

A Target Tracking Method Based on Dynamic Salient Features

H. C. Ke^{1*}, J. Z. Chen², H. Wang³, H. B. Sun¹ and Q. Gu⁴

¹ School of Computer Technology and Engineering, Changchun Institute of Technology, Changchun 130012, China

² Electrical Engineering College, Northeast Dianli University, Jinlin, 132012, China

³ College of Computer Science and Engineering, Changchun University of Technology, Changchun 130012, China

⁴ School of Computing Science, University of Glasgow, Glasgow, G12 8QQ, United Kingdom

Received 4 May 2015; Accepted 29 June 2015

Abstract

Given that dealing with blocking of traditional target tracking algorithm is not enough, a target tracking method fused into dynamic salient features is proposed by simulating human visual mechanism to ensure accuracy and efficiency. First, the salient features of the bottom layer image, such as color, intensity, orientation, and motion, are extracted. These features are considered feature vectors fused into target tracking algorithm. An improved target tracking algorithm is proposed because local regional histogram of target is influenced by the background pixels of the background region. Experiment results show that the proposed target tracking algorithm is more accurate than the traditional tracking algorithm in dealing with blocking, thereby meeting the needs of complex scenes.

Keywords: Target tracking, Human visual mechanism, Dynamic salient features, Feature vectors

1. Introduction

Target tracking is a challenging topic in the field of computer vision. In addition to video surveillance, it is also widely used in the fields of human-computer interaction, video compression, and medical image processing [1]. Target tracking is used in sequence images to identify the objects while exactly locating them to obtain target motion parameters, such as position, velocity, acceleration, and motion trajectory for further processing and analysis [2]. Many types of target tracking algorithms are available. The first type of target tracking algorithm is tracking technology based on motion model of the filtering, including extended Kalman filtering tracking method and particle filtering target tracking method [3].

At present, these methods are widely applied to organs of the human body in tracking and vehicle tracking fields. However, particle filters tend to produce particle degradation problem in the iterative process. Thus, many scholars have studied improved particle filter and proposed a particle filtering method based on the regular theory, a particle filtering method based on evolutionary theory, and other improved methods to overcome the natural degradation of particles and plaque spent defects; however, the effect is not particularly well [4], [5]. The second type of target tracking algorithm is correlation tracking algorithm. Multipoint correlation tracking algorithms based on pixel grayscale and region template correlation algorithm are typical. These methods must not only match the grayscale characteristics of reference template and search the region image but also consider the match result of image landscape location to

increase the robustness of the algorithm. When the tracking of sequence image is complete, these algorithms can obtain superior results, and an update and lost locking detection of the template can be achieved; however, these algorithms have larger computational complexity [6], [7]. The third type of target tracking algorithm is based on the optical flow characteristics of tracking algorithm. This method assumes that the grayscale values of the continuous two frames image pixel are the same. Then, the size and direction of the motion are calculated based on the continuous frames of image pixel, and sports field is used to distinguish background from moving targets. The disadvantage of this method is having excessively large computational complexity [8]. The distribution of optical flow field is not accurate and reliable because of noise, source, and blocking, and the calculation of the majority of optical flow methods is complex and time consuming [9].

In recent years, many scholars have conducted many works. Savani, et al. [10] presented a processing advantage associated with analytic perceptual tendencies and proposed a European Americans outperform Asians on multiple object tracking method, which improved the accuracy of global motion compensation. The judgment of the target movement direction is reliable. However, the effect is not good when the target moves faster. Hu, et al. [11] researched a single and multiple object tracking using log-euclidean riemannian subspace and block-division appearance model and proposed a fast correlation tracking algorithm for the features of image data through the new encoding method. A new algorithm of crossover and mutation was also defined. However, it was only fit for the obvious profile of target.

In the present paper, targets are located quickly through traditional algorithms based on theory of visual attention. Then, the probability of each feature value of target region is calculated based on the improved algorithm at the current frame of the image containing all pixel points of the target

* E-mail address: khch_2000@163.com

region. At the next frame of the image, which may exist in the candidate area of target, each feature value of the target region is calculated through intercomparison to track targets effectively. At target detection stage, each feature of the bottom layer image, including motion features, is fused according to theory of visual attention. The experiment results show that the proposed target tracking algorithm is accurate in dealing blocking, thereby meeting the needs of complex scenes.

2. Salient features extraction based on visually salient features

Visual attention is a psychological phenomenon, in which visual information is used for attention selectivity [12], [13]. It can allot in advance the limited processing resources of the system to a few salience visual regions [14], [15]. The calculation mode of visual attention provides a fast calculation mechanism in selecting some salient regions and in eliminating redundancy [16], [17]. Visual attention applied to target tracking can establish efficient target tracking algorithm fit for human visual characteristics, thereby contributing to the detection and tracking of moving target; thus, its use has a great significance in intelligent video surveillance system [18], [19].

According to visual attention theory, the intensity, color, and orientation features of the image are extracted, and then the center-surround difference operator is used to obtain multiscale features, which are fused into the static saliency map. Dynamic features are extracted from the current frame and previous frame to filter and obtain the multiscale differential map of the two frame images. Finally, the basic saliency map is eventually obtained according to the static saliency map and dynamic saliency map via the different scenes. According to the features of the input video image, the features of static saliency map and dynamic saliency map must be modified properly to generate the basic saliency map by meeting complex scenes. In this paper, two modified factors are defined to adjust static saliency map and dynamic saliency map. Specific models are discussed further in another paper of the author [20].

2.1 Static salient features extraction

Intensity features $I(c, s)$ and color features $C(c, s)$ can be obtained through Gaussian differential operator to extract static features. Orientation features $O(c, s, \theta)$ can be obtained through Gabor filtering. The static saliency map can be obtained by fusing into three feature maps,

$$S_1 = \alpha N(\tilde{I}) + \beta N(\tilde{C}) + \delta N(\tilde{O}) \quad (1)$$

where $N(\cdot)$ is normalized function; α , β , and δ are coefficients that meet $\alpha + \beta + \delta = 1$; and \tilde{I} , \tilde{C} , and \tilde{O} are intensity saliency map, color saliency map, and orientation saliency map, respectively.

2.2 Dynamic salient features extraction

Dynamic features extraction is performed between two adjacent frames, which are extracted from difference points of local and global motions. Two sequence images apply

filtering among different scales, and the image I^σ can be obtained.

Thus, $\sigma \in \{0, 1, 2, 3, 4\}$. In this paper, the size of the original image I^0 is defined as $w_k \times h_k$. Then, we can define other layer images through recursive form as follows,

$$I^\sigma(x, y) = \frac{1}{4} I^{\sigma-1}(2x, 2y) + \frac{1}{8} (I^{\sigma-1}(2x + \gamma, 2y) + I^{\sigma-1}(2x, 2y + \gamma)) + \frac{1}{16} I^{\sigma-1}(2x + \gamma, 2y + \gamma) \quad (2)$$

where σ is the scale of the image, and the value of γ is 1 or -1.

Moreover, the size of the image meets the following,

$$w_k^\sigma \leq \frac{w_k^{\sigma-1} + 1}{2} \quad (3)$$

$$h_k^\sigma \leq \frac{h_k^{\sigma-1} + 1}{2}$$

where the current frame image is $I_t(x, y)$, and the previous frame image is $I_{t-1}(x, y)$. Then, the filter template can be obtained via filtration. The motion features of the image $M_t(c, s)$ consist of point-by-point subtraction of different scales of sequence feature images I_t and I_{t-1} . Each motion feature map is normalized and then fused into the dynamic saliency map. Finally, the static saliency map and the dynamic saliency map are fused into the basic saliency map.

$$S = N(S_1 \times \Delta S_1) \oplus N(S_2 \times \Delta S_2) \quad (4)$$

Where ΔS_1 and ΔS_2 are modified matrix, according to different complex scenes to choose.

3. Target tracking algorithm based on visually salient features

n , which is the given feature of discrete points in the image, x_i , exists, where $i = 1, K, n$.

$$\nabla f(x) = \frac{1}{n} \sum_{i=1}^n \nabla K(x - x_i) \quad (5)$$

$f(x)$ is the gradient of probability density estimation of the kernel function $K(x - x_i)$ at point x .

The sample points along the feature space are iterative, which are repeatedly searched through density gradient direction so that each sample point finally converges to the local density limit point in the approach.

The kernel function is shown as follows,

$$\nabla K(x - x_i) = C \nabla k \left(\frac{x - x_i}{h} \right)^2 \quad (6)$$

where C is the normalization constant, and h is the radius of the kernel function.

The derivatives of kernel function are as follows,

$$K(x-x_i) = \frac{2C}{h^2} (x-x_i) k' \left(\frac{x-x_i}{h} \right)^2 \quad (7)$$

In this paper, $g(x) = -k'(x)$, then,

$$\begin{aligned} \nabla K(x-x_i) &= \frac{2C}{h^2} (x-x_i) g \left(\left\| \frac{(x-x_i)}{h} \right\|^2 \right) \\ &= \frac{2C}{nh^2} \sum_{i=1}^n g \left(\left\| \frac{(x-x_i)}{h} \right\|^2 \right) \left(\frac{\sum_{i=1}^n x_i g \left(\left\| \frac{(x-x_i)}{h} \right\|^2 \right)}{\sum_{i=1}^n g \left(\left\| \frac{(x-x_i)}{h} \right\|^2 \right)} - x \right) \end{aligned} \quad (8)$$

where,

$$S(x) = \left(\frac{\sum_{i=1}^n x_i g \left(\left\| \frac{(x-x_i)}{h} \right\|^2 \right)}{\sum_{i=1}^n g \left(\left\| \frac{(x-x_i)}{h} \right\|^2 \right)} - x \right) \quad (9)$$

$S(x)$ is consistent with probability density gradient direction, which is a convergence. Then, it is iterated to $S(x) \rightarrow 0$ or to a specific threshold value.

The sampling points in the model image are represented as $I_x = \{x_i, u_i\}_{i=1}^n$, where the two-dimensional coordinates of pixels in the image are presented, and u_i stands for eigenvectors corresponding with the pixels. Sampling points in the candidate image are represented as $I_y = \{y_j, v_j\}_{j=1}^m$, and the two-dimensional coordinates and eigenvectors v_j of sampling points are presented. The color features of the object are less vulnerable to effects of light, and strong robustness results in small amount of computing benefits. Color is regarded as feature space of a target, which has a wide range of applications. The color space of a target is quantized to m disjoint feature subspace, and $b(x_i)$ is a feature subspace belonging to the quantitative feature space at a pixel.

Then, the probability of the color is $u = 1, K, m$ in target model. However, kernel function histogram can be shown as follows,

$$q_u = C \sum_{i=1}^n k \left(\left\| \frac{(x_i-x_0)}{h} \right\|^2 \right) \delta(b(x_i)-u) \quad (10)$$

where x_0 is the vector value of the center point; x_i is the vector value of point i in window; m is the number of feature values, which can be matched to understand the number divided by the level of gray in image processing; eigenvalue u is grayscale; and δ is Kronecker delta function. Only the pixels with eigenvalues u must be contributed to the probability distribution. δ can be considered as weighted frequency of u .

In the candidate image, center coordinate location is y of kernel window, and kernel function histogram can be expressed as follows,

$$p_u(y) = C_h \sum_{i=1}^{n_h} k \left(\left\| \frac{(y-x_i)}{h} \right\|^2 \right) \delta(b(x_i)-u) \quad (11)$$

C in Eq. 10 and C_h in Eq. 11 are obviously the normalization constants that meet $\sum_{u=1}^m q_u = 1$ and $\sum_{u=1}^m p_u = 1$.

In the first frame image, the target model is initialized. In the adjacent sequence image, the location of the target model is searched around the candidate region to maximize the similarity function, which is the position of the tracking target in the current frame.

Bhattacharyya coefficient is used for similarity function between p and q ,

$$\rho(y) \equiv \rho[p(y), q] = \sum_{u=1}^m \sqrt{p_u(y)q_u} \quad (12)$$

where $\rho(y) \equiv \rho[p(y), q]$ is the similarity function between p and q , and $\rho(y)$ represents the similarity of the kernel histogram in the model image and candidate image. Matching procedure involves optimizing the procedure to determine the maximum of the similar function, which determines the maximum of $\rho[p(y), q]$. The first and second terms of Taylor series are taken in $\rho(y_0)$, and Eq. 10 is substituted into Eq. 11, thereby making $\rho(y)$ iterate to maximum, as long as the search direction of y is in accordance with the gradient direction.

Similar to those in Eq. 12, offset coordinates can be obtained as follows,

$$y_1 = \frac{\sum_{i=1}^{n_h} x_i w_i}{\sum_{i=1}^{n_h} w_i} \quad (13)$$

where w_i is the weighting coefficient.

$$w_i = \sum_{u=1}^m \sqrt{\frac{q_u}{p_u(y_0)}} \delta(b(x_i)-u) \quad (14)$$

The iteration of the algorithm is completed when y_1 is obtained. Then, let $y_0 \leftarrow y_1$ to the next shift. According to the properties of mean shift, which always moves to the local extreme value of Bhattacharyya coefficient, $y_0 \equiv y_1$ is complete when the target is located. The location of the local extreme value is the location of the target.

The improved target tracking algorithm is proposed. When a target model contains few background pixels, traditional algorithm can obtain efficient track effect because the local regional color histogram of the target is influenced by the background pixel of the background region. However, majority of the background pixels cause error in target locating. In this paper, pixel value u is defined as a feature

value, and normalization histogram is established according to the feature values of the target and background to reduce the effect of background pixels on the target model. The difference degree of target and background is also defined, and weight coefficient is introduced to weight target model. The smaller the difference degree of the target features and background features is, the smaller the weight of the target pixel is. The greater the difference degree of the target features and background features is, the smaller the weight of the target pixel is

$$D(u) = \log \frac{\max\{t'(u), \sigma\}}{\max\{b'(u), \sigma\}} \quad (15)$$

where $t'(u)$ is the value of u , which is the feature value of the target region; and $b'(u)$ is the value of u , which is the feature value of the background region. The background size is 20 pixels, which is outside the target widow. σ is constant, and the value is 10^{-4} . The greater the value of $D(u)$ is, the higher the difference degree of the target and background is. This observation shows that the feature values describe the target. The less the value of $D(u)$ is, the lower the difference degree of the target and background is. This observation shows that the feature values describe the background. The range of $D(u)$ is $[\log \sigma, -\log \sigma]$. λ_u can be obtained with normalized $D(u)$ as follows:

$$\lambda_u = \frac{D(u) + \log \sigma}{2 \log \sigma} \quad (16)$$

Equation 10 is redefined as follows:

$$q_u = C \sum_{i=1}^n k \left(\left\| \frac{(x_i - x_0)}{h} \right\|^2 \right) \delta(b(x_i) - u) \lambda_u \quad (17)$$

Equation 11 is redefined as follows:

$$p_u(y) = C_h \sum_{i=1}^{n_h} k \left(\left\| \frac{(y - x_i)}{h} \right\|^2 \right) \delta(b(x_i) - u) \lambda_{uk} \quad (18)$$

where,

$$C = \frac{1}{\sum_{i=1}^n k \left(\left\| \frac{(x_i - x_0)}{h} \right\|^2 \right) \lambda_u} \quad \text{and} \quad C_k = \frac{1}{\sum_{i=1}^n k \left(\left\| \frac{(x_i - y)}{h} \right\|^2 \right) \lambda_{uk}}$$

are the normalized constant meeting $\sum_{u=1}^m q_u = 1, \sum_{u=1}^m p_u = 1$.

Other steps are the same as those of traditional algorithm.

The following are the steps of the specific algorithm:

Step 1: The sequence image is inputted.

Step 2: The intensity features, color features, and orientation features of the frame image are extracted, and intensity saliency map \mathcal{I} , color saliency map \mathcal{C} , and orientation saliency map \mathcal{O} are generated.

Step 3: The static saliency map can be obtained by fusing three feature maps, namely, \mathcal{I} , \mathcal{C} , and \mathcal{O} .

Step 4: According to the k-1-frame, the motion features of the k-frame image are extracted by filtering. Then, the dynamic map can be obtained.

Step 5: The basic saliency map S is generated.

Step 6: In S , the focus of attention for the tracking target is extracted.

Step 7: Morphological filtering is applied on the marked target region of the original k-frame image.

Step 8: For marked region HSV transformation, the target border and center point are obtained.

Step 9: The marked region is the template of target tracking algorithm. Then, the template is initialized.

Step 10: The target at each image is tracked, and the similarity measure w , namely, the relative change value, is judged whether it is greater than threshold T . When $w > T$, go to Step 2. Otherwise, the target at the next frame is tracked until the end.

4. Experiment results

Three groups of color image sequences of the video are used as a test sample set in the experiment. First, three sets of the image visually salient features are extracted to generate static saliency map and dynamic saliency map through single target tracking and multiple target tracking. The proposed target tracking method is verified and then compared with existing target tracking algorithm.

In the first experiment, visually salient features of the video image are extracted, and parts of the result are shown in Figure 1. Fig. 1(a), 1(d), and 1(g) are a frame of the selected video image, which is an original image of the test sample set. Fig. 1(b), 1(e), and 1(h) are the corresponding static saliency maps of the image sequences. The static saliency map is generated as in Eq. 1, where $\alpha = \beta = \delta = 1/3$. Fig. 1(b) and 1(c) illustrate that the static saliency map and the dynamic saliency map are the same because the motion features of the players are obvious. However, Fig. 1(e) illustrates that except for the salient features of moving target, trees and the stationary car in scene also have salient features. Fig. 1(h) illustrates that buildings, billboards, and street marking in the scene have salient features. The current frame image and the previous frame image are filtered to extract the moving features effectively. Then, dynamic saliency maps are obtained by making a point-to-point difference for different scales of feature images. Fig. 1(f) and 1(i) illustrate that the static features in the scene of trees, cars, buildings, and street marking are effectively suppressed, and the dynamic features of the targets are effectively strengthened. Then, the moving targets are detected efficiently and accurately.

In the second experiment, the players are tracked in the video sequences. The size of the image is 256×192 pixels with 50 frame images. Similar to that in Eq. 4, the modified

matrix is $\Delta S_1 = \Delta S_2 = \frac{1}{2} E$. The parts of tracking results are shown in Figure 2. The proposed tracking algorithm tracks the targets effectively, particularly in the case of block, by moving the video image sequences to the next. Fig. 2(d) illustrates that the algorithm shows good robustness.

In the third experiment, three targets are tracked in community surveillance video, which is shown in Figure 3. The size of the video image is 512×384 pixels with 100 frame images. Similar to that in Eq. 4, the modified matrix is

$\Delta S_1 = \frac{1}{4}E$, $\Delta S_2 = \frac{3}{4}E$. The video shows that the marked target has three moves. Fig. 3(c) illustrates that in the 45th frame of the image, the woman who is walking is completely blocked by the riders. However, Fig. 3(d) shows that in the 49th frame of the image, the woman who is walking reappears. The proposed algorithm can continue to track very well.



Fig. 1. Saliency maps of sequence images.

Note: Saliency maps based on visually salient features. (a), (d), (g) Original image. (b), (e), (h) Static salient image. (c), (f), (i) Dynamic salient image.

In the fourth experiment, multiple targets are tracked in traffic surveillance video to test the tracking effect at complex scenes, as shown in Figure 4. The size of the video image is 512×384 pixels with 100 frame images. Similar to that in Eq. 4, the modified matrix is $\Delta S_1 = \frac{1}{4}E$, $\Delta S_2 = \frac{3}{4}E$. Many moving targets and one rider



Fig. 2. The effect images of tracking players.

Note: The frame of video sequence images. (a) 10th frame. (b) 22nd frame. (c) 26th frame. (d) 29th frame.



Fig. 3. The effect images of tracking targets in community surveillance.

Note: The frame of video sequence images. (a) 15th frame. (b) 40nd frame. (c) 45th frame. (d) 49th frame.

are shown in the image. Fig. 4(c) illustrates that before the 61st frame of the image, the white car running on the opposite lane is completely blocked. In the 61st frame, the white car appears on the opposite lane, as shown in Fig. 4(c). The proposed algorithm begins to track it. The yellow car appears on the right lane, which is ready to be tracked, while the furthest black car on the right lane moves away. The proposed algorithm does not track the black car. In the 90th frame of the image, the yellow car completely appears, as shown in Fig. 4(d). Then, the proposed algorithm begins to track it.

The tracking errors of the three sets of experiments are shown in Figure 5. Horizontal coordinates are the frame indexes of video sequence images. Vertical coordinates are the average error of the target centroid, and the unit is pixel. The offset pixels are the sum of x-axis offset pixels and y-axis offset pixels, which is the average of the pixel offset of multiple targets. The average error of the target centroid in player video image sequences is shown in Fig. 5(a). Compared with traditional mean-shift algorithm, the average error of the proposed algorithm does not make a great difference because the moving object in the video sequences is obvious. The average error of the target centroid in closed community surveillance video image sequences is shown in Fig. 5(b). At the 45th frame of the image, the woman is completely obscured, and the average error of the proposed algorithm is considerably smaller than the average error of the traditional mean-shift algorithm. The average error of the target centroid in traffic video image sequences is shown in Fig. 5(c). At the 61st frame of the image, the white car target appears on the opposite lane. Then, it is tracked with the proposed algorithm. Compared with the traditional mean-shift algorithm, the average error of the proposed algorithm has improved. At the 90th frame of the image, the yellow car completely appeared. Then, it is tracked with the proposed algorithm. Compared with the traditional mean-shift algorithm, the average error of the proposed algorithm has also improved.



Fig. 4. The effect images of tracking targets in traffic surveillance.

Note: The frame of video sequence images. (a) 20th frame. (b) 28th frame. (c) 61st frame. (d) 90th frame.

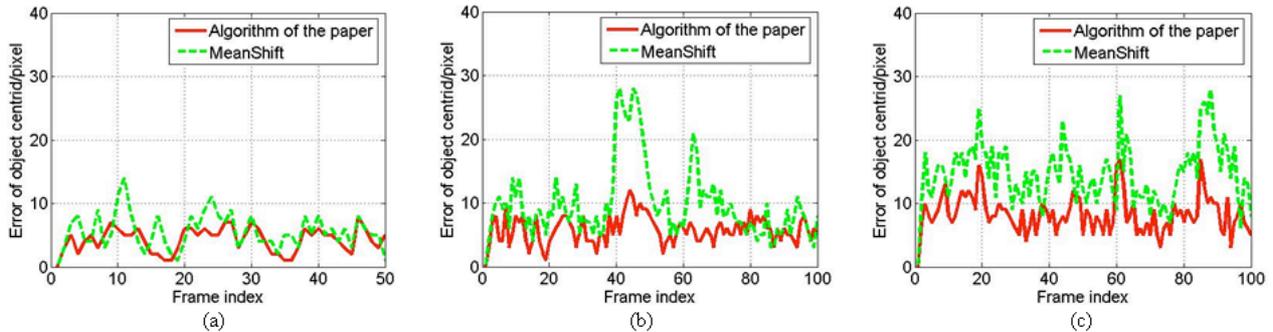


Fig. 5. Comparison of the average tracking errors in sequences for two algorithms.

Note: The different result of each group are compared. (a) Player sequences. (b) Community surveillance sequences. (c) Traffic surveillance sequences.

5. Conclusions

According to the computer model simulating the human visual mechanism and features of image sequences, a salient feature extraction method is proposed. The static salient features are extracted from the features of color, intensity, and orientation. The dynamic salient features are extracted by filtering and iterated to obtain the motion features. These features are eventually fused into the basic saliency map to determine the position of the expected target. Then, the improved target tracking algorithm is proposed. The experiment results show that the proposed target tracking algorithm is more accurate than the traditional tracking algorithm in dealing with blocking, thereby meeting the

needs of complex scenes and real-time requirements. However, the proposed algorithm does not perform well in multiple targets tracking when the targets suddenly appear. This matter is the direction of our next work.

Acknowledgements

The work is supported by (1) the Scientific Research Fund of Jilin Provincial Education (20140307), (2) the Scientific and Technological Planning Project of Jilin Province (20120332), (3) A Project Supported by Scientific Research Fund of Jilin Provincial Education (20120268), (4) A Project Supported by Scientific and Technological Planning Project of Jilin Province (20140414064GH), (5) Jilin province Development and Reform Commission projects (2013C040).

References

- Huang, C., Liu, Q., and Yu, S., "Regions of interest extraction from color image based on visual saliency", *The Journal of Supercomputing*, 58(1), 2011, pp. 20-33.
- Collins, R., T., Liu, Y., and Leordeanu, M., "Online selection of discriminative tracking features", *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 27(10), 2005, pp. 1631-1643.
- Yan, L., Yu, Z., Han, N., and Liu, J., "Improved Image Fusion Algorithm for Detecting Obstacles in Forests", *Journal of Digital Information Management*, 11(5), 2013, pp. 379.
- Koldewyn, K., Weigelt, S., Kanwisher, N., and Jiang, Y., "Multiple object tracking in autism spectrum disorders", *Journal of autism and developmental disorders*, 43(6), 2013, pp. 1394-1405.
- Tristán-Vega, A., García-Pérez, V., Aja-Fernández, S., and Westin, C., F., "Efficient and robust nonlocal means denoising of MR data based on salient feature matching", *Computer Methods and Programs in Biomedicine*, 105(2), 2012, pp. 131-144.
- Bahl, A., Ammer, G., Schilling, T., and Borst, A., "Object tracking in motion-blind flies", *Nature neuroscience*, 16(6), 2013, pp. 730-738.
- Kang, W., S., Na, J., H., and Choi, J., Y., "Salient feature selection for visual tracking", *Electronics letters*, 48(18), 2012, pp. 1123-1125.
- Yao, J., C., and Goh, K., L., "A Refined Algorithm for Multisensor Image Registration Based on Pixel Migration", *Image Processing, IEEE Transactions on*, 15(7), 2006, pp. 1839-1847.
- Tsitiridis, A., Richardson, M., and Yuen, P., W., "Salient feature-based object recognition in cortex-like machine vision", *Engineering Intelligent Systems*, 20(1-2), 2012, pp. 59-72.
- Savani, K., and Markus, H., R., "A processing advantage associated with analytic perceptual tendencies: European Americans outperform Asians on multiple object tracking", *Journal of Experimental Social Psychology*, 48(3), 2012, pp. 766-769.
- Hu, W., Li, X., Luo, W., Zhang, X., Maybank, S., and Zhang, Z., "Single and multiple object tracking using log-euclidean riemannian subspace and block-division appearance model", *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 34(12), 2012, pp. 2420-2440.
- Vellingiri, V., Thirumarimurugan, M., and Kannadasan, T., "An Efficient Salient Feature based Histology Image Retrieval", *International Review on Computers and Software*, 9(5), 2014, pp. 793-802.
- Heo, Y., J., Lim, J., H., and Lee, S., J., "EKF-based SLAM Using Sonar Salient Feature and Line Feature for Mobile Robots", *Journal of the Korean Society for Precision Engineering*, 28(10), 2011, pp. 1174-1180.

14. Itti, L., Koch, C., and Niebur, E., "A model of saliency-based visual attention for rapid scene analysis", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(11), 1998, pp. 1254-1259.
15. Župan, R., Sruck, D., and Frangeš, S., "Experiment for determination of map graphics segment standard for handheld crisis maps management", *Tehnički vjesnik*, 19(4), 2012, pp. 917-921.
16. Iodice, S., Petrosino, A., "Salient feature based graph matching for person re-identification", *Pattern Recognition*, 48(4), 2015, pp. 1070-1081.
17. Chua, F., "Learning to inhibit a salient non-target feature", *Journal of Vision*, 14(10), 2014, pp. 1042-1052.
18. Zhao, H., W., Wang, H., Liu, P., P., and Dai, J., B., "A Computer Model of Directional Visual Attention", *Journal of Computer Research and Development*, 46(7), 2009, pp. 1192-1197.
19. Nieves, J., Santos, I., and Bringas, P., G., "Enhancing the prediction stage of a model predictive control systems through meta-classifiers", *Dyna*, 88(3), 2013, pp. 290-298.
20. Ke, H., C., Sun, H., B., Gao, L., and Wang, H., "A Video Image Compression Method based on Visually Salient Features", *Journal of Digital Information Management*, 12(5), 2014, pp. 333-339.