

Moving Targets Tracking Based on Improved Compressive Sensing Algorithm

Songtao Jiang*

*School of Electrical Engineering, Tianjin University of Technology, Tianjin
300384, China
446707284@qq.com*

Abstract

Aiming at the problems of tracking failure and high error rates in the case of target occlusion and textures change in moving targets tracking, an improved compressive sensing algorithm is proposed. The threshold of a sigmoid function was calculated to judge whether occlusion exist. This threshold determined the classifier update occasion in order to avoid object model mistake update when heavy occlusion. To solve the tracking instability cause by single feature, a feature fusion method under rules that combined grey features and texture features was presented. Tracking was guided by the fused feature. Experiments show that the improved algorithm has 17.84% higher success tracking rates, and 11.59% lower average error rate compares with the traditional algorithm.

Keywords: *Compressive Sensing, Target Tracking, Object Model, Feature Fusion*

1. Introduction

Target tracking technology is an integral part of the computer vision system link, in military guidance, navigation, robot, intelligent transportation and other fields are widely used. The common tracking algorithms: background difference method, optical flow method, the inter frame difference method and particle filter algorithm and so on.

In many tracking algorithms, compressed sensing algorithm is a research hotspot in recent years, first proposed by Donoho [1] and Candes [2], the theory breakthrough the traditional concept of sampling, not according to the signal frequency and bandwidth, but depends on the structure of the signal itself is. Its theory is: if the sampling signal in an orthogonal space is sparse or compressible, discrete samples can be much less than the sampling rate of Nyquist under the condition of obtaining the signal, and can reconstruct high probability accurately the signal. By citing the theory and classification of thinking, compressed sensing is applied to tracking field, produce compressed sensing tracking algorithm (referred to as CT algorithm), to realize moving target feature extraction and real-time tracking. But the traditional CT algorithm in object occlusion and light texture, dramatic changes still exist defects.

Update mode parameter classifier in CT algorithm is fixed, namely: after determining the tracking position, positive and negative samples characterized by the current frame and the previous frame compression domain collected to update the model parameters. But if this happens occlusion, the background error will be in the update process is introduced to the model parameters in the end, gradually Many a little make a mickle. Lead to tracking failure. Therefore, to solve this problem, this paper presents an improved method, makes the appearance of large occlusion when the appearance model stop updating. In addition, the signal characteristics of the CT algorithm in independent and sampling of the original signal and the one corresponding, can quickly extract feature tracking, good real-time performance. But the sampling characteristics of a single [3], when the target texture

* Corresponding Author

changes prone to drift, tracking stability is not good enough. This paper presents the integration method of gray features and texture features, is solved by tracking instability caused by a single feature.

2. CT Algorithm

First of all, the positive and negative samples simultaneously projection; Secondly, compression perception algorithm is adopted to improve the sampling, and uses a RIP conditions very sparse matrix measurement of the high-dimensional feature dimension reduction compression matrix; Finally, through a simple bayesian classifier to extract the discriminant classification, the characteristic of each frame for online learning and update the classifier. Tracking using computed tomography CT algorithm has good real-time performance, the small amount of calculation after compression, fast processing speed, can reach 40 cycles per second.

Work process is shown in Figure 1:

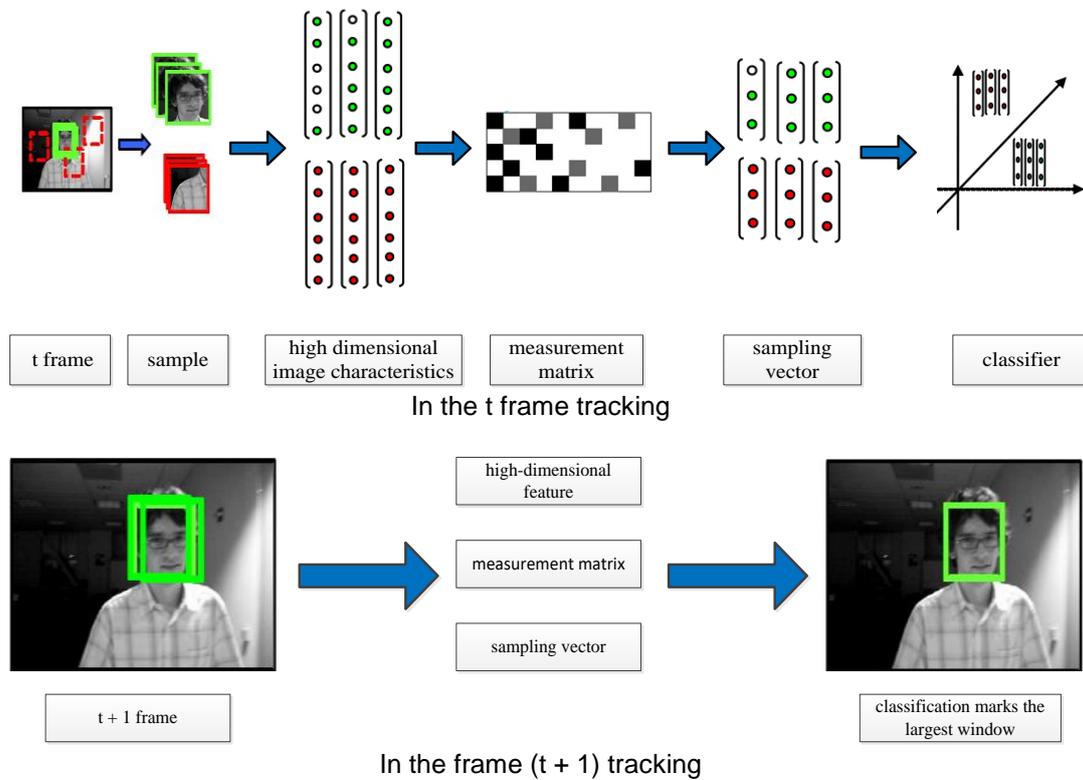


Figure 1. Algorithm for Tracking Work Flow Chart

As shown in Figure 1, t frame, the first scan the whole image sampling, in order to get some virlisen samples (target) and negative samples (background), won the high dimensional image characteristics of these samples at the same time, and through the high-dimensional sparse measurement matrix image feature dimension reduction compression for the low dimensional sample characteristics, finally through the classifier to classify sample characteristics.

T + 1 frame, according to a frame around the identified target location of the scan, so that you can avoid repeated every time the whole image sampling, greatly reduce the amount of calculation and to improve processing speed, get a number of positive and negative samples, and on the same frame of operation process, through sparse dimension measurement matrix, and get the samples after dimension reduction characteristics of direct sampling frame t trained classifier to its classification, take

the classification score the largest window is the need to track the target. Therefore, thus to realize the compression perception tracking algorithm for moving target from $t + 1$ frame to t frame the tracking.

2.1. Feature Extraction.

Compressed sensing theory [4] is proposed: R is a $n \times m$ random measurement matrix, $n \ll m$. It can make use of the formula (1) to compress a high dimensional image signal x (m dimension) into a low dimensional signal v (n dimension):

$$v = Rx \tag{1}$$

Johnson-Lindenstrauss inference [5] show that: if the random selection of a suitable subspace, two projection vector space to the subspace, then there is the distance relationship between a high probability they will be retained. Baraniuk [6] proves that inferences about the Johnson-Lindenstrauss random matrix also satisfy the compressed sensing theory of Rip (restricted isometry property conditions. That is to say, if the random measurement matrix R satisfies the Johnson-Lindenstrauss inferences and x is compressible or sparse, can with high probability from the v under the condition of minimum error will be x reconstruction.

CT characteristic of this algorithm is to select a suitable measurement matrix, because R is a random measurement matrices satisfy the RIP condition, but the dimension of m is very high, matrix is relatively dense, the computation and storage consumption is very large. So the sparse measurement matrices [7] the for dimensionality reduction, $r_{ij} \sim N(0,1)$, Its matrix elements are defined as follows:

$$r_{ij} = \sqrt{s} \times \begin{cases} 1 & p = \frac{1}{2s} \\ 0 & p = 1 - \frac{1}{s} \\ -1 & p = \frac{1}{2s} \end{cases} \tag{2}$$

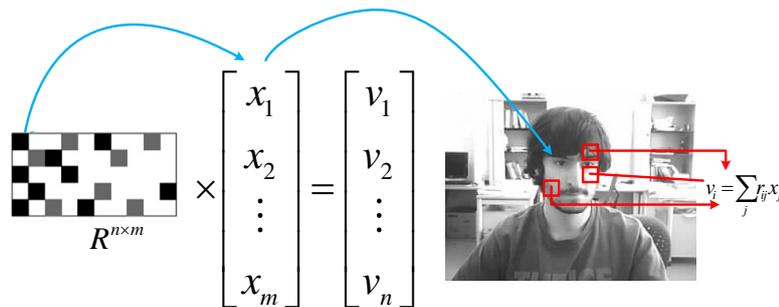


Figure 2. Graphical Representation of Feature Extraction

When $s=3$, $2/3$ probability is 0. Thus, the calculation of the original characteristics of data $1/3$, very low computation complexity. In fact, compression features on the perception of extracted essence is to extract the features of a small part of the original regional out, then random add to or subtract the final composition of various features, feature generation process as shown in Figure 2.

The figure shows that R on the original image characteristic x mapping, to get the sample feature v . In the measurement matrix R , gray, black and white representing matrix elements is positive, negative and zero. Arrows illustrate each row of the matrix R nonzero elements perceived x in an element.

2.2. Construction and Update of Classifier

Assuming that the v of each element in the v_i distribution is independent, and then use the naive Bias classifier [8] to model:

$$H(v) = \log\left(\frac{\prod_{i=1}^n p(v_i|y=1)p(y=1)}{\prod_{i=1}^n p(v_i|y=0)p(y=0)}\right) = \sum_{i=1}^n \log\left(\frac{p(v_i|y=1)}{p(v_i|y=0)}\right) \quad (3)$$

$p(v_i|y=1)$ and $p(v_i|y=0)$ according to the Gaussian distribution, $y \in \{0,1\}$ for binary variables, used to represent the sample of positive and negative, when $y=0$ for negative samples, when $y=1$ is positive sample, assuming that the two classes of prior probability $p(y=1) = p(y=0)$. Diaconis and Freedman demonstrated: high dimensional random vector of random projection almost all obey the Gaussian distribution [9]. So $H(v)$ classifier in the distribution of $p(v_i|y=1)$ and $p(v_i|y=0)$ are assumed to be in accord with the Gaussian distribution of the $(\mu_i^1, \delta_i^1, \mu_i^0, \delta_i^0)$ parameter.

$$\begin{cases} p(v_i|y=1) \sim N(\mu_i^1, \delta_i^1) \\ p(v_i|y=0) \sim N(\mu_i^0, \delta_i^0) \end{cases} \quad (4)$$

The μ_i^1 and δ_i^1 in the formula are representative of the mean and standard deviation of the target samples, μ_i^0 and δ_i^0 represent the mean and standard deviation of the background samples.

$$\begin{cases} \mu_i^1 \leftarrow \lambda \mu_i^1 + (1-\lambda) \mu^1 \\ \delta_i^1 \leftarrow \sqrt{\lambda(\delta_i^1)^2 + (1-\lambda)(\delta^1)^2 + \lambda(1-\lambda)(\mu_i^1 - \mu^1)^2} \end{cases} \quad (5)$$

$\lambda > 0$ for learning factors, it affects the speed of the update, the smaller the λ that the faster the speed of the classifier update.

3. Improvement of Compressed Sensing Tracking Algorithm

Classifier parameters in CT algorithm update model is fixed, namely: when the position tracking, in the current frame and a frame of the positive and negative samples were collected in the compressed domain feature to update the model parameters. But if the condition, the background error will be introduced into the model in the process of updating parameters, gradually many a mickle makes a muckle, and eventually led to the tracking failure. In addition, CT algorithm characteristics between independent and sampling signals and the original one to one correspondence, able to quickly extract the features of real-time tracking is very good. But sample characteristics of a single, easy to happen when the target texture dramatic changes in drift, tracking stability is not good enough. So this article will be to improve the compression perception tracking algorithm, to solve these two problems.

3.1 Improvement of Occlusion

Through the classifier updating mechanism, we find that, if in the update process error will be piled up caused the gradual drift. Therefore, when the target is occluded, will cause some error is introduced to model parameters. In the follow-up process, if not corrected, the error will accumulate, which led to the failure of target tracking. Therefore, this paper puts forward the according to the classifier response value of $H(v)$ and sigmoid function occlusion detection method. In the target was severe occlusion or problems in the new template, stop to update the model parameters.

Based on the detection of tracking can be seen as a two yuan classification problem[4],The classifier response value $H(v)$ or sigmoid function response value larger as new tracking results. Sigmoid function as shown in the formula (6), definition: response values if greater than 0.5 ($H(v) > 0$), determine the sample as positive samples and negative otherwise samples. In CT tracking algorithm for the detection stage, if the sample classifier response values are less than 0 is classifier target samples (response sample maximum) is negative samples, this is generally in the target was occluded appear. Therefore, the dissertation put forward that at this time does not update the target is more appropriate, so as to avoid the errors introduced by the model parameters, resulting in a target drift.

$$\sigma(H(v)) = \frac{1}{1 + e^{-H(v)}} \quad (6)$$

$$H_{Max} = \max(H(x_i)) \quad (7)$$

When $H_{Max} < 0$, the classifier is not updated.

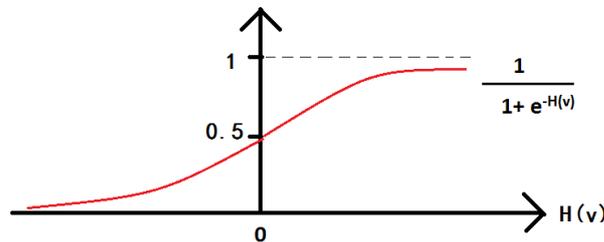


Figure 3. Sigmoid Function Diagram

3.2. Improvement on Single Feature Problem

Because the number of coefficients of each row in the random measurement matrix R used in CT is randomly selected between 2,3,4, when the feature extraction formula (2) is $s = 3$, the probability of each line in the random measurement matrix is only 1 or -1 is as follows:

When nonzero coefficient is 2: $r = \frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$

When nonzero coefficient is 3: $r = \frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} = \frac{1}{8}$

When nonzero coefficient is 4: $r = \frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} = \frac{1}{16}$

So the probability of only 1 or -1 exists in the random measurement matrix is:

$$r_i = 2 \times \left[\frac{1}{3} \times \left(\frac{1}{4} + \frac{1}{8} + \frac{1}{16} \right) \right] = \frac{7}{24} \approx 0.3 \quad (8)$$

Accordingly, sparse measuring each line in the matrix R nonzero vector is the probability of one or both to 1 30%, sampling characteristics embodied in grayscale characteristics of images; Each line and sparse in the measurement matrix R nonzero vector 1 and 1 there is a 70% chance of at the same time, the sampling feature for image texture feature. As a result, the vast majority of cases, the characteristics of the sample are for texture feature, the image texture feature reflects the difference between sample piece, when great changes occurred when the target texture, the tracker will appear unstable problem.

If two characteristics of guiding track at the same time, the tracking robustness is relatively better. Therefore, in order to solve the problem of characteristics of a single, gray features and texture features of the proposed fusion method, by improving the measure to increase the probability characteristic of gray level matrix R, as shown below:

$$g_1 r_i = r_i = \sqrt{s} \times \begin{cases} 1 & p = \frac{1}{2s} \\ 0 & p = 1 - \frac{1}{s} \\ -1 & p = \frac{1}{2s} \end{cases} \quad (9)$$

Can be seen from the type, 1 and 1 sampling matrix scenario will not appear at the same time, as a result, after sampling the features only embodied in gray. Set the second sampling matrix to only embodied in the texture features, such as formula (10)

$$g_2 r_i = r_i = (-1)^k \sqrt{s} \times \begin{cases} 1 & p = \frac{1}{2s} \\ 0 & p = 1 - \frac{1}{s} \\ -1 & p = \frac{1}{2s} \end{cases} \quad (10)$$

K for the sparse coefficient of the bank. Will merge the two transition matrix will constitute the additional new sampling matrix, the formula as shown in (11):

$$c r_i = \begin{cases} g_1 r_i & 70\% \\ g_2 r_i & 30\% \end{cases} \quad (11)$$

Is the new characteristic equation as shown in formula (12):

$$v_1 = CR \times x \quad (12)$$

Additional sample is obtained by sampling characteristics, of which 70% of gray-scale features, 30% of the texture features, and compared with the original characteristic equation to get the characteristics of the fusion, guiding track at the same time, realize the characteristics of sampling, the characteristics of both the texture features as well as gray, makes the tracking effect is improved, the stability was strengthened.

3.3. Algorithm Flow

1) Initialization of the selected target (automatic selection of moving objects, can also manually selected tracking area), the collection of target samples and background samples;

2) According to the formula (2) to generate the feature extraction matrix r_1 , then according to the formula (9) and the value of matrix r_1 to generate the feature extraction matrix r_2 ;

3) According to the formula (1) to generate a feature vector v , using the formula (4) to calculate the sample characteristics of the mean and the variance of μ^0 and δ^0 , μ^1 and δ^1 ;

4) Read a new frame image, the search radius and find the candidate region and calculate the feature and then bring them into type (3), using classifier H (V) to classify the feature. Find calculated values the largest sample as the current tracking to the target location. If $H_{Max} < 0$ is not updated, return the (1) step to the next frame were sampling and processing;

5) When $H_{Max} > 0$, classifier update, get the new target location l_t , with l_t as the center α for radius acquisition positive sample set: $D^\alpha = \{z | \|l(z) - l_t\| < \alpha\}$; $\alpha < \zeta < \beta$ is the radius of the collection of negative samples: $D^\beta = \{z | \zeta < \|l(z) - l_t\| < \alpha\}$, according to the characteristics of the two sample set, update the classifier parameters;

6) Jump to 1) for next frame processing;

Algorithm flow chart is as follows:

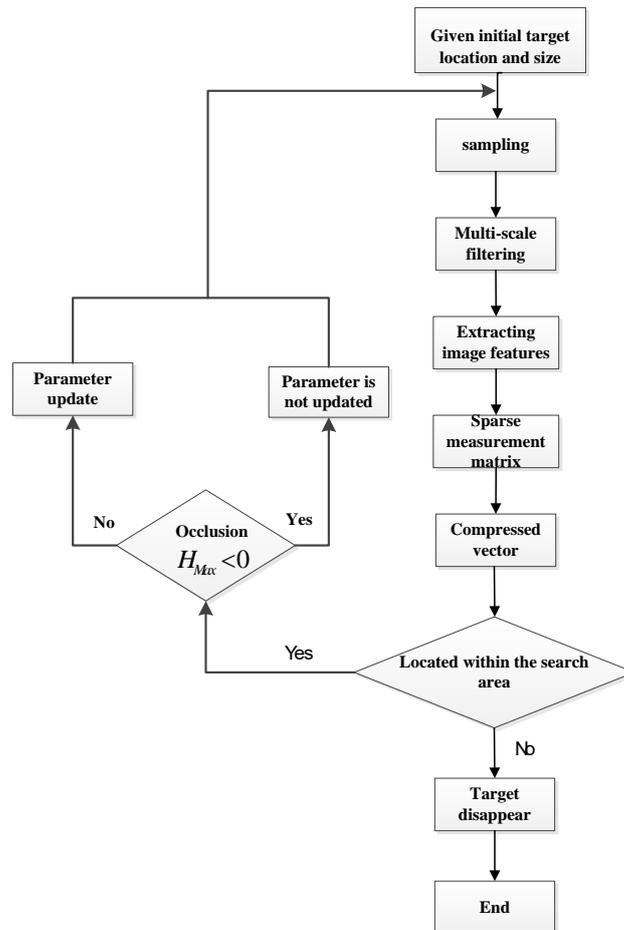


Figure 4. Algorithm Flow Chart

4. Experimental Results and Analysis

In order to verify the performance of the proposed algorithm and the traditional CT algorithm, a series of experiments were carried out to verify the performance of the tracking target, and compared with the CT algorithm. Video sequence contains some effects often appear in the tracking process, for example, shaded with large scale, light illumination, fast motion problems. Algorithms operate in Windows 7 64 bit system and under the environment of MATLAB 2010a. The accepted benchmark video database in the video sequence is verified, listed below the face, suv, car scale three typical video sequences are described, as shown in Figure 5-7, which the tracking performance of the solid line represents the algorithm, dashed line represents the tracking performance of conventional CT algorithm.

Experiment 1: occlusion and posture change

Face video sequence as shown in Figure 5 is the tracking effect of human face detection, a total of 812th frames. In the tracking process, the occurrence of a small amplitude posture transform, face is obscured. From the video sequence can be seen in the tracking process at first, two algorithms remain relatively stable tracking effect. In the 492th frame by most facial occlusion, drift phenomenon occurred two algorithms, then the traditional CT algorithm model is updated, the error parameters continue to accumulate, so in the 522th frame when the occlusion is over, tracking the target again, this algorithm does not update the appearance model, still can better track targets, and by CT algorithm blindly update the result error information accumulated, the final target tracking failure.

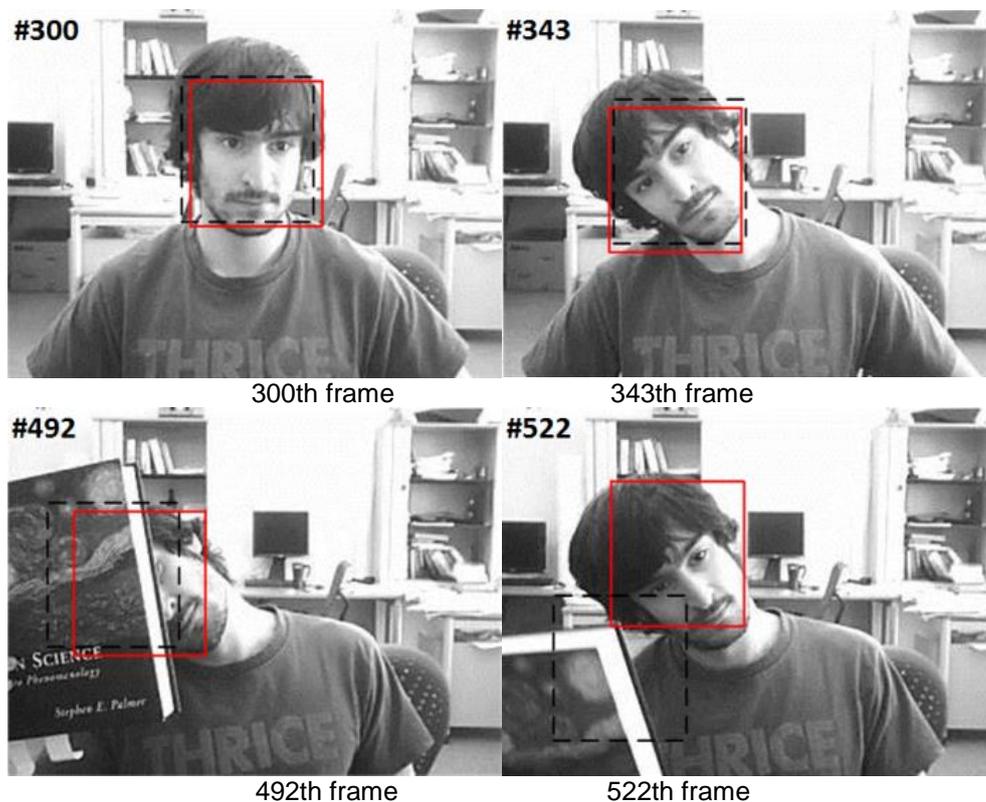


Figure 5. Tracking Effect of Face Video Sequence

Experiment 2: illumination and texture changes

As Figure 6 shows the SUV video sequence is the verification in the tracking effect of light and texture changes were. After the 240th frames, light from dark to bright and the car began changing, traditional CT algorithms only use one kind of feature tracking. Therefore, there are a drift phenomenon, 275th frame due to the cumulative error model resulted in complete with lost goal. Algorithm in this paper due to the use of gray and texture feature fusion can still continue to follow.

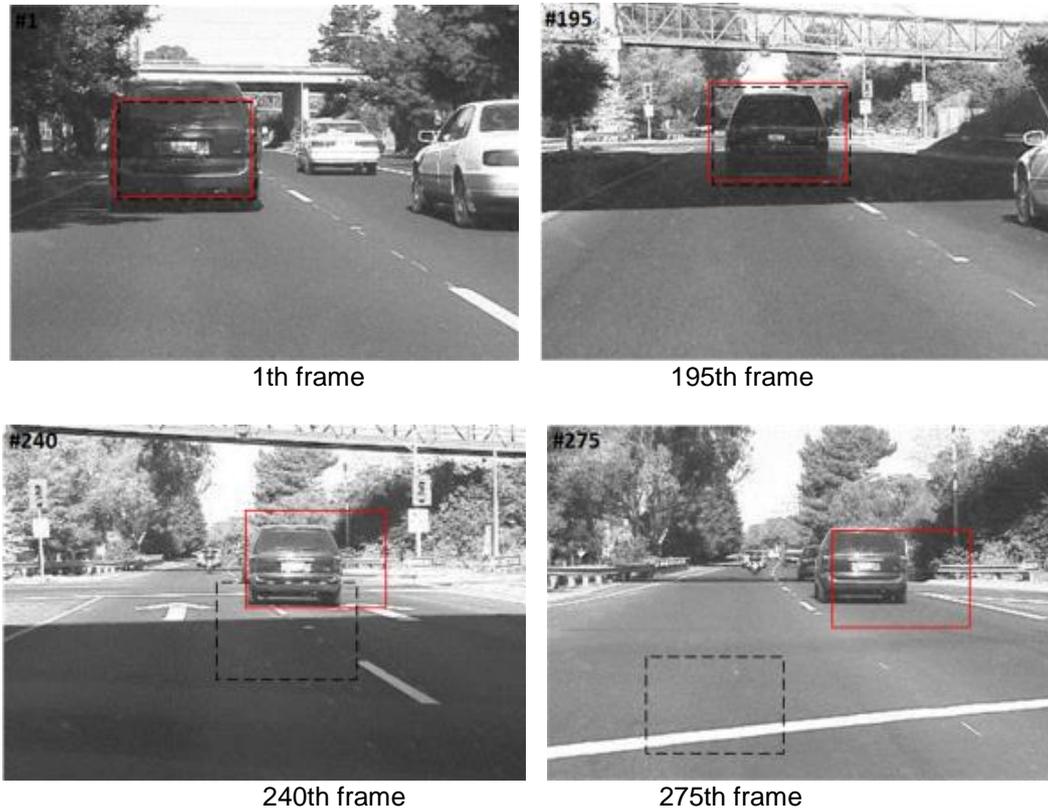


Figure 6. Tracking Effect of SUV Video Sequence

Experiment 3: fast motion and occlusion

As shown in Figure 7 car scale video sequence is a fast moving test target tracking effect. When partial occlusion occurred in 154th frame, CT algorithm shift phenomenon and the errors introduced by the appearance model, 173th frame can be seen in the traditional CT algorithm has the background information into the model (tracking areas appear in the tree, and not the front, resulting in unstable tracking). This algorithm is also capable of tracking. But in the 173th frame after the target to start turning, appearance also changed dramatically, because the two algorithms are not to deal with the problem that the target size, the size of the tracking window always keep unchanged, directly affect the positive and negative samples collected, along with the video sequence the second half of the appearance of the car is more and more small, the two algorithm finally occurred with the lost phenomenon.



5th frame



123th frame



134th frame



173th frame

Figure 7. Tracking Effect of Scale Car Video Sequence

For the quantitative evaluation of the experimental results, two evaluation indexes are introduced, namely the success rate and the average error:

$$S = \frac{|R_t \cap R_a|}{|R_t \cup R_a|} \quad (13)$$

Here R_t is the target of the actual area; R_a is the algorithm of the tracking area; if $S > 0.5$, you can think of tracking success.

If a sequence of video frames for n , the i frame tracking to the target center coordinates for \hat{x} , \hat{y} , and standard coordinates of the center of the target is x, y . Tracking error is defined as:

$$E(i) = \sqrt{(\hat{x} - x)^2 + (\hat{y} - y)^2} \quad (14)$$

Average error is

$$\overline{E(i)} = \sum_{i=1}^n E(i) / n \quad (15)$$

The formula (15) shows that the average error is smaller, and the tracking accuracy of the whole video sequence is higher.

Table 1 shows the success rate and the average error data. Compared with the traditional CT algorithm, the algorithm in dealing with occlusion and illumination, the performance of texture changes are superior to the traditional algorithm. The experimental statistics, this algorithm in tracking success rates improve 17.84%, the average error rate reduces 11.59%.

Table 1. Evaluation Index

Video Sequence	Tracking success rate (%)			Average error rate (%)		
	Traditional	This paper	↑	Traditional	This paper	↓
face	73.38	87.69	14.13	13.19	8.56	4.63
suv	44.34	71.49	27.15	41.23	16.84	24.39
Car scale	52.07	64.30	12.23	51.19	45.43	5.76
Statistics			17.84			11.59

5. Conclusion

This paper in compressed sensing tracking algorithm based on the proposed in in the process of tracking target occlusion, illumination and texture changes of solution. After the classic test video sequence experiments. The results show that this algorithm in the event of occlusion and texture, light illumination changes, the tracking evaluation index is significantly improved. But in some complex tracking scene, fast moving targets and severe deformation, the algorithm needs further improvement, but also the future direction of research.

References

- [1] DONOHO DL. Compressed sensing [J]. IEEE Transaction on Information Theories. (2006), 52(4): 1289-1306.
- [2] Candes E, Tao T. Near optimal signal recovery from random projections and universal encoding strategies [J]. IEEE Transactions on Information Theory(S0018-9448), (2006), 52: 5406-5425.
- [3] WANG S, LU HCH YANG F, *et al.* Superpixel tracking[C]. Computer Vision(ICCV), (2011): 1323-1330.

- [4] Zhang K H, Zhang L, Yang M H. Real-time compressive tracking [C] //Proc of the 12th European conference on Computer Vision. Berlin: Springer-Verlag, **(2012)**: 864-877.
- [5] Achlioptas D. Database-friendly random projections: Johnson-Lindenstrauss with binary coins [J]. Journal of Computer and System Sciences(S0022-0000), **(2003)**, 66: 671-687.
- [6] Baraniuk R, Davenport M, DeVore R, *et al.* Wakin M. A simple proof of the restricted isometry property for random matrices[J]. Constructive Approximation(S0176-4276), **(2008)**, 28 253-263.
- [7] ORON S, AHARON B H, LEVI D *et al.* Locally orderless tracking [C]. Computer Vision and Pattern Recognition, IEEE Computer Society Conference, **(2012)**: 1940-1947.
- [8] Ng A, Jordan M. On discriminative vs. generative classifier: a comparison of logistic regression and naïve bayes [J]. Neural Information Processing Systems(S2249-7110), **(2002)**, 52: 841-848.
- [9] Diaconis P, Freedman D. Asymptotics of graphical projection pursuit [J]. The Annals of Statistics(S0090-5364), **(1984)**, 12(3): 228-235.

Author



Name: Songtao Jiang

Gender: Male

Date of birth: 1990.08.31

Native place: Tianjin

Graduate School: Tianjin University of Technology

Professional: Control Engineering

Telephone: 13602027951

E-mail: 446707284@qq.com