ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

Robust Object Tracking using compressive sparse vectors

Pranathi Shirga
ME,Student,
Stanley College of Engineers & Technology for Women, Hyderabad.

Mrs.Udayini Chandana, MS(USA)
Asst Professor,
Stanley College of Engineers & Technology for Women, Hyderabad.

Abstract:

It is a challenging task to develop effective and efficient appearance models for robust object tracking due to factors such as pose variation, illumination change, occlusion, and motion blur. Existing online tracking algorithms often update models with samples from observations in recent frames. Despite much success has been demonstrated, numerous issues remain to be addressed. First, while these adaptive appearance models are data-dependent, there does not exist sufficient amount of data for online algorithms to learn at the outset. Second, online tracking algorithms often encounter the drift problems. As a result of self-taught learning, misaligned samples are likely to be added and degrade the appearance models. In this paper, we propose a simple yet effective and efficient tracking algorithm with an appearance model based on features extracted from a multiscale image feature space with data-independent basis. The proposed appearance model employs non-adaptive random projections that preserve the structure of the image feature space of objects. A very sparse measurement matrix is constructed to efficiently extract the features for the appearance model. We compress sample images of the foreground target and the background using the same sparse measurement matrix. The tracking task is formulated as a binary classification via a naive Bayes classifier with online update in the compressed domain. A coarse-to-fine search strategy is adopted to further reduce the computational complexity in the detection procedure. The proposed compressive tracking algorithm runs in real-time and performs favorably against state-of-the-art methods on challenging sequences in terms of efficiency, accuracy and robustness.

I.Introduction:

Despite that numerous algorithms have been proposed in the literature, object tracking remains a challenging problem due to appearance change caused by pose, illumination, occlusion, and motion, among others.

An effective appearance model is of prime importance for the success of a tracking algorithm that has attracted much attention in recent years [2]–[15]. Numerous effective representation schemes have been proposed for robust object tracking in recent years. One commonly adopted approach is to learn a low-dimensional subspace, which can adapt online to object appearance change. Since this approach is data-dependent, the computational complexity is likely to increase significantly because it needs eigen-decompositions. Moreover, the noisy or misaligned samples are likely to degrade the subspace basis, thereby causing these algorithms to drift away the target objects gradually. Another successful approach is to extract discriminative features from a high-dimensional space.Since object tracking can be posed as a binary classification task which separates object from its local background, a discriminative appearance model plays an important role for its success. Online boosting methods [6], [10] have been proposed to extract discriminative features for object tracking. Alternatively, high-dimensional features can be projected to a lowdimensional space from which a classifier can be constructed. The compressive sensing (CS) theory, shows that if the dimension of the feature space is sufficiently high, these features can be projected to a randomly chosen low-dimensional space which contains enough information to reconstruct the original high-dimensional features. The dimensionality reduction method via random projection (RP) ,is data-independent, non-adaptive and information-preserving. In this paper, we propose an effective and efficient tracking algorithm with an appearance model based on features extracted in the compressed domain [1]. The main components of the proposed compressive tracking algorithm are shown by Figure 1. . We use a very sparse measurement matrix that asymptotically satisfies the restricted isometry property (RIP) in compressive sensing theory, thereby facilitating efficient projection from the image feature space to a low-dimensional compressed subspace.For tracking, the positive and negative samples are projected (i.e., compressed) with the same sparse measurement matrix

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

and discriminated by a simple naive Bayes classifier learned online. The proposed compressive tracking algorithm runs at real-time and performs favorably against state-of-the-art trackers on challenging sequences in terms of efficiency, accuracy and robustness.

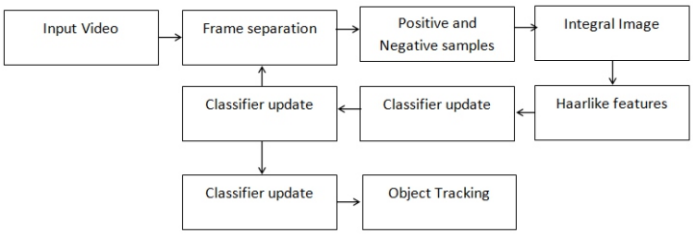


Fig. 1: Main components of the proposed compressive tracking algorithm.

II.PRELIMINARIES:

We present some preliminaries of compressive sensing which are used in the proposed tracking algorithm.

2.1 Random projection and compressive sensing:

In random projection, a random matrix whose rows have unit length projects data from the high-dimensional feature space to a lower-dimensional space where n <<m. Each projection v is essentially equivalent to a compressive measurement in the compressive sensing encoding stage. The compressive sensing theory ,states that if a signal is K-sparse (i.e., the signal is a linear combination of only K basis), it is possible to near perfectly reconstruct the signal from a small number of random measurements. The encoder in compressive sensing (using (1)) correlates signal with noise (using random matrix R) , thereby it is a universal encoding which requires no prior knowledge of the signal structure. In this paper, we adopt this encoder to construct the appearance model for visual tracking.

Ideally, we expect R provides a stable embedding that approximately preserves the salient information in any K- sparse signal when projecting from to . A necessary and sufficient condition for this stable embedding is that it approximately preserves distances between any pairs of K-sparse signals that share the same K basis. That is, for any two K-sparse vectors x1 and x2 sharing the same K basis,

$$(1-\varepsilon)\|x_1 - x_2\|_{l_2}^2 \le \|Rx_1 - Rx_2\|_{l2}^2 \le (1+\varepsilon)\|x_1 - x_2\|_{l2}^2 \quad (2)$$

The restricted isometry property ,in compressive sensing shows the above results. This property is achieved with high probability for some types of random matrix R whose entries are identically and independently sampled from a standard normal distribution, symmetric Bernoulli distribution or Fourier matrix. Furthermore, the above result can be directly obtained from the Johnson-Lindenstrauss (JL) lemma.

2.2 Very sparse random measurement matrix:

A typical measurement matrix satisfying the restricted isometry property is the random Gaussian matrix where (i.e., zero mean and unit variance), as used in recent work [11]. However, as the matrix is dense, the memory and computational loads are very expensive when m is large. In this paper, we adopt a very sparse randommeasurement matrix with entries defined as

$$r_{ij} = \sqrt{\rho} \times \begin{cases} 1 & with\ probability\ \frac{1}{2\rho} \\ 0 & with\ probability\ 1 - \frac{1}{\rho} \\ -1 & with\ probability\ \frac{1}{2\rho} \end{cases} \quad (3)$$

III.PROPOSED ALGORITHM:

In this section, we present the proposed compressive tracking algorithm in details. The tracking problem is formulated as a detection task and the main steps of the proposed algorithm are shown in Figure 1. We assume that the tracking window in the first frame is given by a detector or manual label.

At each frame, we sample some positive samples near the current target location and negative samples away from the object center to update the classifier. To predict the object location in the next frame, we draw some samples around the current target location and determine the one with the maximal classification score.

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

3.1 Image representation:

To account for large scale change of object appearance, a multiscale image representation is often formed by convolving the input image with a Gaussian filter of different spatial variances. The Gaussian filters in practice have to be truncated which can be replaced by rectangle filters. Bay et al. show that this replacement does not affect the performance of the interest point detectors but can significantly speed up the detectors via integral image method. For each sample , its multiscale representation is constructed by convolving Z with a set of rectangle filters at multiple scales defined by

$$F_{w,h}(x,y) = \frac{1}{wh} \times \begin{cases} 1 & 1 \le x \le w, 1 \le y \le h \\ 0 & otherwise \end{cases} \quad (4)$$

where w and h are the width and height of a rectangle filter, respectively. Then, we represent each filtered image as a column vector in and concatenate these vectors as a very high-dimensional multiscale image feature vector where .The dimensionality m is typically in the order of to . We adopt a sparse random matrix R in (7) to project x onto a vector in a low-dimensional space. The random matrix R needs to be computed only once offline and remains fixed throughout the tracking process. For the

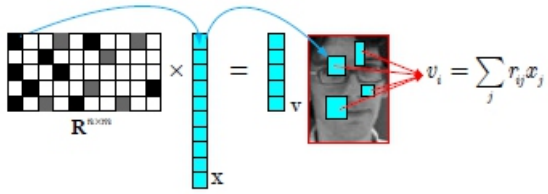


Fig. 2: Graphical representation of compressing a high-dimensional vector x to a low-dimensional vector v. In the matrix R, dark, gray and white rectangles represent negative, positive, and zero entries, respectively. The blue arrows illustrate that one of nonzero entries of one row of R sensing an element in x is equivalent to a rectangle filter convolving the intensity at a fixed position of an input image.sparse matrix R in (7), the computational load is very light. As shown in Figure 2, we only need to store the nonzero entries in R and the positions of rectangle filters in an input image corresponding to the nonzero entries in each row of R. Then, v can be efficiently computed by using R to sparsely measure the rectangular features which can be efficiently computed using the integral image method.

3.2 Analysis of compressive features:

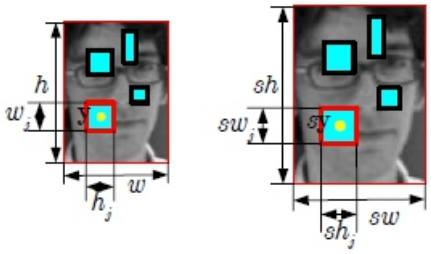
3.2.1 Relationship to the Haar-like features:

Each element vi in the low-dimensional feature is a linear combination of spatially distributed rectangle features at different scales. Since the coefficients in the measurement matrix can be positive or negative (via (7)), the compressive features compute the relative intensity difference in a way similar to the generalized Haar-like features [10]. The Haar-like features have been widely used for object detection with demonstrated success [10]. The basic types of these Haar-like features are typically designed for different tasks. There often exist a very large number of Haar-like features which make the computational load very heavy.

This problem is alleviated by boosting algorithms for selecting important features. Recently, Babenko et al. [10] adopt the generalized Haar-like features where each one is a linear combination of randomly generated rectangle features, and use online boosting to select a small set of them for object tracking. In this work, the large set of Haar-like features are compressively sensed with a very sparse measurement matrix. The compressive sensing theories ensure that the extracted features of our algorithm preserve almost all the information of the original image, and hence is able to correctly classify any test image because the dimension of the feature space is sufficiently large (106 to1010). Therefore, the projected features can be classified in the compressed domain efficiently and effectively without the curse of dimensionality.

3.2.2 Scale invariant property:

It is easy to show that the low-dimensional feature v is scale invariant. As shown in Figure 2, each feature in v is a linear combination of some rectangle filters convolving the input image at different positions. Therefore, without loss of



ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

Fig. 3: Illustration of scale invariant property of low-dimensional features. From the left figure to the right one, the ratio is s. Red rectangle represents the j-th rectangle feature at position y.

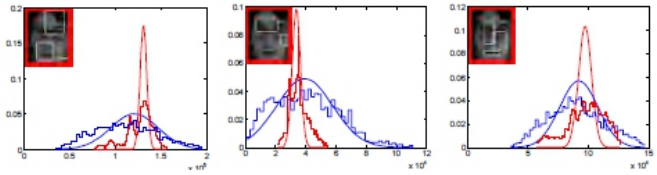


Fig. 4: Probability distributions of three different features in a lowdimensional space. The red stair represents the histogram of positive samples while the blue one represents the histogram of negative samples. The red and blue lines denote the corresponding estimated distributions by the proposed incremental update method.generality, we only need to show that the j-th rectangle feature in the i-th feature in v is scale invariant. From Figure 3, we have

$$
\begin{aligned}
&= F_{sw_j,sh_j}(sy) \otimes Z(sy) \\
&= F_{sw_j,sh_j}(a) \otimes Z(a)|_{a=sy} \\
&= \frac{1}{s^2 w_i h_i} \int_{u\in\Omega_s} Z(a-u)du \\
x_j(sy) &= \frac{1}{s^2 w_i h_i} \int_{u\in\Omega_s} Z(y-u)|s^2|du \\
&= \frac{1}{w_i h_i} \int_{u\in\Omega_s} Z(y-u)du \\
&= F_{w_j,h_{j(y)}}(y) \otimes Z(y) \\
&= x_j(y),
\end{aligned} \tag{5}
$$

Where $\Omega = \{(u_1,u_2) \mid 1 \le u_1 \le w_i, 1 \le u_2 \le h_i\}$

And $\Omega_s = \{(u_1,u_2) \mid 1 \le u_1 \le sw_i, 1 \le u_2 \le sh_i\}$

IV.Classifier construction and update:

We assume all elements in v are independently distributed and model them with a naive Bayes classifier,

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

$$
= \sum_{i=1}^{n} \log\left(\frac{p(v_i \mid y=1)}{p(v_i \mid y=0)}\right) \tag{6}
$$

where we assume uniform prior, p(y = 1) = p(y = 0), and is a binary variable which represents the sample label.

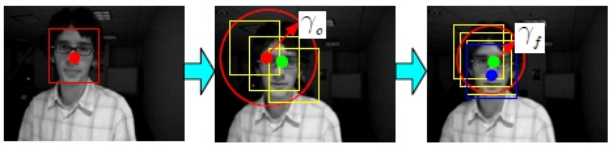


Fig. 5: Coarse-to-fine search for new object location. Left: object center location (denoted by red solid circle) at the t-th frame. Middle: coarse-grained search with a large radius and search step based on the previous object location. Right: fine-grained search with a small radius and search step based on the coarse grained search location (denoted by green solid circle). The final object location is denoted by blue solid circle. Diaconis and Freedman show that random projections of high dimensional random vectors are almost always Gaussian. Thus, the conditional distributions and in the classifier H(v) are assumed to be Gaussian distributed with four parameters ,

$$
P(v_i \mid y=1) \sim N(\mu_i^1, \sigma_i^1), \quad P(v_i \mid y=0) \sim N(\mu_i^0, \sigma_i^0) \tag{10}
$$

Where $\mu_i^1(\mu_i^0)$ and $\sigma_i^1(\sigma_i^0)$ are mean and standard deviation of the Positive (negative) class. The scalar parameters in (10) are incrementally Updated by

$$
\mu_i^1 \leftarrow \lambda\mu_i^1 + (1-\lambda)\mu^1 \text{ Learning parameter,}
$$

$$
\sigma^1 = \sqrt{\frac{1}{n}\sum_{k=0|y=1}^{n-1}(v_i(k)-\mu^1)^2} \text{ And } \mu^1 = \frac{1}{n}\sum_{k=0|y=1}^{n-1} v_i(k)
$$

Parameters and are updated with similar rules. Theabove equations can be easily derived by maximum likelihoodestimation . Figure 3 shows the probability distributionsfor three different features of the positive and negative samplescropped from a few frames of a sequence for clarity ofpresentation. It shows that a Gaussian distribution with onlineupdate using (11) is a good approximation of the features inthe projected space where samples can be easily separated. Because the variables are assumed to be independent in ourclassifier, the n-dimensional multivariate problem is reducedto the n univariate estimation problem. Thus, it requires fewertraining samples to obtain accurate estimation than estimatingthe covariance matrix in the multivariate estimation. Furthermore,several densely sampled positive samples surroundingthe current tracking result are used to update the distributionparameters, which is able to obtain robust estimation evenwhen the tracking result has some drift.

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

In addition, the usefulinformation from the former accurate samples is also usedto update the parameter distributions, thereby facilitating theproposed algorithm to be robust to misaligned samples. Thus,our classifier performs robustly even when the misaligned orthe insufficient numbers of training samples are used.

V.Experimental results::

The proposed algorithm is termed as fast compressive tracker (FCT) with one fixed scale, and scaled FCT (SFCT), with multiple scales in order to distinguish from the compressive tracker (CT) proposed by our conference[1]. The FCT and SFCT methods demonstrate superior performance over the CT method in terms of accuracy and efficiency, which validates the effectiveness of the scale invariant features and coarse-to-fine search strategy. Furthermore, the proposed algorithm is evaluated with other 15 state-of-the-art methods on 20 challenging sequences among which 14 are publicly available and 6 are collected on our own.

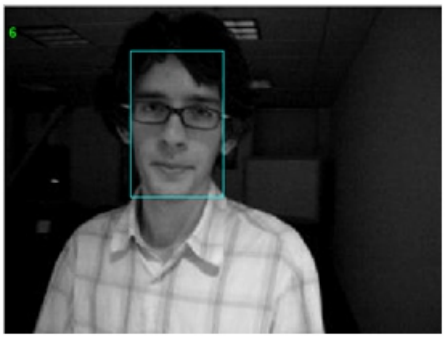


Fig 6: Result for tracking using proposed method.

VI.Conclusion:

In this paper, we propose a simple yet robust tracking algorithm with an appearance model based on non-adaptive random projections that preserve the structure of original image space. A very sparse measurement matrix is adopted to efficiently compress features from the foreground targets and background ones.

The tracking task is formulated as a binary classification problem with online update in the compressed domain. Numerous experiments with state-ofthe- art algorithms on challenging sequences demonstrate that the proposed algorithm performs well in terms of accuracy, robustness, and speed.

Our future work will focus on applications of the developed algorithm for object detection and recognition under heavy occlusion. In addition, we will explore efficient detection modules for persistent tracking.

VII.References:

1)K. Zhang, L. Zhang, and M. Yang, "Real-time compressive tracking," in Proceedings of European Conference on Computer Vision, pp. 864–877, 2012. 1, 8.

2)M. Black and A. Jepson, "Eigentracking: Robust matching and tracking of articulated objects using a view-based representation," International Journal of Computer Vision, vol. 26, no. 1, pp. 63–84, 1998. 1, 2, 7.

3)A. Jepson, D. Fleet, and T. El-Maraghi, "Robust online appearance models for visual tracking," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 25, no. 10, pp. 1296–1311, 2003. 1, 2.

4)S. Avidan, "Support vector tracking," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 26, no. 8, pp. 1064–1072, 2004. 1, 2.

5)R. Collins, Y. Liu, and M. Leordeanu, "Online selection of discriminative tracking features," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 27, no. 10, pp. 1631–1643, 2005. 1, 2.

6)H. Grabner, M. Grabner, and H. Bischof, "Real-time tracking via online boosting," in Proceedings of British Machine Vision Conference, pp. 47–56, 2006. 1, 2, 6, 7, 8.

7)D. Ross, J. Lim, R. Lin, and M. Yang, "Incremental learning for robust visual tracking," International Journal of Computer Vision, vol. 77, no. 1, pp. 125–141, 2008. 1, 2, 7, 8.

8)H. Grabner, C. Leistner, and H. Bischof, "Semi-supervised on-line boosting for robust tracking," in Proceedings of European Conference on Computer Vision, pp. 234–247, 2008. 1, 2, 8.

9)J. Kwon and K. Lee, "Visual tracking decomposition.," in Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, pp. 1269–1276, 2010. 1, 2, 8.

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

10)B. Babenko, M. Yang, and S. Belongie, "Robust object tracking with online multiple instance learning," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 33, no. 8, pp. 1619–1632, 2011. 1, 2, 4, 6, 7, 8, 10.

11)H. Li, C. Shen, and Q. Shi, "Real-time visual tracking using compressive sensing," in Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, pp. 1305–1312, 2011. 1, 2, 3, 6, 8.

12)X. Mei and H. Ling, "Robust visual tracking and vehicle classification via sparse representation," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 33, no. 11, pp. 2259–2272, 2011. 1, 2, 6, 8.

13)S. Hare, A. Saffari, and P. Torr, "Struck: Structured output tracking with kernels," in Proceedings of the IEEE International Conference on Computer Vision, pp. 263–270, 2011. 1, 2, 3, 8.

14)J. Kwon and K. M. Lee, "Tracking by sampling trackers," in Proceedings of the IEEE International Conference on Computer Vision, pp. 1195– 1202, 2011. 1.

15)J. Henriques, R. Caseiro, P. Martins, and J. Batista, "Exploiting the circulant structure of tracking-by-detection with kernels," in Proceedings of European Conference on Computer Vision, pp. 702–715, 2012. 1, 2, 3, 8.