

L_0 -Regularized Object Representation for Visual Tracking

Jinshan Pan¹

sdluran@gmail.com

Jongwoo Lim²

jlim@hanyang.ac.kr

Zhixun Su¹

zxsu@dlut.edu.cn

Ming-Hsuan Yang³

mhyang@ucmerced.edu

¹ School of Mathematical Sciences

Dalian University of Technology Dalian, China

² Division of Computer Science & Engineering

Hanyang University

Seoul, Korea

³ Electrical Engineering and Computer Science

University of California at Merced

California, USA

Introduction & Motivation: Visual tracking is a highly researched topic in the computer vision community since it has been widely applied in visual surveillance, driver assistant system, and many others. Although much progress has been made in the past decades, designing a practical visual tracking system is still a challenging problem due to numerous challenges in real world.

Very recent efforts have been made to improve this method in terms of both speed and accuracy by using APG algorithm [1] or modeling the similarity between different candidates [6]. The works in [4, 5] point out that the aforementioned methods do not exploit rich and redundant image properties which can be captured compactly with subspace representations. Thus, they propose combining the strength of subspace learning [3] and sparse representation for modeling object appearance. In their work the raw pixel templates used in [1, 2] are replaced with the orthogonal basis vectors (e.g., PCA basis), and the coefficients for an image are obtained by a least square (LS) method. However, we empirically find that such linear combination of the orthogonal basis vectors sometimes include redundant parts (e.g., background portions), which will interfere with the accuracy of object representation.

We in this paper address this problem by proposing a tracking method based on an L_0 regularized object representation. The L_0 regularized object representation is able to reduce the redundant features while keeping the most important part. Furthermore, the estimation of the L_0 regularized parameters can be efficiently conducted by the proposed APG algorithm.

L_0 Regularized Object Representation: We assume that the target region $\mathbf{y} \in \mathbb{R}^{d \times 1}$ can be represented by an image subspace with corruption,

$$\mathbf{y} = D\boldsymbol{\alpha} + \mathbf{e}, \quad (1)$$

where the columns of $D \in \mathbb{R}^{d \times n}$ are orthogonal basis vectors of the subspace, $\boldsymbol{\alpha}$ is the sparse coefficient vector, and \mathbf{e} represents additive errors modeled by a Laplacian noise.

We propose an L_0 regularized prior to select useful features, which is defined as

$$\min_{\boldsymbol{\alpha}, \mathbf{e}} \frac{1}{2} \|\mathbf{y} - D\boldsymbol{\alpha} - \mathbf{e}\|_2^2 + \lambda \|\mathbf{e}\|_1 + \gamma \|\boldsymbol{\alpha}\|_0, \quad (2)$$

where $D^\top D = I$, $\|\cdot\|_0$ denotes the L_0 norm which counts the number of non-zero elements, $\|\cdot\|_2$ and $\|\cdot\|_1$ denote L_2 and L_1 norms, respectively, γ and λ are regularization parameters, and I is an identity matrix. The term $\|\mathbf{e}\|_1$ is used to reject outliers (e.g., occlusions), while $\|\boldsymbol{\alpha}\|_0$ is used to select the useful features. We note that if we set $\gamma = 0$, (2) is reduced to [4].

Analysis on the Effectiveness of L_0 Representation: The benefit of the L_0 norm regularized prior is that it is able to reduce the redundant features while keeping the most important part, thereby facilitating the tracking result.

When there are no errors (e.g., occlusion) in the observation \mathbf{y} , i.e., $\mathbf{e} \approx 0$, we can think of L_p regularized error metric in general,

$$\min_{\boldsymbol{\alpha}} \frac{1}{2} \|\mathbf{y} - D\boldsymbol{\alpha}\|_2^2 + \gamma \|\boldsymbol{\alpha}\|_p^p, \quad \text{where } D^\top D = I, \quad (3)$$

and the solutions for different p are given in the following theorem.

Theorem 1 Assume that $D \in \mathbb{R}^{d \times d}$ and $D^\top D = I$. The solution of (3) when p is 0 is given by

$$\boldsymbol{\alpha} = H_{2\gamma}(D^\top \mathbf{y}), \quad (4)$$

when p is 1, the solution is

$$\boldsymbol{\alpha} = S_\gamma(D^\top \mathbf{y}), \quad (5)$$

and when p is 2, the solution becomes

$$\boldsymbol{\alpha} = \frac{D^\top \mathbf{y}}{1 + 2\gamma}. \quad (6)$$

Here $S_\theta(x) = \text{sign}(x) \max(|x| - \theta, 0)$, and $H_\theta(x)$ is a hard thresholding operator, which is defined as $H_\theta(x) = x$; if $x^2 > \theta$ and 0 otherwise.

Based on Theorem 1, we have the following corollary.

Corollary 1 We assume D is redundant and contains all possible basis. Let \mathbf{u}^* denote the non-zero elements of $D^\top \mathbf{y}$. If we set $\gamma = \frac{1}{2} \min_i \{|\mathbf{u}_i^*|^2\}$, the solution of L_0 regularized error metric (i.e., (3) when $p = 0$) can exactly recover the data \mathbf{y} .

Figure 1 shows the tracking results by using LS method [4] (i.e., $\gamma = 0$ in (2)), L_0 and L_2 norm under the same dictionary D , respectively. We note that using L_0 regularized method is able to find the good candidate when there exists occlusion, then facilitating the tracking results.

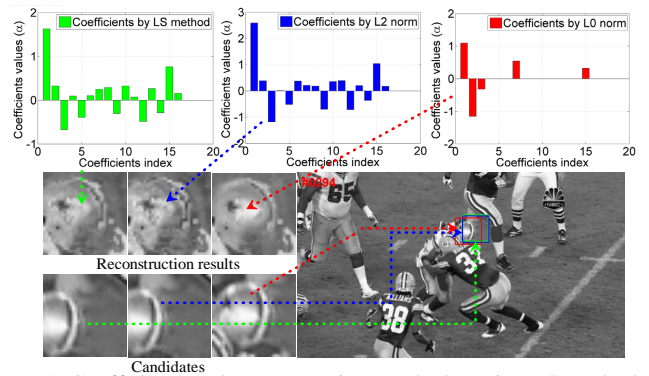


Figure 1: Coefficients and reconstruction results by using LS method, L_0 and L_2 norm under the same dictionary D , respectively. The coefficients by using L_0 norm are more sparse than those by L_2 norm and LS method, and the reconstruction result and the best candidate are also better. The rectangles in the last image represent MAP states for particle filters.

- [1] Chenglong Bao, Yi Wu, Haibin Ling, and Hui Ji. Real time robust ℓ_1 tracker using accelerated proximal gradient approach. In *CVPR*, pages 1830–1837, 2012.
- [2] Xue Mei and Haibin Ling. Robust visual tracking using ℓ_1 minimization. In *ICCV*, pages 1436–1443, 2009.
- [3] David A. Ross, Jongwoo Lim, Ruei-Sung Lin, and Ming-Hsuan Yang. Incremental learning for robust visual tracking. *IJCV*, 77(1-3): 125–141, 2008.
- [4] Dong Wang, Huchuan Lu, and Ming-Hsuan Yang. Least soft-threshold squares tracking. In *CVPR*, pages 2371–2378, 2013.
- [5] Dong Wang, Huchuan Lu, and Ming-Hsuan Yang. Online object tracking with sparse prototypes. *IEEE TIP*, 22(1):314–325, 2013.
- [6] Tianzhu Zhang, Bernard Ghanem, Si Liu, and Narendra Ahuja. Robust visual tracking via structured multi-task sparse learning. *IJCV*, 101(2):367–383, 2013.