物体の状況を考慮した適応制御を備えた相関フィルターに基づく移動物体追跡

Academic Year 2021
Graduate School of Advanced Mathematical Sciences

Resume of Doctoral Dissertation

Visual Tracking based on Correlation Filter with Adaptive Control

Considering the Situation of Objects

Major: Frontier Media Science Program

Name: TANG ZHAOQIAN

1 Research goal

Computer vision plays a significant role in various areas, and it is a trendy research direction at present. It is an arduous task in computer vision systems that accurately recognizes and tracks objects as human vision systems. However, in some high-level applications of computer vision, visual tracking is an essential component. In recent ten years, the research of visual tracking has achieved outstanding progress. Support vector machines (SVM), a discriminative correlation filter (DCF), and deep learning are used for visual tracking. It is difficult for the trackers based on SVM and DL to achieve real-time performance. Therefore, discriminative correlation filters are widely used in visual tracking, which has made significant progress. Discriminative correlation filter trackers achieve a natural balance between excellent performance and real-time, making visual tracking technology easier to apply in actual life. However, there are still many challenges in visual tracking, such as deformation, illumination variation, occlusion, scale variation, etc. These challenges make it difficult for visual tracking models to perform detection and tracking effectively. Some trackers take different methods to solve these problems, and this leads the trackers is difficult to balance outperformance and real-time.

This thesis proposed three improved trackers based on a discriminative correlation filter to obtain the performance and real-time trackers. The first one is based on the most straightforward CF framework, which solves some common problems, such as scale variation, object rotation, etc. Moreover, the proposed CF tracker achieves robust performance and keeps a high running speed. The second CF framework enhances the robustness of the CF model and has been further improved in many challenges. The proposed tracker obtains a real-time outperformance on different benchmarks. The latest algorithm is aimed at UAV, a popular research direction. The proposed tracker improves the CF model to meet video tracks with low frame rates and high frame rates. The relationship among the three is an extension. That is to say, and the latter is further improved based on the idea of the former.

2 Summary of the chapters

In this thesis, I propose three trackers based on the discriminative correlation filter. These trackers balance real-time and outperformance in tracking. The main structure of this thesis is as follows. The main structure of this thesis is as follows.

In chapter 1, the introduction of visual tracking is presented, and the earlier research of visual tracking is illustrated.

In chapter 2, the classic algorithms (the baseline of the proposed trackers) based on discriminative correlation filters are reviewed. In this chapter, the theory and the unsolved problems of DCF trackers are introduced.

In chapter 3, to improve the performance of the kernelized correlation filter tracker, Correlation Filter tracker using Confidence map and Adaptive model (CFCA) is proposed. Firstly, an improved scale estimation is used to deal with scale variation and accurately estimate a small object's scale variation. Besides, the kernelized correlation filter tracker locates the object position by the maximum score of the confidence map. When the object is occluded or the object deformation occurs, the reliability of the confidence map will decrease. The position corresponding to the maximum score is not necessarily the object's location. The proposed tracker combines the multiple high scores of the confidence map and the similarity of the luminance histogram to improve the precision of detection. And then, CFCA adaptively adjusts the learning rate of the model training by estimating the object's state. In the worst case, CFCA tries to relocate the target object with the peak of the confidence map. The analysis experiment on different databases verifies the validity of the proposed tracker.

In chapter 4, Correlation Filter tracker using a spatial-temporal regularized with Advanced Scale Estimation (CFASE) is proposed, a more robust DCF framework is introduced into the proposed tracker. In temporal regularization, CFASE trains the correlation filter model more precisely using temporal prediction from the previous two filters, the robustness of the DCF model is enhanced. CFASE trains the scale estimation model by the HOG feature and the localization model by the hand-crafted feature to solve scale variation. In tracking, the obtained scale is used to locate the object. This method increases the precision of the scale estimation. Average peak-to-correlation energy (APCE) is introduced to evaluate the accuracy of scale estimation and object location. The outstanding performance of the proposed tracker on different benchmarks demonstrates the validity of the tracker.

In chapter 5, Background-Aware correlation filter tracker with Spatial-Temporal Regularized (BASTR) is proposed, the framework of the proposed tracker is complex. Firstly, adaptive spatial regularization and temporal regularization are proposed to enhance the robustness of the DCF model, and this leads to some kinds of challenges being solved. In general, a tracker can obtain outstanding performance on the benchmark database with the same frame rate, cannot perform well on other databases with different frame rates. The generalize of the proposed tracker is enhanced by selecting the scale estimation method accurately. For videos with a high frame rate, scale pool technology can obtain better performance. In contrast, DSST is better for videos with a low frame rate. By analyzing the experiment, the proposed tracker can meet the requirement of the different benchmarks and obtain excellent performance.

Chapter 6 demonstrates the result of the evaluation experiments on different benchmark databases. From three terms (location, scale estimation, and framework) to analyze the proposed trackers' advantage and relation. This chapter combines the core of the three proposed algorithms and analyzes and discusses the algorithms in depth. Based on the results of the experiment, we elaborate on the connections and differences between the three algorithms. From the term of location, CFCA proposes a robust method to improve detection accuracy, combining the four maximum scores of the confidence map with the luminance histogram similarity. CFASE and BASTR adopt the same method as SRDCF to increase the location estimation accuracy. In section 6.1, the influence of location estimation is analyzed by the experiment about these two location estimation methods. In section 6.2, the influence of scale estimation is analyzed for visual tracking. The qualitative analysis experiments are conducted on different benchmark databases to evaluate the reliability of the proposed scale estimation methods. In section 6.3, the connection among the mathematical model of the three proposed trackers is demonstrated.

In the final chapter, Chapter 7 summarizes this paper.