Advances in Computer Science Research (ACSR), volume 61
7th International Conference on Education, Management, Computer and Society (EMCS 2017)

A Novel Target Tracking Algorithm based on Visual Saliency Model

Wenzhun Huang[1, a] and Baohui Zhao[1, b*]

[1]School of Information Engineering, Xijing University, Xi'an 710123, China

[a]huangwenzhun@xijing.edu.cn, [b]hui546837282@qq.com

*The corresponding author

Keywords: Target tracking, Visual saliency, Algorithm, Model description, Math optimization

Abstract. In this paper, we propose the novel target tracking algorithm based on the visual saliency model. Dynamic object tracking problem of the scene is the current hot issues in the field of machine vision research. The main task of it from the dynamic target detection, recognition and tracking in image sequence, or even able to understand and describe the target object behavior that can be widely used in visual surveillance, human-computer interaction, robot soccer, military industry, etc. In view of the tracking object and the different application scenarios, multiple targets tracking problem has always been a difficulty in research of the target tracking. Tracking methods are divided into point tracking, nuclear track, contour tracking, and mixed scenarios. Our paper combines the concepts of the visual saliency to propose the enhance feature extraction methodologies that help to assist us get better tracking accuracy that is innovative.

Introduction

Target tracking is a hot research direction in the field of computer vision, is widely used in many areas such as security monitoring, behavior recognition, target tracking can be divided into three categories: by calculating the parameterized motion model for tracking, only the target location information and ignore the specific shape and at the same time obtain the target location and shape information, such as the target contour tracking, etc.

Used for general target tracking information, such as the distance from radar have been the target of bearing, etc. is affected by the noise pollution, the need for the filtering processing, target state to meet specific application requirements. Filter is generally based on the mathematical model for describing the motion rule of target, if we can accurately on the target movement model, target motion state can accurately figure out. However when how should target maneuver is unknown, uncertain, so difficult to unified modeling of target movement process, make time to time can accurately the movement rule of the target. Target model built by the kinematics equation and the target's actual motion mode there is a mismatch becomes a basic problem of maneuvering target tracking problem, which can lead to filter divergence that cause the target tracking lost [1-2].

For the mentioned issues, existing scholars have proposed the corresponding solutions for better performance. Fuzzy reasoning is based on fuzzy logic, a kind of uncertainty reasoning it is relatively on the basis of the basic classical logic deductive reasoning. Deductive reasoning is a kind of strict reasoning from general to specific. It can be expressed as a syllogism, is the major premise and minor premise and a conclusion. The concept used by deductive reasoning is abstract, clear not ambiguous, the inference rules are used by the certain theory framework is absolutely correct theorem, thus the conclusion is absolutely reliable. The reliability is a major advantage of deductive reasoning as form and logical rigor makes deductive reasoning particularly easy to machine [3].

Biology, especially people with a strong image understanding and pattern recognition, computer vision model inspired by biology has always been a hot spot in the field of image processing. Visual attention mechanism is an organism selectively ability to deal with interested in visual area, its prey and escape on organisms has important significance. Focus areas are usually marginal strong area, at this time due to contrast masking effect, the noise in the region that are largely, so become very not obvious, highlight the importance of edge, instead it gives the local edge comparison model larger weights, give

local statistical mean and variance of smaller weight, in the focus areas, vice versa. Area that due to the human eye is more focused on attention, that is, more sensitive to the focus on regional distortion, so the focus area weight increased obviously is reasonable. On the basis of classifying the distortion, first of all, the residual image local area of the mean and variance statistics, these indicators can very good detection noise distortion to better reflect the statistical features of images. In the Fig. 1, we demonstrate the saliency guided target tracking methodology.

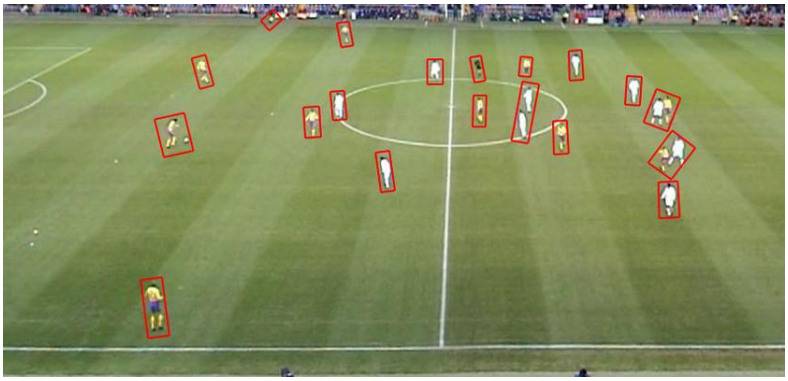

Figure 1. The Demonstration of the Target Tracking Methodology

In this research paper, we propose the novel target tracking algorithm based on the visual saliency model. Visual selective attention is unique to the primate that is an important and effective visual mechanism. This mechanism can be seen as a kind of biological information processing, the choice of a vast amount of visual information is the most important part of the priority for processing. So we can quickly from complex visual scene of priority to obtain interest or important information. In the later sections, we will apply it into the general scenarios for better countermeasures.

The Proposed Algorithm and Methodology

The Image Feature and Selection Methods. In many research fields of computer vision, it is often necessary to map the image to a certain feature space analysis and general processing. As one of the commonly used image features, histogram feature has the characteristics of simple and practical, and has a mature application in many fields [4-5].

The technique is using the geometry of the affine transformation and color interpolation to realize the fusion, its essence is to use the existing image pixel information search features feature matching primitives to keep corresponding relationship between the geometry, and according to the geometric relation between two linear interpolation to generate transition between images, image color realize smooth transition from the source image to the target image and the feature extraction and matching algorithm is the most difficult problems, the new algorithm is proposed in order to solve this problem, the fusion of feature extraction and matching through multiple aspects, and the combination of image distortion, to realize automatic image deformation to get a visual and reasonable in the middle of the interpolation image. Based on the above image deformation and feature point extraction and matching of research shows that did not see a specific feature extraction and matching algorithm is applied to image morphing algorithm based on grid, the automatic image deformation. So this article will feature extraction and matching and image deformation based on grid, the combination of the same scenario, the Angle between the two images in image interpolation conditions, put forward a kind of automatic image morphing algorithm based on feature fusion.

In this paper, the method includes edge detection, feature point extraction, feature point matching and perspective transformation matrix estimation. According to the differential exchangeability to the sequence of convolution calculation, we could express the feature as the follows.

$$\nabla^2 I(x,y) = \frac{\partial^2 I}{\partial x^2} + \frac{\partial^2 I}{\partial y^2} \tag{1}$$

Gauss filtering of edge image median smoothing has resulted in a rapid decline in the number of extremum point, edge image so need to build a set of edge image pyramid will mainly feature points to find to complete registration. How to increase the spatial information of image features is the key to solve the problem. In basic view of the histogram features fast exhaustive studies show that if the calculation of characteristics meets the integral area additive property, the computing efficiency can be greatly. In spite of this, the gradient direction histogram the local block to enhance the character description of the space through the ability to distinguish, to some extent that embodies the structure information of the image, in the case of a gradient to the obvious, therefore, the ability to distinguish between image features have improved significantly as the follows [6].

$$A(predict) = \sum w(s,t)\left|D_I(m+s,n+t)\right| \tag{2}$$

Spatial histogram, edge direction histogram and gradient direction histogram have in common is a part of the use of the structure of the image information that achieve the enhancement purpose to the feature differentiation. However, the degrees of the differentiation of growing at the same time, the robustness of image rotation changes are affected.

The Visual Saliency Model. For an image, the user is only interested in parts of the image area, the region of interest this part represents the user's query intention, and are not interested in most of the rest of the region has nothing to do with the user's query intention. Significant regional image is the most can cause the interest, the user can show the region of the image content. In fact, a significant area of choice is very subjective, because of the different user tasks and knowledge background as for the same image, different users may choose different areas as significant areas.

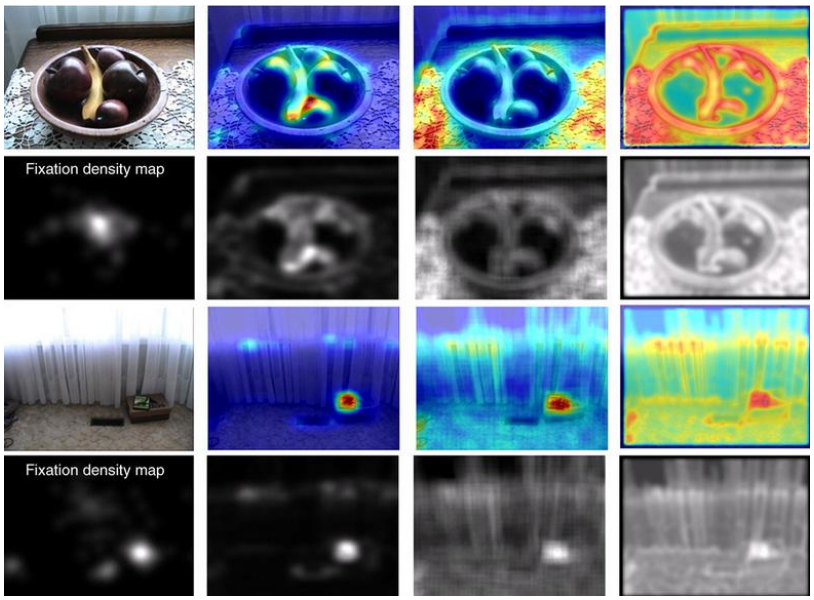

Figure 2. The Illustration of the Visual Saliency Model

Commonly used method is to use the person's attention mechanism on the basis of calculating the significant of the image. Cognitive psychology research has shown that some areas of the image can significantly attract the attention of people and these areas contain a large amount of the information. Cognitive scientists have proposed many mathematical models to simulate the general concentration mechanism. By using the general law in the process of image cognition, the extraction of significant areas more accord with people's subjective evaluation. We can directly to the image as a brightness characteristic figure that also some scholars to stretch or enhancement processing, get the brightness feature maps that is demonstrated as the follows [7].

$$L(x,y) = I(x,y) \tag{3}$$

Usually uses the edge detection operator, such as Canny, Sobel, Laplace operator, by calculating the difference image to realize the local small area in image gradient feature extraction. But this kind of edge operator is sensitive to noise, and often in the gradient feature extracting of strengthening noise at the

same time. Low sensitivity to noise and morphological gradient, gradient extraction can get good results. The formula four illustrates the features.

$$Grad(f)=(f\oplus g)-(f\ominus g) \tag{4}$$

While each kind of new significance after basic model establishment, by some comparative tests illustrates the superiority of the new model, but the contrast is often one-sided, just emphasizes the advantages of the new model at the different angles. Therefore, more objective and comprehensive evaluation of third party reference value. We propose the feature extraction method as the follows.

$$T^{p,q}(x,y)=\sum\sum L(x-r,y-s)\cdot r^{p}\cdot s^{q} \tag{5}$$

Because the interest area of gray level with the surrounding usually large difference, performance on the image to gray scale is not smooth transition, this paper introduce the gradient characteristic further describe its significance. Gradient value, the greater the show that the differences of the region and the surrounding area is larger, this area is significantly higher, the region of interest, the more likely it is that people could obtain through the formula 6.

$$c_{ij}=D\left[\left(\frac{1}{N}\sum_{p=1}^{N_1}V_p\right),\left(\frac{1}{N}\sum_{p=1}^{N_2}V_p\right)\right] \tag{6}$$

The significance of a pixel in the image depends on its own characteristics with the surrounding environment difference, the greater the differences in features of the surrounding environment, the more significant, as a result, we can consider to use each pixel in the image characteristics of adjacent pixels difference to measure the pixels in a local neighborhood of visual significance, namely local visual significant measure. By describing a significant figure and artificial segmentation image to evaluate the consistency of the significant algorithm when applied to the target segmentation based on this view, based on significant figure and artificial segmentation image contrast evaluation system, evaluation index including three groups: directly to the significant figure and artificial segmentation image contrast indicator; For the fixed threshold segmentation based on significant figure and the artificial segmentation image contrast indicator; For the adaptive threshold segmentation based on the significant figure and artificial segmentation image contrast indicators [8].

The Target Tracking Algorithm. As for our proposed algorithm, we should firstly separate it into the listed aspects. (1) Tracking module is an adaptive tracking module that under the condition of the inter-frame motion target visible, use a frame to track current frame on target location. (2) Learning module assessment tracking module and test the performance of the module by generating positive and the negative training samples to complete the detection module updates, eliminate the error of the detection module to produce. (3) Detection module for efficient cascade classifier, a combination of several simple image characteristics, to detect target in real time, at the same time to participate in the tracking module results are correct.

Particle filter is based on the Monte Carlo and recursive Bayesian estimation of nonlinear filtering method, is widely used in target tracking, the particle damping and dilution problems when used in target tracking. To solve the above problems and the double sampling particle sets. First sampling is introduced in the process of sequential importance sampling resampling algorithm, is used to restrain or eliminate small particle weights, keep weight particles, based on the particle weight redistribution of computing resources, get weights particle set, etc.

$$\frac{\partial l(w)}{\partial h_n}=-2C_{yn}w\left(\max\left(1-w^{T}h_n y_n,0\right)\right) \tag{7}$$

For adaptively adjust and optimize the number of encoder to extract features, between the encoder and the linear classifier adds characteristics of incremental learning, layer optimization feature set in order to obtain the features of compact, said to adapt to the target appearance changes in the process of the tracking and hybrid model objective function models the objective function and the discriminant model was generated by the objective function as the follows.

$$L_{hybrid}(x,y) = L_{disc}(x,y) + \eta L_{gen}(x) \quad (8)$$

Add new features in the process of the original feature set parameters unchanged, only the training characteristics and parameters of the newly added to minimize the generation primary model and the discriminant model of objective function to optimize the network parameters. Integration of similar characteristics in the process, choose the most similar in the existing feature set two characteristics as the candidate feature set of integration, at the same time from the characteristics of the concentrated delete these two similar characteristics. TLD target tracking tickets only and integrated modules of the tracking a single target tracking detection module can detect multiple targets for consolidation with comprehensive tracking module and detection module as a result, through the integration of only one of the biggest output similarity ultimate goal and tracking results.

Summary and Conclusion

In this paper, we propose the novel target tracking algorithm based on the visual saliency model. Multiple target tracking is a hot issue in research of tracking in recent decades that it has important theoretical significance and wide engineering application background, every year there are a large number of relevant papers and monographs published. In multiple target tracking, due to the various interference such as meteorology, electromagnetic, false target, sensor set includes not only is the result of objective measurement and also includes a lot of clutter, and the two are indistinguishable. Therefore, how to use a combination of observation data to estimate target state and number of the multiple target tracking is the key to the problem. Our research combines the novel visual saliency model to propose the novel target tracking algorithm that could enhance the performance of tracking accuracy and the effectiveness that is meaningful.

References

[1] A. Andriyenko, K. Schindler and S. Roth: IEEE Conference on Computer Vision and Pattern Recognition (Providence, Rhode Island, USA, June 16-21, 2012), Vol. 1, p. 1926.

[2] B. N. Vo, B. T. Vo, and D. Phung: IEEE Transactions on Signal Processing: A publication of the IEEE Signal Processing Society, Vol. 62 (2014) No. 24, p. 6554.

[3] O. Demigha, W. K. Hidouci and T. Ahmed: IEEE Communications Surveys & Tutorials, Vol. 15 (2013) No. 3, p. 1210.

[4] A. Ickowicz: Computer Science, arXiv preprint arXiv:1410.4262 (2014).

[5] H. L. Yan, G. H. Huang, H. W. Wang and R. Shu: International Conference on Information Science and Cloud Computing (Guangzhou, China, December 7-8, 2013), p. 61.

[6] G. Chen and H. Ling: SPIE Defense, Security, & Sensing. International Society for Optics and Photonics, 2013:873912.

[7] M. F. Zhang and T. Liu: Journal of Intelligent & Robotic Systems, Vol. 81 (2016) No. 3, p. 505.

[8] K. Benbouabdallah and Q. D. Zhu: Research Journal of Applied Science, Engineering and Technology, Vol. 5 (2013) No. 15, p. 4023.