Predictive visual tracking based on least absolute deviation estimation

Rongtai Cai (蔡荣太)[1,2] and Yanjie Wang (王延杰)[1]

[1] *Image Processing Department, Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences, Changchun 130033*

[2] *Graduate University of Chinese Academy of Sciences, Beijing 100039*

Received April 28, 2007

To cope with the occlusion and intersection between targets and the environment, location prediction is employed in the visual tracking system. Target trace is fitted by sliding subsection polynomials based on least absolute deviation (LAD) estimation, and the future location of target is predicted with the fitted trace. Experiment results show that the proposed location prediction algorithm based on LAD estimation has significant robustness advantages over least square (LS) estimation, and it is more effective than LS-based methods in visual tracking.

OCIS codes: 330.4150, 100.0100, 120.0120, 150.0150.

Visual tracking[1,2] has a wide variety of applications ranging from missile interception to visual monitoring. To cope with the occlusion and intersection of target, trace prediction is employed in the visual tracking system. Predictive target tracking is the process of estimating the location of the target at a given time in the future based upon the current tracking state and the assumed motion model[3]. There are many classical and novel predictive target tracking algorithms such as trace fitting based prediction algorithm[4], Kalman filter[3], particle filter[5], and so on.

The location prediction based on trace fitting is one of the simplest methods and has found successful implementations in engineering. The location prediction scheme based on trace fitting follows two steps.

Step 1: Using the latest m locations of the target to fit the trace. Considering the continuity of location, velocity and acceleration of the target, lower order polynomials are usually employed to fit the trace. A first-in, first-out (FIFO) register is prepared for the polynomial fitting. The FIFO renews and stores the latest m data at every frame during the tracking process.

The polynomial fitting based on least square (LS) criteria is to find the polynomial parameters $a_0, a_1, \cdots, a_n$, such that

$$x(k) = a_0 + a_1 k + a_2 k^2 + \cdots + a_n k^n + \varepsilon_k, \tag{1}$$

and to satisfy the minimization of the L_2 norm

$$\min_{a_i} \sum_{k=k_0}^{k_0-m+1} |\varepsilon_k|^2 = \sum_{k=k_0}^{k_0-m+1} |x(k) - \sum_{i=0}^{n} a_i k^i|^2, \tag{2}$$

where $x(k)$ is the location of the target, k denotes sampling time (frame), $k = k_0, (k_0-1), \cdots, (k_0-m+1)$, k_0 denotes the current frame, ε_k is the deviation of $x(k)$ from the polynomial.

Step 2: Using the polynomial produced from step 1 to estimate the location of the target in the next frame. The location estimation of the target in the (k_0+1) frame according to the polynomial fitted in step 1 is

$$\hat{x}(k_0+1) = a_0 + a_1(k_0+1) + \cdots + a_n(k_0+1)^n. \tag{3}$$

The problem with using the LS estimation is that it is sensitive to outliers. Contrarily, least absolute deviation (LAD) method is known to have significant robustness advantages over LS methods[6]. Inspired by substitution of LAD method for LS method in the solution of the linear equations in computational science[7], the polynomial fitting based on LAD estimation is proposed in this paper. The polynomial fitting based on LAD is to find $a_0, a_1, \cdots, a_n$, according to Eq. (1), and to satisfy the minimization of the L_1 norm

$$\min_{a_i} \sum_{k=k_0}^{k_0-m+1} |\varepsilon_k| = \sum_{k=k_0}^{k_0-m+1} |x(k) - \sum_{i=0}^{n} a_i k^i|. \tag{4}$$

Unlike LS methods[4], the calculation of LAD estimation is usually transformed into a constrained optimization problem[6]. Let $u_k = \frac{1}{2}(|\varepsilon_k| + \varepsilon_k) \geq 0$, $v_k = \frac{1}{2}(|\varepsilon_k| - \varepsilon_k) \geq 0$, then $|\varepsilon_k| = u_k + v_k$, $\varepsilon_k = u_k - v_k$, $k = k_0, (k_0-1), \cdots, (k_0-m+1)$. Let $a_i = b_i - c_i$, $i = 0, 1, 2, \cdots, n$. And the LAD problem is transformed into a linear programming (LP) problem:

$$\min_{a_i} \sum_{k=k_0}^{k_0-m+1} (u_k + v_k),$$

$$\text{s.t. } x(k) = [\sum_{i=0}^{n} (b_i - c_i) k^i] + u_k - v_k, \tag{5}$$

where $k = k_0, (k_0-1), \cdots, (k_0-m+1)$.

The LP problem can be solved by simplex algorithm or the recently appearing inner point algorithm[8]. Considering computational advantages, LAD estimators using simplex algorithms yield timings that are quite competitive with LS methods on problems of moderate size. Theoretically, the LAD-computations based on inner point algorithm can be made strictly faster than LS-computations for problems with the size sufficiently large[8].

It is well known that LAD-computation is much robustness than LS-computation, especially there are outliers. As an illustration, we present a comparing experiment of curve fitting using LS estimation and LAD estimation respectively, see Fig. 1. The moving trace of the target is described by

$$x(t) = 0.04 - 0.035t + 0.03t^2 + 0.005t^3,$$
$$t = 0, 1, \cdots, 15, \quad (6)$$

with a Gaussian noise governed by $N(0, 0.5)$ and an outlier in the 8th sampling time. It is obvious in Fig. 1 that the LS estimator is sensitive to the outlier, and fails to fit the source data, but the LAD estimator is robust to the outlier and fits the source data very well.

We employed a set of real data in the predictive visual tracking experiment, which is the target positions captured by an optical theodolite. The position data of target are composed of horizontal data x and vertical data y. We performed the experiment on both x and y with LS and LAD methods respectively. The 1st, 2nd and 3rd order polynomials were used to fit the trace of the target. For each order of polynomial, $3-8$ data items were used to fit the trace respectively.

Experiment results show that 3 is the best number to fit the 1st order polynomial, 6 is the best number to fit the 2nd order polynomial, and 8 is the best number to fit the 3rd order polynomial. The experiment results with the best data numbers are listed in Table 1, where n is the order of the polynomial, p is the number of data items to fit the polynomial, mean is the average absolute error of the prediction, std is the standard deviation of absolute error, max is the maximum of the absolute errors. It is obvious that the LAD-based prediction algorithm performs better than the LS-based prediction algorithm in most cases. For example, the average absolute errors in y of LAD-based prediction are strictly smaller than those of the LS-based prediction. The predictive target tracking results in this experiment using the 1st order polynomial fitted by 3 data items at every sampling time are shown in Fig. 2. The maximum absolute relative prediction error in x is less than 0.3077%, and that in y is less than 0.3636%. It is obviously seen from Fig. 2 that the proposed prediction algorithm based on LAD estimation is effective and robust in visual tracking.

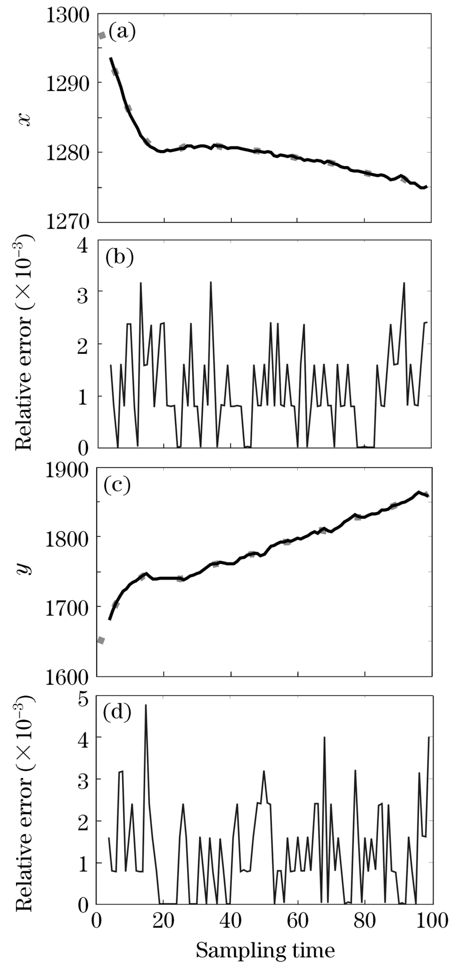

Fig. 2. Predictive visual tracking results of LAD-based method. (a) Prediction results of x; (b) prediction errors of x; (c) prediction results of y; (d) prediction errors of y. In (a) and (c), the solid curves are prediction results, and the dashed curves are real values.

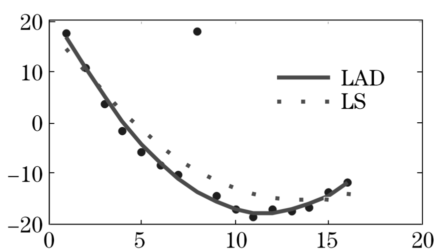

Fig. 1. Comparison of curve fitting based on LAD and LS methods. Dots represent source data.

Table 1. Comparison of Prediction Performance of LS-Based and LAD-Based Methods

	n	p	Method	Mean	Std	Max
	1	3	LS	1.1660	0.8828	3.7100
			LAD	1.1336	0.8273	3.1950
x	2	6	LS	1.2377	1.0343	5.2550
			LAD	1.2401	0.9772	3.8420
	3	8	LS	1.5528	1.3251	6.9564
			LAD	1.5830	1.3143	5.9717
	1	3	LS	1.2940	1.1308	4.7733
			LAD	1.2169	1.0891	4.7750
y	2	6	LS	1.7011	1.2806	6.2260
			LAD	1.4869	1.2389	6.0800
	3	8	LS	1.9673	1.6647	8.3479
			LAD	1.7249	1.5521	7.6674

In conclusion, a LAD-based target tracking prediction algorithm is proposed. The trace of target is fitted under the criteria of L_1 norm minimization first, and the location of the target in the next frame is predicted using the fitted trace. Experiment results show that the proposed method is more effective and robust than LS-based algorithm in predictive visual tracking. Because of the availability and ease of computability, the proposed algorithm will find wide applications in real-time engineering.

This work was supported by the National High Technology Research and Development Program of China under Grant No. 2006AA703405F. R. Cai's e-mail address is gjrtcai@163.com.

References

1. S. Benhimane and E. Malis, The International Journal of Robotics Research **26,** 661 (2007).
2. J. Wang, J. Ren, Y. You, and G. Liu, Acta Opt. Sin. (in Chinese) **14,** 523 (1994).
3. K. A. Kramer and S. C. Stubberud, in *Proceedings of the 2nd International Conference on Intelligent Sensors, Sensor Networks and Information Processing* 199 (2005).
4. S.-J. Zhang and Y.-Z. Liu, Opto-Electronic Eng. (in Chinese) **30,** 24 (2003).
5. B. Zhang, W. Tian, and Z. Jin, Chin. Opt. Lett. **4,** 569 (2006).
6. S. Portnoy and R. Koenker, Statistical Sci. **12,** 279 (1997).
7. J. B. Rosen, H. Park, J. Glick, and L. Zhang, Computational Optimization and Applications **17,** 329 (2000).
8. M. H. Wright, Acta Numerica **1,** 341 (1992).