

An Interactive Visual Object Tracking Technology for Juvenile Fish Motion Trajectory Extraction

Xu Guangzhu^{1,2}, Zhong Chong¹, He Yan¹, Lei Bangjun^{*1,2}, Huang Yingping³ and Shi XiaoTao³

¹College of Computer Science and Information Technology, Three Gorges University, Yichang 443002, P.R.China

²Hubei Key Laboratory of Intelligent Vision Based Monitoring for Hydroelectric Engineering, Three Gorges University, Yichang 443002, P.R.China

³Engineering Research Center of Eco-environment in Three Gorges Reservoir Region, Ministry of Education, China Three Gorges University, Yichang, 443002, P.R.China

xgz@ctgu.edu.cn, z_chong@126.com, bangjun.lei@ieee.com

Abstract

Building hydropower stations continually destroyed the fish habitats and prevented some species of fish from tracing back to spawn. The construction of fishways can effectively alleviate these problems. Therefore, it is necessary to provide practical advices for the construction of fish ladders through analyzing fish motion trajectories. In this paper, we focus on juvenile fish motion tracking and trajectory extraction. Because the target is small and non-rigid, it is difficult to track them correctly during a long period. With the above considerations, this paper presented an interactive object tracking system based on real time compressive tracking algorithm to extract juvenile fish motion trajectories. In order to effectively reduce the number of interactions, a target sample reselection method is proposed with the help of fish swimming direction constrains. The experimental results showed that this system can track a juvenile fish accurately for a long time and extract its trajectory effectively.

Keywords: *Computer Vision, Object Tracking, Motion trajectory, Juvenile Fish,*

1. Introduction

With a long term of evolution to adapt to the natural environment, different kinds of fish will choose different ways to spawn [1]. Migratory fishes have a habit of tracing back to spawn. However as more and more hydropower stations are built, many original natural rivers are blocked which affects the fish spawning badly and leads to the fish resources decline. In order to harmonize the relationship between the construction of dams and ecological environment, researchers have put forward some methods such as building fish-passing facilities, designing cascade hydropower stations. Fishway, as a kind of engineering technique to protect biological resources, can effectively improve the degree of river connectivity and repair river ecosystem. It has become an important method to alleviate the bad effects caused by water conservancy projects.

The design of fish pass facilities is closely related with fish motion behavior, which directly determines the success of fish passing facilities. So the fish way designing needs a long term analysis for fish swimming behavior according to their motion trajectories.

*Corresponding Author.

One typical method for fish motion tracking is using color marks adhered to a fish body [2, 3]. By tracking the marks with computer object tracking technique the fish motion trajectory can be estimated. This method is effective for big fishes but not suitable to juvenile fishes. Another commonly used method is to adapt automatic animal behavior analysis software such as SwisTrack [4]. But the software could not notify users when it misses the tracked object or be confused by multi tracked objects with similar appearance. And the existences of various noises in videos, such as the interactions between multiple fishes, will cause losing tracking target. Therefore, these kinds of software cannot be applied for long term tracking. In order to get more accurate continuous motion trajectories, some researchers record the fish swimming video with video capture card on PC, and track fishes manually offline. Such methods require heavy workload and easy to cause visual fatigue of users. Aiming at above problems, this paper presents a semi-automatic scheme by using an interactive visual object tracking algorithm to extract juvenile fish motion trajectories efficiently.

Target tracking is one of the most important fields of computer vision. Its aim is to estimate the target position on each frame by analyzing the target appearance and motion. According to feature vectors extracted from the target, the same target in different frames can be associated. During this process, moving target can be tracked continuously and the motion trajectory can be obtained [5]. According to the appearance model used, tracking algorithms can be roughly classified as generative model based methods and discriminative model based methods. A generative model based algorithm forms a template to describe the target commonly, and uses it to search for the candidate which has the minimum error of reconstruction [6-7]. A method based on discriminative model combines the foreground and background information to detecting the target, in essence it treats track problem as a classification problem [8-9].

Algorithms based on generative model need a lot of training samples before tracking to learn an appearance model online. These training samples commonly come from first video frames. But most of the algorithms assume that the appearance of the target will not change much during this period. If the appearance of the target changes greatly at the beginning, drift problem is likely to happen. In addition, this kind of algorithm commonly does not use the background information which might improve the tracking accuracy and stability [10]. Algorithms based on discriminative model can achieve more effective prediction performance. However the biggest limitation of this kind of algorithm is that there is a heavy dependency upon selection of training samples. If the selected samples appeared to be confused, this kind of algorithm will fail.

In recent years, the target tracking technology has made a great progress, but by now there is not a universal tracking method because of the target pose and illumination changes, object occlusion and motion blur, and many other factors. The existing tracking algorithms usually perform well in a short time, but it is difficult to maintain a good performance in a long run. Due to the lack of the ability of self-recovery, the existing tracking algorithm often lost target when drift phenomenon happens. SwisTrack proposed by Correll Nikolaus [4, 11-12] is a visual target tracking software with open source codes and it is easy to use in different platform. But in the specific application of juvenile fish tracking, there are still some problems: reflection from water surface will lead to tracking failure; fish swimming action varies dramatically causes fish appearance abrupt changes; the occlusion among multiple juvenile fishes. All these factors lead to the failure of long term tracking juvenile fishes.

In order to extract the juvenile fish motion trajectories during a long term, this paper proposed an interactive fish tracking system based on real-time compressive tracking algorithm. Because the long term tracking process will make mistakes inevitably, this paper presented a target sample reselection strategy with the help of fish swimming direction constrains. By reselecting target samples, the number of interaction is reduced dramatically. The remaining sections of this paper discussed the real-time compression

tracking algorithm, interactive tracking technique, target sample reselection strategy, and some experiments. Section 2 gives the principle of real-time compression tracking algorithm and its advantages in the field of target tracking. Section 3 discusses the interactive object tracking technique for long term target tracking. Section 4 introduces the improved target sample selection strategy and how to reduce the number of interactions with it. Section 5 describes a number of experimental results. Finally, section 6 is a summary of this paper.

2. Real Time Compression Tracking Algorithm

Real Time Compression Tracking Algorithm (CT) was first proposed by Zhang [10]. On one hand, the algorithm is based on a generative model. It uses a sparse measurement matrix that conforms to the compressed sensing restricted isometry property [13] to project multi-scale features of the initial image into a low dimensional feature vector which keeps the most prominent image information. On the other hand, CT is based on a discriminative model as well. It focuses on finding features beneficial to tracking and its purpose is to maximum the difference between the target and the background. In CT, the sparse matrix is applied to extract target features and background features as positive samples and negative samples to update classifier online. Then the updated classifier is applied to find out possible objects that like target but don't like background mostly in follow-up frames. The algorithm is fast and can deal with occlusions, appearances and light change, and movement blur to some extent. So it got much attention of many related researchers [14-18].

The CT algorithm has two main components: classifier updating and target searching. The initial target location is set manually in advance. In each frame, positive samples are selected in the region with a distance less than α from the target center l_i ($\{z | \|l(z) - l_i\| < \alpha\}$), denoted with white bounding boxes shown in Figure 1(a). Negative samples are chosen in the region with a distance less than β but greater than ζ from the target center l_i ($\{z | \zeta < \|l(z) - l_i\| < \beta\}$), denoted with white bounding boxes shown in Figure 1(a)). Then a projection matrix R conforming to the compressed sensing restricted isometry property is used to extract feature vectors of the target and its surrounding background. These features then are used to update a naive Bayes classifier online. Figure 1(a) shows the process of classifier updating. In the next frame (here noted with $t+1$), features are extracted with the same sparse measurement matrix, then the updated classifier evaluates samples with that features, and selects an object that like target most as the target for tracking. Figure 1(b) illustrates the process of selecting target. In addition to the first frame, each frame needs to update the classifier and search for candidate targets.

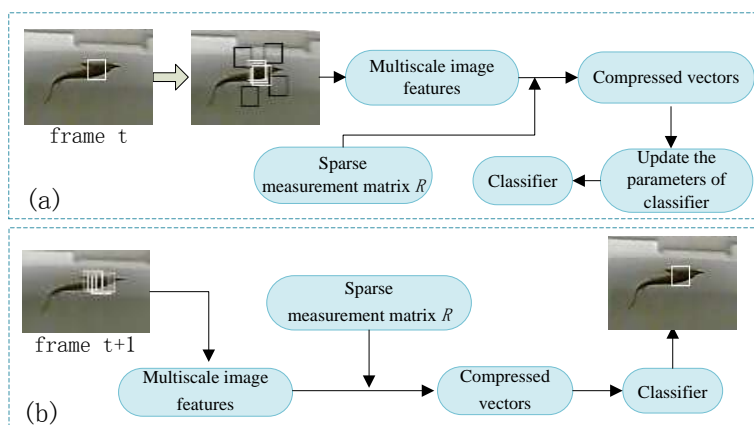


Figure 1. Process of Updating a Classifier (a) and Selecting the Target (b)

3. Interactive Object Tracking Technology

Long term tracking is a very challengeable problem, and most tracking algorithms could not maintain a good performance in a long time. One of the reasons is that most of the algorithms assume that the appearance of the target does not change much in the first few frames. In reality, the target appearance may have changed suddenly. In order to keep track stability during a long time, the information of target appearance changes must be collected timely which is caused by sudden object pose changes or illumination changes. But it is hard to include so many important information in general model, especially in real time tracking. Another important reason is that most of the existing algorithms could not detect the drift phenomenon and recover from it. Once the target location is initialized, most of the tracking is becoming a series of prediction of the target motion in following continuous frame, so that this cumulative drift phenomenon is inevitable unless the location of the target can be reinitialized again [19]. This situation has provided us with a new idea. For a long time video, reinitializing the location of the target can be applied by manual when tracking error happens. With the help of users, we can insure the accuracy of juvenile fish motion trajectories extracted from the long term tracking results. After analyzing the juvenile fish tracking problem, this paper proposed a CT based interactive tracking technology for long term juvenile fish swimming trajectories extraction.

Interactive object track technology applied CT algorithm to extract the features of the juvenile fish head, then track the fish head and get the trajectories of the fish. The speed of the fish can be calculated with the frequency of the camera and the distance from the position of the camera to the fish in practice. When tracking error is found, users can return to the error frame and reinitialize fish location to continue the tracking. Figure 2 is the user interface of our system. The implementation process of the interactive object tracking will be described as follows:

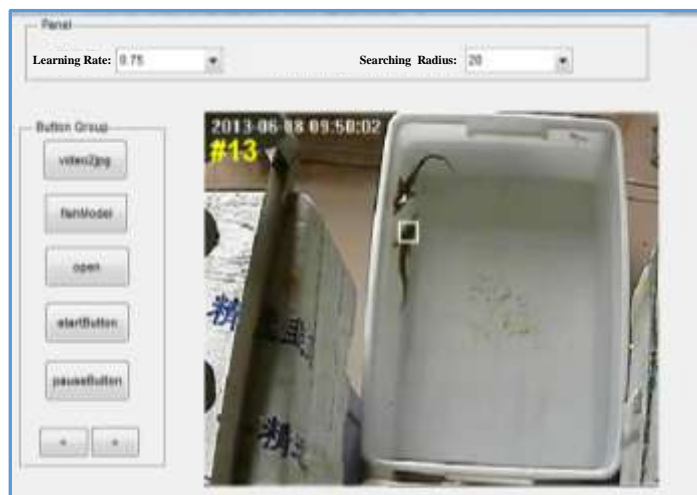


Figure 2. The GUI of the Proposed Interactive Object Tracking System

- (1) First the fish video should be converted to image frames, which can decrease the memory occupation of the video file. This can be done by clicking the button of 'video2jpg', and selecting video that need to be converted. Then the software will convert the video into images.
- (2) Click on the 'open' button to open the folder that contains converted images. Select the first frame to load.
- (3) Click on the 'startButton' button and draw the track rectangle when the mouse appear cross shape (adjust the size of the box). Then double click on the rectangle to confirm track object.

- (4) Positive samples are selected in the region around the target with a distance less than α from the target center l_i ($\{z | \|l(z)-l_i\| < \alpha\}$). The value of α is 4. Negative samples are chosen in the region around the target with a distance less than β but greater than ζ from the target center l_i ($\{z | \zeta < \|l(z)-l_i\| < \beta\}$). The value of ζ and β is 8 and 30 respectively. Then multi-scale features of positive samples and negative samples will be extracted with the projection matrix R . Finally the naive Bayes classifier will be updated with these features.
- (5) In the next frame, candidate fish head patches samples will be searched around the target location of last frame within a region with the distance less than γ from the target center l_i ($\{z | \|l(z)-l_i\| < \gamma\}$). The value of γ is 20. Then the multi-scale features of these samples will be extracted with the same projection matrix R . Finally the classifier selects the sample that is like most the target in previous frame as the target. Repeat step 4 and step 5 until tracking error happens.
- (6) Click on the 'pauseButton' button when tracking fails which can be judge by user manually and the tracking will be suspended.
- (7) Due to the real-time performance of CT algorithm, when user find the tracking error the frame could not just the one with error. So, click the button labeled '>' or '<' to find the error frame. Jump to step 3 when locating the error frame. Then start step 4 and step 5 to continue the tracking. Tracked results are recorded in the date file named data.txt. It can be used to extract the juvenile fish motion trajectories.

4. Target Sample Reselection

In CT algorithm the target searching space is within a region with the distance less than γ from the target center l_i ($\{z | \|l(z)-l_i\| < \gamma\}$). But for the specific application of fish tracking, there is a certain swimming direction for the fish and the swimming direction of the fish in two continuous frames cannot be opposite. And the two fishes will overlap in one point and then swim in the direction of their own when there are two or more fishes. Since the CT algorithm choose samples randomly around the target in previous frame, there will be a lot of non-target samples when two fish overlap. In addition, the target is small (about 20×20 pixels). It is likely to take another fish head as the current frame target when two fish overlap. This will cause tracking errors. Through the observation of multiple videos, the frequency of fish overlapping is relatively high. Therefore, it is necessary to reduce the tracking error caused by the overlapping.

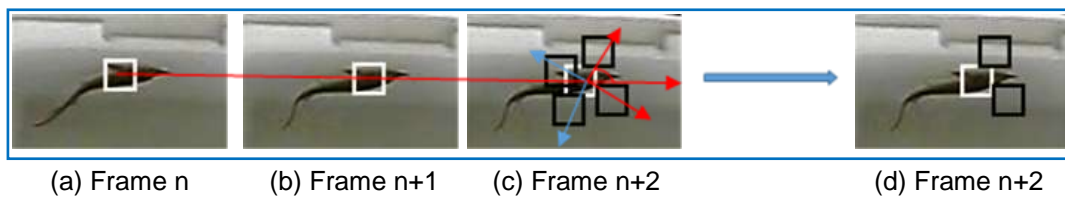


Figure 3. Illustration of the Target Samples Reselection Strategy

Figure 3(a) and Figure 3(b) is the frame n and frame $n+1$. Candidate samples in frame $n+2$ will be selected around the fish position in the frame $n+1$ in order to get the fish position of the frame $n+2$ (as shown in Figure 3(c)). In our study samples selection is according to a direction constraint consideration that the direction of the fish swimming cannot change abruptly in continuous frames. A vector $\overline{(n, n+1)}$ can be formed by linking the center point of the frame n to that of the frame $n+1$. Another vector $\overline{(n+1, n+2)}$ links

the center point of the frame $n+1$ to that of all samples in the frame $n+2$. If the dot product of vector $\overline{(n,n+1)}$ and vector $\overline{(n+1,n+2)}$ is greater than zero, the sample will be retained to be identified. Otherwise, the sample would be thrown away. Finally the target samples searching area is where the vector dot product is greater than zero, as shown in Figure 3(d).

5. Experimental Results and Analysis

The algorithm flow of the proposed interactive juvenile fish tracking technology is shown in Figure 4 below.

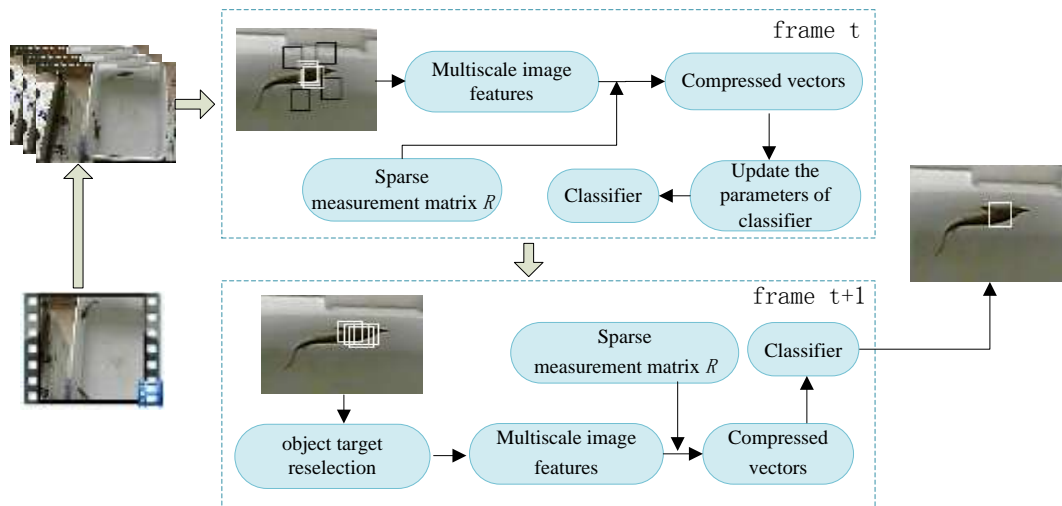


Figure 4. The Flow of the Interactive Juvenile Fish Tracking Technology

5.1. Experimental Results

Some experimental results with and without target sample reselection strategy in the same video are shown in Figure 5(a) and (b). The initial location and the size of the tracking box are set to the same.



Figure 5(a). Frame #62 to frame #69 in one test video. Dashed Box Shows the Tracking Results without Sample Reselection, and the Solid line Box Illustrates the Tracking Results with Sample Reselection



Figure 5(b). Frame #113 to frame #120 in one test video. Dashed Box Shows the Tracking Results without Sample Reselection, and the Solid Line Box Illustrates the Tracking Results with Sample Reselection

Figure 5(a) shows a set of frames in a fish video, these frames show that two fishes are swimming, and overlap in some frames, but they swim along their original direction after overlapping. The tracking results with and without considering swimming direction constraints are shown with bounding boxes with solid line and dashed line respectively. As shown from frame #62 to #65, both tracking algorithms maintained good performance before two fishes overlapped. When the two fishes overlapped in the frame 66, the original algorithm tracked the fish swimming in the opposite direction and the improved algorithm filtered the samples with opposite direction with the help of the direction constraint. This improves the tracking accuracy greatly, and decreases the interaction frequency dramatically. Figure 5(b) shows frames 113 to 120 in a video. There are two fishes in the video, they swam along their own direction and overlapped at frame 115. It is obvious that the original algorithm tracked another fish head falsely and the improved algorithm could keep tracking correctly.

5.2. Juvenile Fish Motion Trajectories Extraction based on Tracking Results

The position data of the fish head obtained with the proposed interactive fish tracking technology is stored in the file named data.txt. Each tuple inside the position data consists of seven fields, these fields are frame number, the x coordinate of the center point of the tracking rectangle, the y coordinate of the center point of the tracking box, the x coordinate of the point in the upper left corner of the tracking rectangle, the y coordinate of the point in the upper left corner of the tracking box, the width of the tracking box, and the height of the tracking box. Figure 6 shows an example of the data file. The data can be used for further analysis. Figure 7 visualizes the part of the data. Figure 7(d) is the first frame of a video, and the resolution is 352×288 . Two fishes swims within a white gutter in this video. During the whole period, the fish keeps swimming around the perimeter zone of the gutter and cross the tank some times. Figure 7(a), (b), and (c) present the position of the fish and show that the fish swimming trajectory is periodic, which is consistent with the real swimming trajectory. These data verifies the correctness of the results.

1	202.24	176.02	120.70	104.02	22.107	22.200
2	177.24	104.22	202.70	104.02	22.107	22.200
3	202.24	176.02	120.70	104.02	22.107	22.200
4	177.24	104.22	202.70	104.02	22.107	22.200
5	202.24	176.02	120.70	104.02	22.107	22.200
6	177.24	104.22	202.70	104.02	22.107	22.200
7	202.24	176.02	120.70	104.02	22.107	22.200
8	177.24	104.22	202.70	104.02	22.107	22.200
9	202.24	176.02	120.70	104.02	22.107	22.200
10	177.24	104.22	202.70	104.02	22.107	22.200
11	202.24	176.02	120.70	104.02	22.107	22.200
12	177.24	104.22	202.70	104.02	22.107	22.200
13	202.24	176.02	120.70	104.02	22.107	22.200
14	177.24	104.22	202.70	104.02	22.107	22.200
15	202.24	176.02	120.70	104.02	22.107	22.200
16	177.24	104.22	202.70	104.02	22.107	22.200
17	202.24	176.02	120.70	104.02	22.107	22.200
18	177.24	104.22	202.70	104.02	22.107	22.200
19	202.24	176.02	120.70	104.02	22.107	22.200
20	177.24	104.22	202.70	104.02	22.107	22.200

Figure 6. Fish Head Position Data

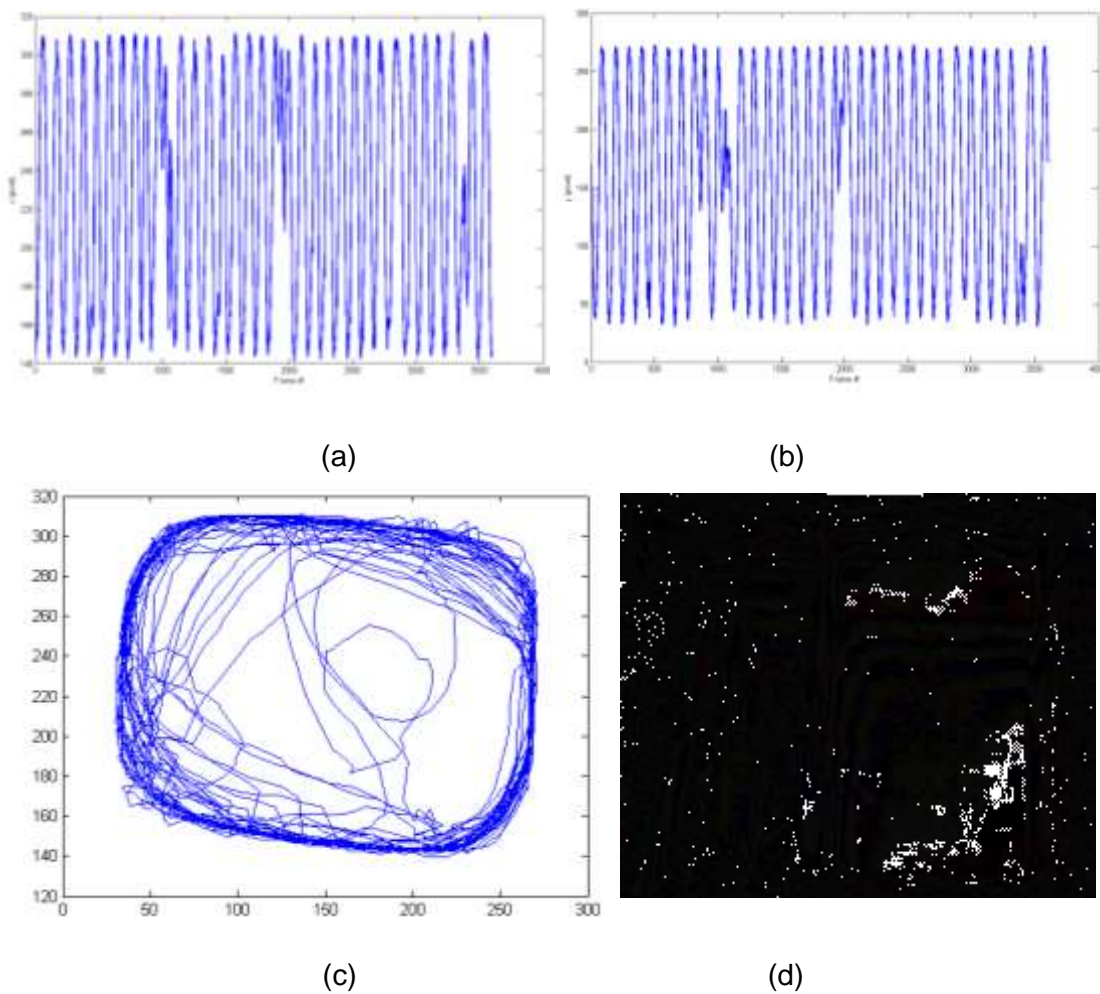


Figure 7. The Analysis of Fish Head Position with Interactive Fish Tracking Technology (a) Visualization of the x Coordinates (b) Visualization of the y Coordinate (c) The Fish Motion Trajectories (d) First Frame of a Video

In these videos, the phenomenon that two fish overlap appears frequently. This paper puts forward a target sample reselection strategy which can reduce the number of interaction effectively and improve the accuracy of the tracking.

6. Conclusion

This paper presented an interactive visual object tracking system to extract the motion trajectory of juvenile fishes. For the fish tracking, we found that fish swimming movement owns a certain direction which cannot be opposite in two continuous frames. At the same time, the frequency of two fish overlapping and swimming along their original direction in a video is high. In order to reduce the interactive frequency and improve tracking accuracy, a target sample reselection strategy is proposed. Experimental results show that the method of target sample reselection can reduce the number of interaction dramatically and improves the trajectories extraction accuracy.

Acknowledgment

This work was financially supported by national natural science foundation of china (61402259, U1401252), Hubei key laboratory of intelligent vision based monitoring for hydroelectric engineering open found project (2014KLA10, 2014KLA02), Engineering Research Center of Eco-environment in Three Gorges Reservoir Region, Ministry of Education, China Three Gorges University open found project (KF2015-10).

References

- [1] Y. Y. Watanabe, Q. W. Wei, H. Du, L.X. Li and N. Miyazaki, "Swimming behavior of Chinese sturgeon in natural habitat as compared to that in a deep reservoir: preliminary evidence for anthropogenic impacts", *Environmental Biology of Fishes*, vol.96, no. 1, (2013), pp. 123-130.
- [2] H. Yan, Y. M. Su and L. Yang, "Experimentation of Fish Swimming Based on Tracking Locomotion Locus", *Journal of Bionic Engineering*, vol.5, no. 3, (2008), pp. 258-263.
- [3] <http://bioold.science.ku.dk/jfsteffensen/software.htm>
- [4] <http://en.wikibooks.org/wiki/SwisTrack>
- [5] A. S. Jalal, V. Singh, "The state-of-the-art in visual object tracking", *Informatica*, vol.36, no.3, (2012), p. 227-248.
- [6] Y. M. Li, "On Incremental and Robust Subspace Learning", *Pattern Recognition*, vol.37, no.7, (2004), p. 1509-1518.
- [7] D. Skocaj, A. Leonardis, "Weighted and Robust Incremental Method for Subspace Learning", *Proceedings of the IEEE International Conference on Computer Vision, Nice, France*, (2003) October 14-17.
- [8] T. Parag, F. Porikli, A. M. Elgammal, "Boosting Adaptive Linear Weak Classifiers for Online Learning and Tracking", *Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition, Alaska, USA*, (2008) June 24-26.
- [9] S. Avidan, "Ensemble Tracking", *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol.19, no.2, (2007), pp. 261-271.
- [10] K. H. Zhang, L. Zhang, M.H. Yang, "Real-Time Compressive Tracking", *Proceedings of the European Conference on Computer Vision, Firenze, Italy*, (2012), October 7-13.
- [11] L. Thomas, R. Pierre, C. Chris, C. Nikolaus, J. Jacques, M. Alcherio, "SwisTrack - A Flexible Open Source Tracking Software for Multi-Agent Systems", *Proceedings of the IEEE/RSJ 2008 International Conference on Intelligent Robots and Systems, Nice, France*, (2008) September 22-26.
- [12] N. Correll, G. Sempo, "SwisTrack: A Tracking Tool for Multi-Unit Robotic and Biological Systems", *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems, Beijing, China*, (2006) October 9-15.
- [13] E. J. Candes, T. Tao, "Decoding by Linear Programming", *Information Theory, IEEE Transactions on Information Theory*, vol.51, no.12, (2005), pp. 4203-4215.
- [14] K. H. Zhang, L. Zhang, M. H. Yang, "Fast Compressive Tracking", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.36, no.10, (2014), pp. 2002-2015.
- [15] J. Jing, G. Z. Xu, B.J. Lei, "Using Appearance Re-matching to Improve Real-time Compressive Tracking", *Proceedings of the Fifth ACM International Conference on Internet Multimedia Computing and Service, HuangShan, China*, (2013) August 17-19.
- [16] Y. He, G. Z. Xu, B.J. Lei, "An Improved Real-time Compressive Tracking Method", *Proceedings of the Fifth ACM International Conference on Internet Multimedia Computing and Service, HuangShan, China*, (2013) August 17-19.
- [17] S. Zhang, H. Yao, X. Sun, "Sparse coding based visual tracking: Review and experimental comparison", *Pattern Recognition*, vol.46, no.7, (2013), pp. 1772-1788.

- [18] H. Xu, F. Yu, "Improved Compressive Tracking in Surveillance Scenes", Proceedings of the Seventh IEEE International Conference on Image and Graphics, QingDao, China, (2013), July 26-28.
- [19] M. H. Yang, H. Jeffrey, "Toward Robust Online Visual Tracking", Edited B. Bhanu, Springer, New York, (2010), pp. 119-136.

Authors



Xu GuangZhu, (1979.1-) is currently an associate professor at University of China Three Gorges University. He received his M.S. and Ph.D. degrees in Circuits & Systems and Radio Physics at Lanzhou University in 2004 and 2007 respectively. His main research interests are visual object detection and tracking, medical image processing and biometrics.



Zhong Chong, (1992.8-) graduated at the China Three Gorges University in 2014. Currently he is a MSc. student at the College of Computer Science and Information Technology of China Three Gorges University. His research interests are oriented to visual object detection and pattern recognition.



Lei BangJun, (1973.6-) received his Ph.D. from TUDelft, the Netherlands in 2003. He is currently a professor at University of China Three Gorges University. His majors include low-level image processing, 3-D imaging and multi-dimensional intelligent computer vision. He is an IEEE senior member and have published more than 60 papers.