A Review of Visual tracking Technology based on Infrared Image

Shijie Ren [a)], Xiaogang Yang [b)] and Naixin Qi [c)]

Xi'an High Technology Research Institute, Shanxi 710025, China.

[a)] 314419468@qq.com
[b)] doctoryxg@163.com
[c)]qinaixin2015@sina.com

Abstract. Visual tracking is a hot topic in the field of computer vision, this paper firstly describes the classification and application of visual tracking and highlights the advantages of object tracking based on infrared image by simple comparison. Combined with the research of this technology, tracking technology is divided into five categories, which are named region-based tracking, model-based tracking, activity-based tracking, feature-based tracking, and filtering-based tracking respectively. Finally, based on the characteristics of infrared images, this paper expounds the difficulties in the visual tracking of infrared images and prospects the research prospect of infrared image visual tracking.

Key words: Computer Vision, Visual Tracking, Infrared Image.

INTRODUCTION

With the continuous development of computer and internet technology, the computational ability of computers has made a large leap, it makes computer plays an increasingly important role in the academic research, of which visual function is one of the hot issues in the field of computer [1]. But the visual tracking problem is an important issue in the field of computer vision. The essence of this method is to detect, extract and identify the target in the video sequence, and then to track the target, extract the target position, velocity, acceleration, trajectories and other movement parameters in order to achieve a better understanding of the target behavior and complete a higher level of tasks.

There are many classifications of visual tracking. HAN Chong Zhao [2] summarize several methods of visual tracking classification are reviewed in this paper. Firstly, according to the number of cameras used, they are classified into single-camera target tracking and multi-camera tracking. Secondly, on the basis of the motion state of the camera, it is divided into static camera target tracking and moving camera target tracking; third is based on the number of tracking targets, divided into single-target tracking and multi-target tracking; Fourthly, according to the structural attributes of tracking target, divided into rigid object tracking and non-rigid object tracking. And the fifth is based on the type of sensor imaging, one is tracking based on visible image and the other one is tracking based on infrared image. Figure 1.1 shows the visual tracking problem-driven classification.
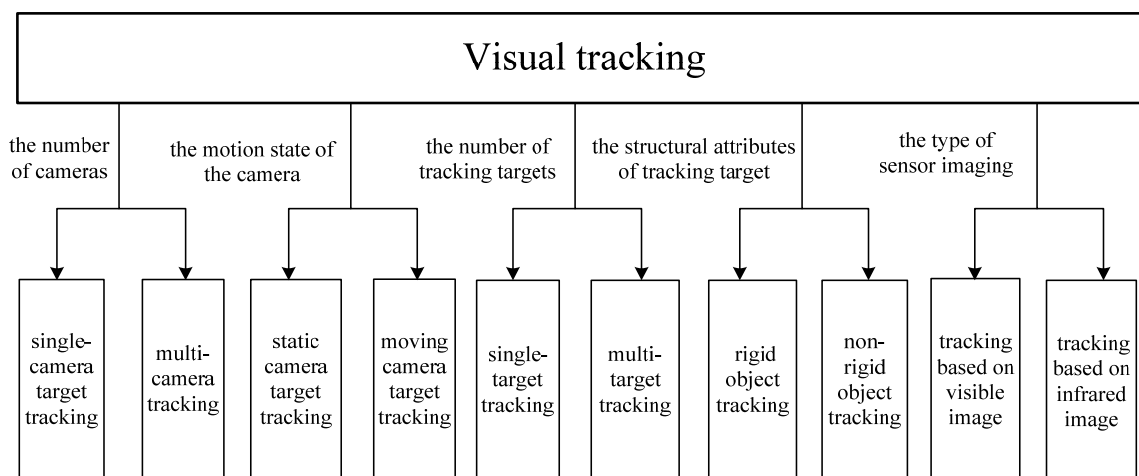
FIGURE 1. Target tracking driven by problem.

This article we mainly discuss the type of infrared sensor imaging tracking. Visible image tracking has many advantages, such as abundant spectral information, high resolution, and large dynamic range and so on. But in the night or low visibility conditions, it has poor effect and can't deal with the demand. However, the infrared image is made by using different objects or different parts of one object, due to the different temperatures produced by different infrared radiation intensity. The process shows in Figure 1.2. It has the advantages of good sensitivity, high accuracy, strong anti-interference ability, good concealment, capable of working at any time and any weather and so on, which offsets the defect of visible image tracking. Therefore, many domestic and foreign research teams have done many researches on it and have useful achievements.
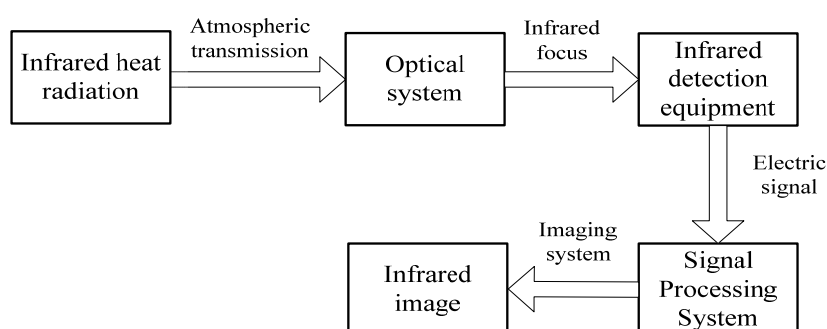

FIGURE 2. Infrared imaging flow chart.

The reason why infrared images have drawn much attention is because of they have very high value for many military and civilian applications. It is mainly used in the following three aspects: (1) Medical diagnosis: Human body has a certain temperature, so various inspections can be done by infrared thermal imager, it's common that many developed countries using infrared technology to medical diagnosis and is of very significant meaning in early diagnosis of some cerebrals, tumors and other serious diseases. (2) Video monitoring: It is mainly used in the monitoring of public places such as banks, communities, parking and other public places. It also can servile suspicious persons who carrying dangerous goods, in order to prevent theft, vandalism, and force to maintain public order. With the continuous development of technology, visual tracking has also been widely used in transportation systems. In haze or other bad weather, it can control traffic flow and track the behavior of vehicles and pedestrians, reducing the probability of accidents. (3) Navigation, control and guide: In the battlefield conditions, it is necessary to detect and track combat targets such as aircraft and warships. Target tracking occupies more time proportion in the terminal guidance of the missile. Therefore, the superiority of the tracking algorithm often determines the

accuracy of missile strike. Especially in the night and other harsh environments using infrared image tracking can greatly reduce the target detection time and tracking time and improve attack precision of the missile.

CLASSIFICATION OF TRACKING METHODS

From the point of the tracking method and technology, the target tracking methods can be roughly divided into the following five categories: region-based tracking, model-based tracking, active contour-based tracking, feature-based tracking and filtering-based tracking.

Region-Based Tracking

The region-based tracking algorithm is on the basis of the similarity measure of the target template and the real-time graph in the sequence image, according to the set criteria detects the target and its location, so as to achieve accurate tracking. The template matching method is the most typical method based on area tracking, which has a long history in the field of target tracking. In general, this method can build the target template with fewer initial graphs, then use the sliding window to search the target area in the image space to get the candidate target set and give a measure of pertinence between Candidate target and constructed target, select the candidate target set with the largest correlation coefficient as the target state. This method is also called related tracking method. Product correlation (Prod) [3] [4], differential correlation (MAD) [5] [6], and sequential similarity detection (SSAD) are the three most commonly used correlation tracking algorithms. Because the relevant tracking method does not need image segmentation and feature extraction, it can track the target well in low signal-to-noise ratio images and is suitable for guided missile guidance devices, which is suitable for infrared image target tracking. As early as the 1980s Lucas and Kaneda [7] proposed a classic tracking method based on image template matching. This method saves the target in the window area of the first frame image as a template. Then, based on the image block matching strategy, the Sum of Squared Differences (SSD) [8] [9]is used as the criterion of image matching, after Taylor expansion, an iterative gradient descent method can be derived to find the matching target region.

However, the disadvantages of relevant tracking method are obvious, the target location and strong interference have a great influence on the target tracking. Moreover, the correlation algorithm is a poor match, and the search space will increase dramatically with the increase of image size and template size, and the tracking efficiency is lower. In order to reduce the amount of calculation and improve the matching tracking speed, a lot of literatures in order to have the best matching position are used in a variety of accelerated algorithms, such as the SSDA, diamond search, three-step search, multi-resolution matching.

In addition, template matching can also be performed in feature space, the matching is made through the contour, gradient and optical flow field of the target. Comaneci and Meer [10] and other people introduced the histogram to describe the feature distribution of the target region, adopted the Bhattacharyya coefficient as the similarity matching criterion, and achieved the target tracking through the mean-shift iteration. Avian [11] proposed a moving target tracking algorithm based on support vector machine, Determine the best matching area by the maximum classification result, A recursive calculation method which is similar to the optical flow equation is used to find the local optimal matching region and avoid universal Search effectively.

The problems to be solved based on the region tracking methods are the preparation of the template and the update of the template. In recent years, the region-based tracking method is more concerned with how to deal with changes in the template, this change is caused by changes in the attitude of the moving target, if the prediction of the target attitude changes is correct, and we can achieve stable tracking. Relevant research on this issue can be seen in the literature. [12 13.14.15.16.17-20].

Model-Based Tracking

The basic idea of model-based tracking is as follows: First, the target model is established based on prior knowledge, and then a certain criterion is used to match the target and determine its position in the image so as to achieve target tracking and update the target model in real time during the tracking process. Model-Based Tracking Algorithm In tracking rigid body targets and non-rigid targets, the update strategy of the target model is also different, because the shape of the rigid target does not change, the target template only needs to be updated when

the number of frames or the target similarity has reached a certain threshold. The shape of the non-rigid target changes time by time and the target model needs to be updated in the matching tracking in real time.

In target tracking based on infrared image, most of the targets we track are rigid targets. According to the expression of the visual target, the target model can be divided into three common models: line model, 2D model and 3D model [21] [22]. Among them, the line graph model approximates all parts of the target by points or straight lines. The tracking method based on the line graph model is simple and the real-time performance of the algorithm is good. However, the tracking robustness is poor due to it ignore the target apparent information; the 3D model uses the generalized frustum, elliptic cylinder and ball to describe the target. The model-based tracking needs to construct the three-dimensional surface model and the motion model by using the prior knowledge of the target, and then calculate the target according to the current image sequence. This method can accurately estimate the three-dimensional trajectory of the target and has good robustness to the three-dimensional rotation of the target. However, the accuracy of the method depends greatly on the accuracy of the construction of target model, and it is difficult to obtain all the targets during the actual practice. At the same time, this method often requires a lot of computing time, which is not conducive to real-time tracking of the target. The 2D model uses the interconnected plane area to represent the target implementation. The visual tracking based on 2D images utilizes the apparent features of the target in the tracking, which is more accurate than the point-based visual tracking and the 2D images can be obtained directly. The storage and computational costs are significantly lower than the target tracking based on 3D model images. Therefore, the target tracking based on 2D images is on the focus of current visual tracking research.

Active Contour-Based Tracking

The outline of the object is the image feature used by the visual target tracking. Generally speaking, the gray value changes greatly from the target area to the background area, so the outline usually contains the important information of the target. The basic idea of tracking based on the active contour model is to first determine the initial contour line which is near the target boundary, and then define the internal force, external force and binding force on the contour line, explain further, the internal force to ensure smooth and continuous contours, external force to promote the movement of the outline of edge of object, the binding force makes the contour move towards the significant edge of the feature. Followed by minimizing the energy curve so that the outline of the initiative to move the true contour of the target, so as to achieve a robust target tracking. In general, the choice of the initial contour position has an important effect on the performance of the contour-based tracking method. In visual tracking, the algorithm usually takes the search result of the previous frame as the starting search position of the next frame.

In order to improve the performance of the algorithm based on active contour tracking, the active contour model based on the level set method has been proposed and has been well applied. The level set method, proposed by Ocher and Seething [23] is a novel method for solving the geometric curve evolution and was first applied to the target segmentation algorithm. Bertalmio [24] et al. realized target tracking based on level set algorithm. Gao Jing and XuJianjun [25] established the velocity function of fast level set algorithm and then used this function evolution to evolve the target boundary curve to achieve fast Infrared body contour tracking.

Because the tracking algorithm based on the active contour can directly obtain the shape of the target, the target description is more concise and effective, the computational complexity is also greatly reduced, and it also has some robustness under the interference conditions. However, the algorithm is sensitive to the initial position, In the vicinity of the moving target, it is necessary to ensure the initial position of the active contour. Due to the non-convexity of the contour, the contour curve may converge to a local extreme point or even diverge.

Feature-Based Tracking

The basic idea of feature-based tracking is as follows: Firstly, the features of the target image are extracted from the initial frame image (common features of infrared image are point feature, edge feature, corner feature and texture feature). The key of this kind of algorithm is select the feature. Generally, extraction of point feature is relatively simple. Methods which are commonly used as follow: Harris [26], SURF [27], ORB [28], SUSAN [26], SIFT [29] [30] and Hessian-Laplace. Among them, the feature points extracted by the SIFT algorithm have scale invariance, which can extract effective and stable features, and have strong matching ability even when images are taken at any angle in the same scene such as rotation, affine, perspective, which is very suitable for tracking the target of scale changes. However, the computational cost is very large and the algorithm always takes a long time, which seriously affects the execution efficiency of the tracking algorithm. For this reason, many scholars have made some

improvements to the SIFT algorithm. For example, Sinha proposed a SIFT efficient algorithm based on the GPU. In the literature, principal component analysis (PCA) is used instead of the histogram description in SIFT, and then get a variable number of descriptors. At the same time, SIFT was accelerated by integral histogram and Determination of Hessian matrix, and SURF (Speeded up Robust Features) algorithm was proposed.

Since the feature-based tracking algorithm uses the targets or some local features as related objects, the target tracking can still be achieved when the target part is occluded. By establishing a stable feature descriptor and selecting the appropriate dimension, fast and accurate tracking can be achieved. If the features of size, translation and rotation invariance are chosen, the problem of target tracking can be greatly simplified and the automatic tracking can be realized more easily.

Filtering-Based Tracking

Since the essence of target tracking is to find the target of interest in the sequence image, and the target can be described by its own state, so the target tracking problem is equivalent to the estimation of the target state. The target state in the continuous image sequence includes space position, moving speed, rotation angle, scale changes and so on.

The state of the target can be estimated through the observation of the target. At the same time, when dealing with the tracking problem, it is generally considered that the target's cognition has some priori characteristics. By combining the target's prior knowledge, the target's observation and the target's state, the Bayesian probability model can be constructed by the posterior probability Density and have an effective representation.

Typical Bayesian filtering methods include Kalan filtering and particle filtering. Kalan filter is first proposed by Kalan and Buy. It is an effective method for state estimation. By minimizing the error of system state estimation, it can obtain a set of equations of recursive estimation. Kalan filter has been introduced into the field of video tracking has quickly been widely used. However, the conventional Kalan filter algorithm cannot deal with the target tracking problem of nonlinear and non-Gaussian systems. For tracking non-linear and non-Gaussian systems, Izard pioneered the particle filter algorithm in the field of computer vision and proposed a target tracking algorithm based on particle filter. This kind of algorithm is in the framework of Bayesian estimation theory, through the sequential Monte Carlo sampling to obtain a set of weighted samples (particles) to approximate the posterior probability of distribution of target, this algorithm can effectively solve tracking problem in any nonlinear, non-Gaussian system state space model, its precision can approximate achieve the optimal estimate.

THE DIFFICULTY AND PROSPECT OF INFRARED IMAGE VISUAL TRACKING

The Difficulty of Infrared Image Visual Tracking

At present, the target tracking based on infrared technology mostly comes from related research based on visible light technology, and the special properties of infrared target images are not fully considered, so some visible light algorithms are not suitable for infrared target image sequences, the main reasons lead to this problem are:

Infrared image has a smaller amount of information and a relatively simple scene structure. Visible images have rich color and texture information, while the infrared imaging of the color, texture information is relatively scarce, which led to the infrared image feature extraction only can select less features, and the selection of features is relatively high. This condition leads to the description of the target and feature extraction more difficult. Moreover, the target that needs infrared image tracking mostly is the moving target, and the appearance of the target changes with time, such as changes in scale, rotation, non-rigid deformation, changes in perspective and changes in appearance which are caused by their own movement or even occlusion, etc. As a result, the feature loss will lead to the tracking failure, which greatly increases the tracking difficulty and reduces the tracking accuracy.

Low SNR and contrast. There are many sources of noise in infrared images, such as the molecular movement of nature, which will introduce a variety of noise, atmospheric radiation and solar radiation into the imaging system itself. When the noise is high, the SNR will be low. Instead of the object or infrared target, the surroundings, including air is also have radiation, the high temperature of objects continue to lose heat, low temperature objects are abort heat, the temperature continues to converge, resulting in objects and ambient temperature is not much difference, which leads to the contrast of infrared image is low and the level is not obvious, affecting the performance of the image, bringing a feeling of vague. The above two reasons can make the infrared device unable

to detect the existence of the target, or the target can be detected but the capturing ability is degraded, which affects the robustness of the tracking.

Inhomogeneity of gray distribution. Infrared imaging system has a wide temperature range, while the actual scene temperature is narrower than the scope of the object. So, the temperature distribution has small differences, which resulting in the gray distribution of image is narrow, and reflected in the histogram, some gray value has a higher proportion. This will also lead to inconspicuous characteristics of the image, with high requirements on the feature extraction algorithm, it is increasing the difficulty of tracking.

Prospect of Infrared Image Visual Tracking

In recent years, the development of infrared target tracking technology has made great progress. People have been developed a number of excellent tracker. However, each method still needs continuous improvement and development, and the algorithms also need to complement and learn from each other.

Improve the robustness of tracking algorithm :Robustness refers to the tracking algorithm can achieve continuous and stable tracking of the target in different environmental conditions, and the current tracking algorithm are basically based on specific application scenarios and data sets, but there are some typical infrared target tracking ,the common problem affects the robustness of the tracking algorithm, mainly because of the influence of the attitude change of the tracked target and the interference of environmental noise. Among them, the occlusion problem is the most typical one. How to effectively solve the problem of obscuring the target of tracking has yet to be put forward a better solution. For an obscured target, the external information it presents does not accurately descript the target, so the description of the target characteristic needs to be improved, carrying out efficient and dynamic characterizations of the target. At present, there are mainly two ways to deal with this type of problem: Fast-talking into account the region-based tracking algorithm, the state of the target and the scene information are combined, the scene information is analyzed and utilized, and then the target's overall tracking result is obtained. Secondly, it should analyze the occlusion state of the target, predicts the occlusion target's trajectory within a certain range by using its prior knowledge, and reduces the interference of the occlusion on the tracking algorithm, so as to ensure that the target tracking algorithm can obstruct the robustness in this complex situation.

improve the accuracy of tracking: Tracking accuracy is mainly manifested in the accuracy of target detection and tracking of environment tend to be more complex, there are a lot of noise, testing the target should reduce the noise as much as possible which is caused by the phenomenon of false alarm and leakage alarm, in real life, this two cannot be taken into account, we need to find an intermediate value for balance, so as to improve the detection probability of the real goal.

Improve the real-time of tracking algorithm: Nowadays, the length of time is an important parameter to measure whether the algorithm is good or not, in the tracking area is no exception, a practical tracking system must be able to track the target in real time, which requires us to improve the ability to track goals in real time. In the future, we mainly solve the contradiction between algorithmic complexities and tracking accuracy, that is, we must reduce the time for the tracking algorithm to process the objects (huge amount of image data and video sequences) and ensure the accuracy of the algorithm. Making algorithms fast and accurate is also a trend in today's tracking algorithms.

REFERENCES

1. Lu Wei, Yuan Guanglin, Xue Mogen, and Li Congli, "Visual Tracking via -Norm Maximization Principal Component Analysis," Journal of Computer-Aided Design & Computer Graphics, J., Vole. 25, no. 9, pp. 1392－1398, Sep. 2013.
2. Hou ZhiQiang, Han Chong Zhao. "A Survey of Visual Tracking," Act Automatic Sinica, J., Vole. 6, no.4, pp. 603－617, Jun. 2006.
3. Yu, M. L., et al. "Paritaprevir/ritonavir/ombitasvir plus dasabuvir with ribavirin for treatment of recurrent chronic hepatitis C genotype 1 infection after liver transplantation: Real-world experience." Journal of the Formosan Medical Association (2017).
4. Noreen, Shaikh, P. B. Gates, and J. P. Brockets. "The Metis home protein regulates the axolotl Prod 1 promoter during limb regeneration." Gene 484.1-2(2011):69.
5. MLARudolf, Markus, H. J. Welter, and H. Zimmermann. "A linear model for tracking error minimization." Journal of Banking & Finance 23.1(1999):85-103.

6. Declare, J. M., et al. "Evaluation of Left Ventricular Systolic Function Using Automated Angle-Independent Motion Tracking of Mitral Annular Displacement." Journal of the American Society of Echocardiography 18.12(2005):1266-1269.

7. Comaneci D, Ramesh V, Meer P. Kernel-Based Object Tracking [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2003, 25(5): 564-577.

8. Cobras, Dana, and P. Sturm. "3D SSD Tracking with Estimated 3D Planes." Computer and Robot Vision, 2005. Proceedings. The, Canadian Conference on IEEE, 2005:129-134.

9. Han, Youngwood, and H. Hahn. "Visual tracking of a moving target using active contour based SSD algorithm." Robotics & Autonomous Systems 53.3(2005):265-281.

10. Divan S. Support Vector Tracking [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2004, 26(8): 1064-1072.

11. Matthews, Iain, T. Ishikawa, and S. Baker. "The Template Update Problem." IEEE Transactions on Pattern Analysis & Machine Intelligence 26.6(2004):810-815.

12. Comanniciu D, Ramesh V, Meer P. Kernel-based object tracking. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2003, 25(5): 564-577.

13. Hager G, Belhumeur P. Efficient region tracking with parametric models of geometry and illumination. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1998, 20(10): 1025-1039.

14. Zhou S K, Chellappa R, Moghaddan B. Visual tracking and recognition using appearance-adaptive models in particle filters. IEEE Transactions on Image Processing, 2004, 13(11): 1491-1506.

15. Jepson A, Flit D, EI-Marathi T. Robust online appearance models for visual tracking. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2003, 2b (10): 1296-1311.

16. Matthew I, Wahiawa T, Baker S. The template update problem. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2004, 28(6): 810-815.

17. Nguyen H, Smolders A. Fast occluded object tracking by a robust appearance filter. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2004, 28(8): 1099-1104.

18. Avian S. Support vector tracking. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2004, 26(8): 1064-1072.

19. Black M, Jepson A. Eigen tracking: Robust matching and tracking of articulated objects using a view-based representation. In: Proceedings of European Conference of Computer Vision. Cambridge, UK: Springer Verilog, 1996, 329-342.

20. Kyriazis N, Argyrols A. Scalable 3D Tracking of Multiple Interacting Objects: IEEE Conference on Computer Vision and Pattern Recognition, 2014 [C].

21. Xiang Y, Song C, Mottaghi R, et al. Monocular Multitier Object Tracking with 3D Aspect Parts: European Conference on Computer Vision, 2014 [C].

22. Ocher S, Seething J A. Fronts Propagating with Curvature-Dependent Speed: Algorithms Based on Hamilton-Jacobi Formulations [J]. Journal of Computational Physics, 1988, 79(1): 12-49.

23. Bertalmio M, Shapiro G, and Randall G. Morphing Active Contours: a Geometric Approach to Topology-Independent Image Segmentation and Tracking[C]//IEEE 1998 International Conference on Image Processing, 1998: 318-322.

24. Gao Jing, Xu Jianjun,and Bi Dayan, "Fast contour tracking algorithm for human body in infrared images, "Journal of Indian University, J., Vole. 38, no. 5, pp. 95－100, CT. 2011.

25. Sun jiyin, Sun xiangdong, Wang Zhong, and Guo wenpu. "Front view infrared scene matching technology," M, Dec 2011.

26. Liu Yuhui, Yan Qian Qian, Wei Liu, et a1. An effective target tracking algorithm infrared images video[C]//6th International Conference on Wireless Communications Networking and Mobile Computing (Wacom),2010:847-850.

27. Yin, Hong Peng, et al. "A Robust Object Tracking Algorithm Based on Surf and Kalan Filter." Intelligent Automation & Soft Computing 19.4(2013):567-579.

28. Chai, Zheng, and T. Matsumaru. "Feature Tracking and Synchronous Scene Generation with a Single Camera." International Journal of Image Graphics & Signal Processing 8.6(2016):1-12.

29. Jin, Richen, and J. Kim. "Tracking feature extraction techniques with improved SIFT for video identification." Multimedia Tools & Applications (2015):1-10.

30. Zhang, headlong, et al. "SIFT flow for abrupt motion tracking via adaptive samples selection with sparse representation." Neurocomputing 249.C (2017):253-265.