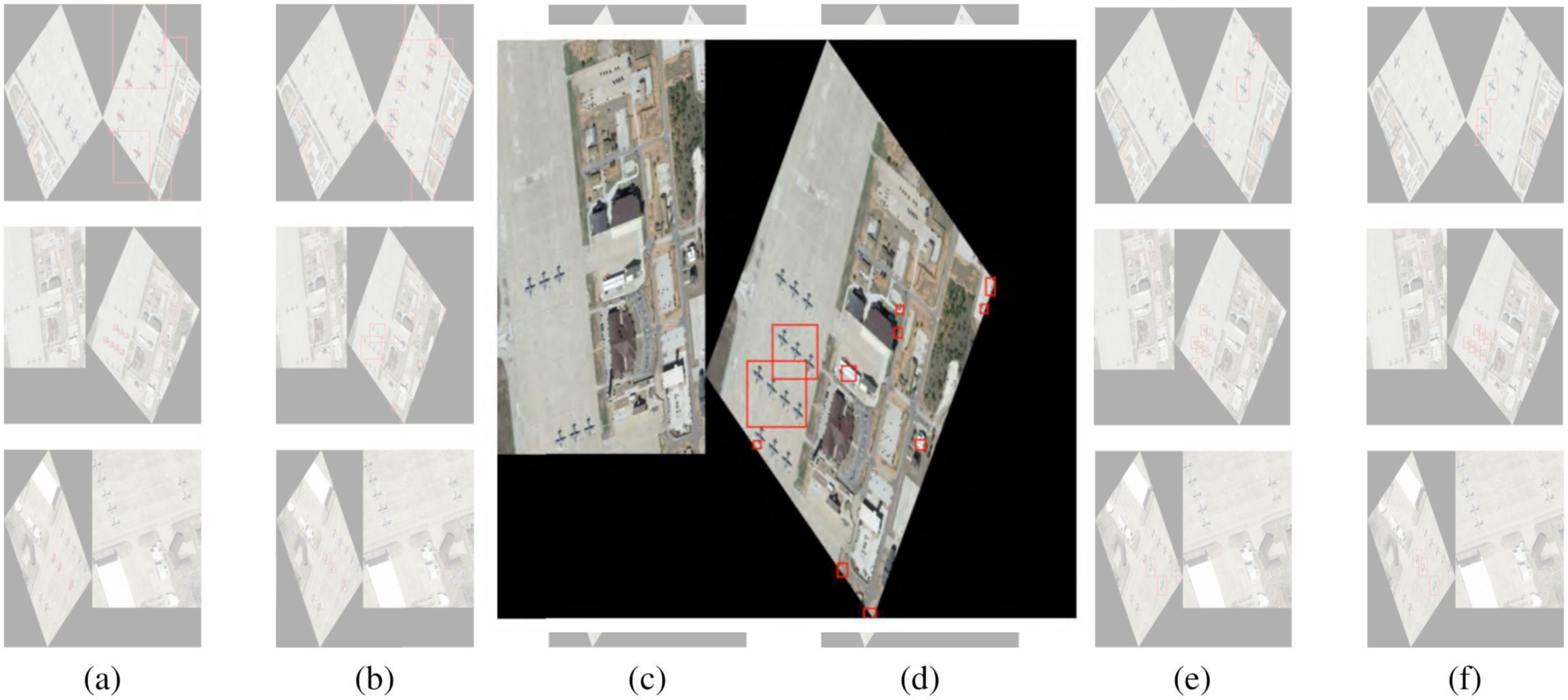
2. Experiments on Change Detection



(a) **Registration-based approach**, (b) DNN-based approach, (c) MSER+ASIFT, (d) AiFRCNN+CPD, (e) AiFRCNN+MR-RPM, (f) AiFRCNN+MR-RPM-SPS.

Experiments

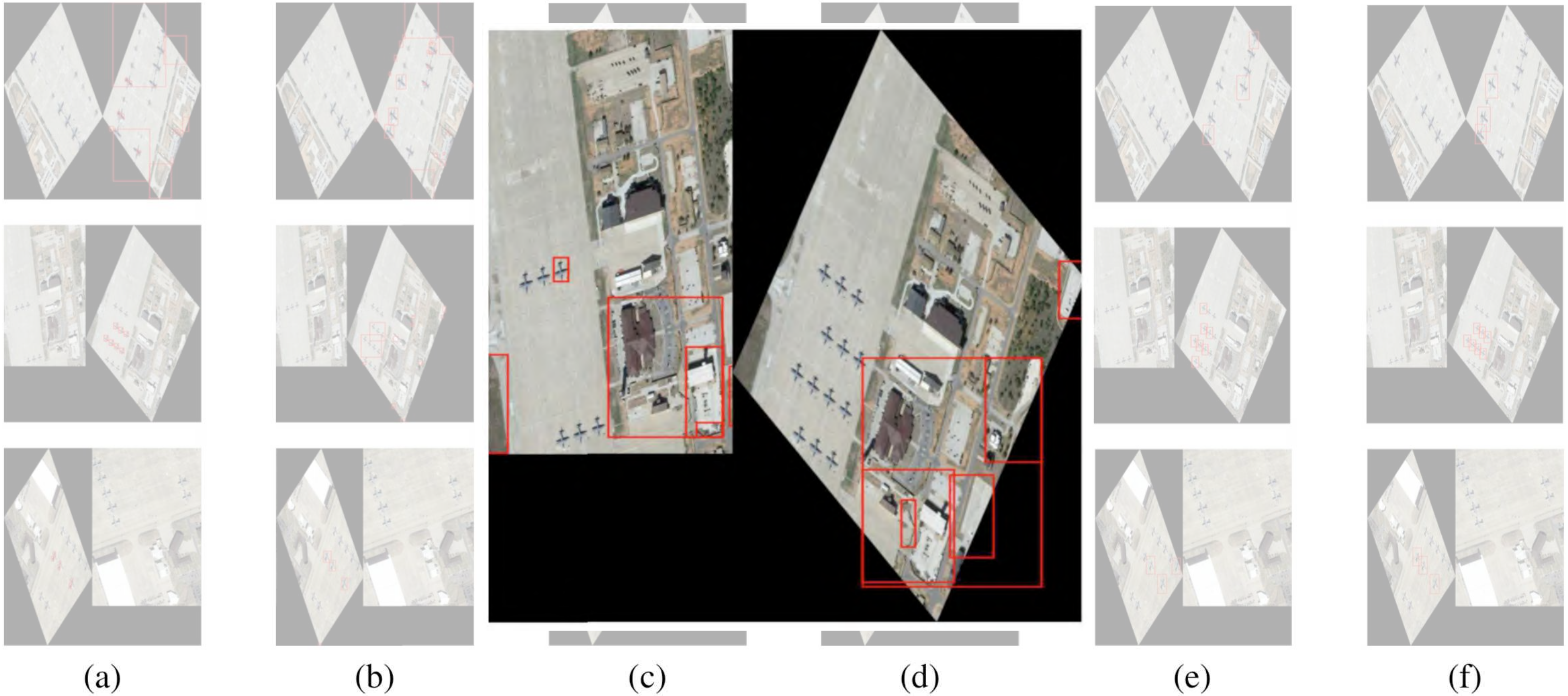
2. Experiments on Change Detection



(a) Registration-based approach, (b) **DNN-based approach**, (c) MSER+ASIFT, (d) AiFRCNN+CPD, (e) AiFRCNN+MR-RPM, (f) AiFRCNN+MR-RPM-SPS.

Experiments

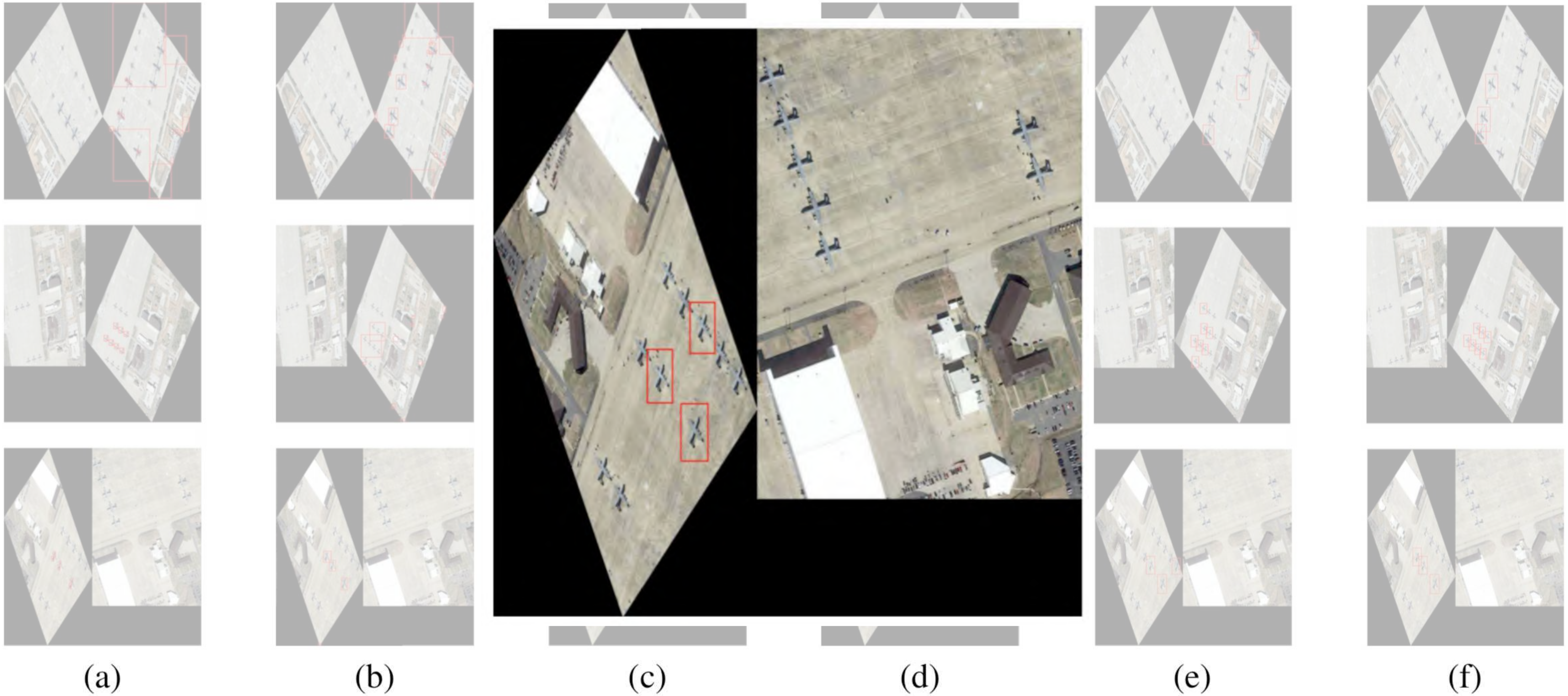
2. Experiments on Change Detection



(a) Registration-based approach, (b) DNN-based approach, (c) **MSER+ASIFT**, (d) AiFRCNN+CPD, (e) AiFRCNN+MR-RPM, (f) AiFRCNN+MR-RPM-SPS.

Experiments

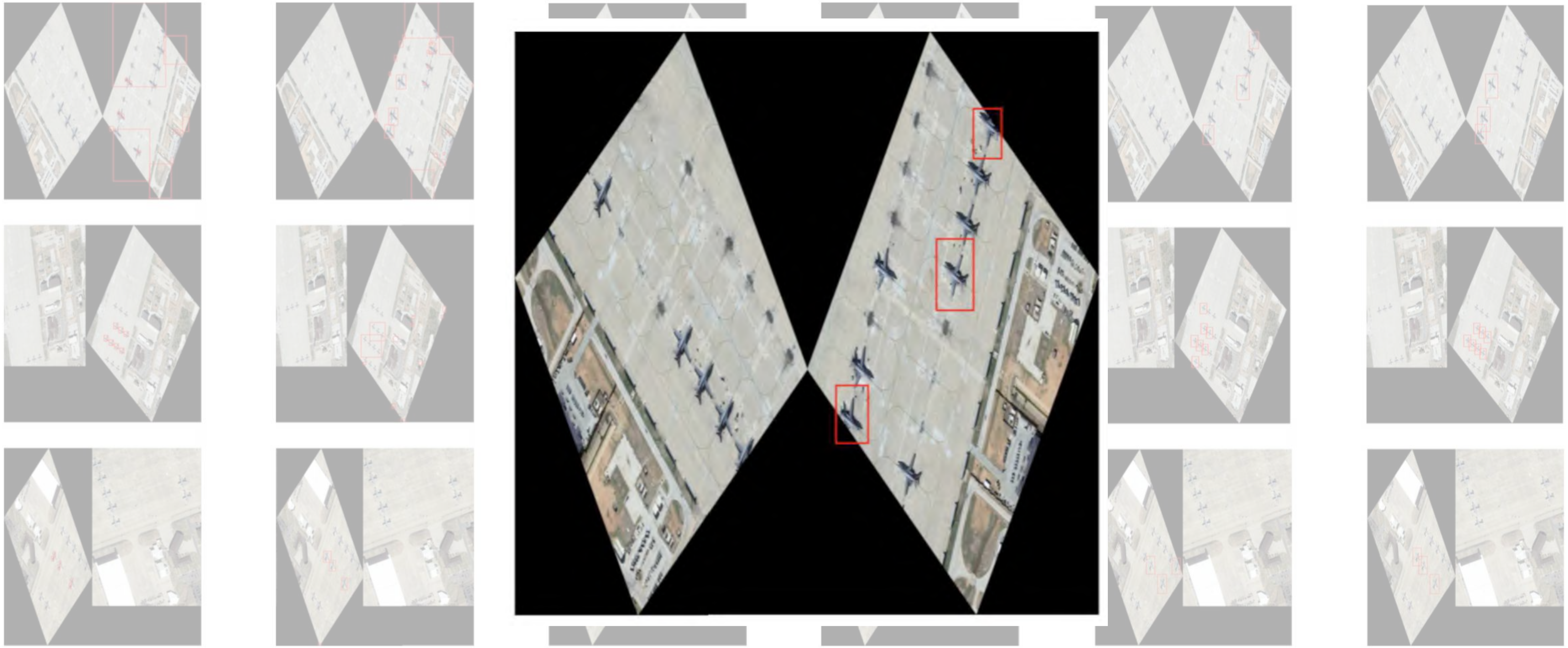
2. Experiments on Change Detection



(a) Registration-based approach, (b) DNN-based approach, (c) MSER+ASIFT, (d) **AiFRCNN+CPD**, (e) AiFRCNN+MR-RPM, (f) AiFRCNN+MR-RPM-SPS.

Experiments

2. Experiments on Change Detection

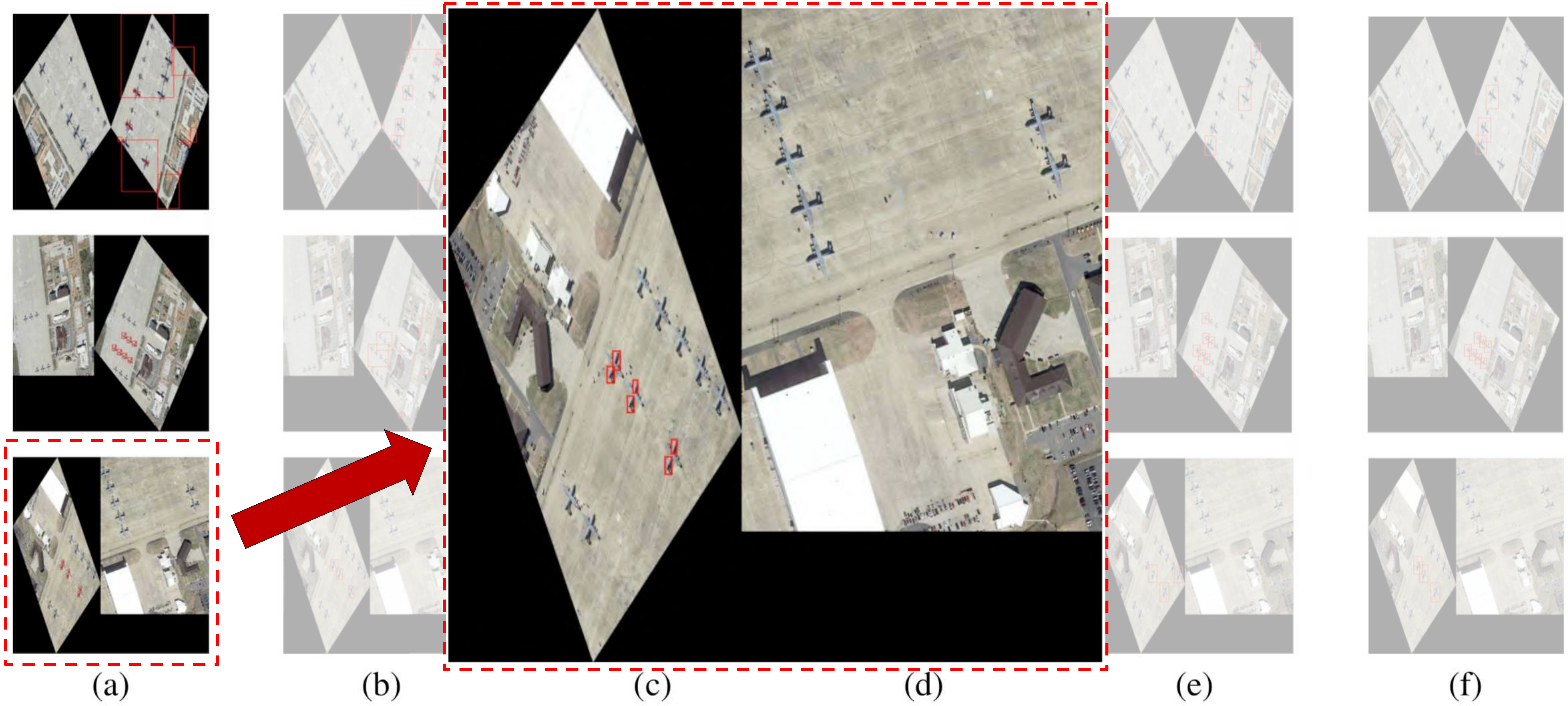


(a) (b) (c) (d) (e) (f)

(a) Registration-based approach, (b) DNN-based approach, (c) MSER+ASIFT, (d) AiFRCNN+CPD, (e) **AiFRCNN+MR-RPM**, (f) AiFRCNN+MR-RPM-SPS.

Experiments

2. Experiments on Change Detection



(a) **Registration-based approach**, (b) DNN-based approach, (c) MSER+ASIFT, (d) AiFRCNN+CPD, (e) AiFRCNN+MR-RPM, (f) AiFRCNN+MR-RPM-SPS.

Conclusion

An **novel object-specific change detection approach** is proposed for objects monitoring task, which decomposes the task into **object detection** and **point sets matching**.

- It is robust to view angle variation **without registration.**
- The result is **rich in semantic information**.

Future Work

The proposed approach will be evaluated on heterogeneous multi-temporal images.