2014 22nd International Conference on Pattern Recognition

Locality-constrained Sparse Reconstruction for Trajectory Classification

Ce Li[1], Zhenjun Han[1], Qixiang Ye[1,2], Shan Gao[1], Lijin Pang[1], Jianbin Jiao[1*]
[1]University of Chinese Academy of Sciences, Beijing, China.
[2]Institute for Advanced Computer Studies, University of Maryland, College Park
[2]qxye@umiacs.umd.edu, [1*]jiaojb@ucas.ac.cn

***Abstract*—Trajectory classification has been extensively investigated in recent years; however, problems remain when processing incomplete trajectories of noises and local variations. In this paper, we propose a Locality-constrained Sparse Reconstruction (LSR) approach that explores both sparsity and local adaptability for robust trajectory classification. A trajectory dictionary with locality constrains is constructed with tracklets partitioned from collected trajectories by control points of cubic B-spline curves. On the dictionary, the proposed LSR is used to calculate a discriminate code matrix. Then, a loss weighted decoding strategy is employed to perform multi-class trajectory classification. In addition, the approach can be used for anomalous trajectory detection with a thresholding strategy. Experiments on two datasets show that the results of the LSR approach improve the state of the art.**

I. Introduction

Visual trajectory classification is an important research topic in computer vision with applications in video analysis, understanding and retrieval [1][2][3]. Many approaches use trajectory classification as a tool to understand and characterize the object behaviors in video scenes. However, trajectory classification is still an open research problem with challenges from the variation of trajectory length [4], the trajectory noises [5], the local variation of the trajectories [6] and the limited sizes of sample sets [7]. It is required to explore more effective trajectory representation and classification approaches.

For trajectory representation [8], motion vectors, positions, velocities, acceleration information etc. are commonly used as feature vectors. For trajectories of variable length, re-sampling and linear interpolation strategies [9] are typically used to align the feature vectors, while function approximation [10][11] could be used to improve the local adaptability of the representation.

In [10], an efficient trajectory representation is proposed by using function approximation algorithms of least square polynomial, Cheybyshev polynomial and Discrete Fourier Transform (DFT). In [11], Haar wavelet coefficients and Least-square Cubic Spline Curves Approximation (LCSCA) are adopted to represent a trajectories. LCSCA holds better fidelity to the original trajectory and is less sensitive to the trajectory length, for the use of the least-square fitting procedure. Despite the simplicity of these approximation approaches, however, they often require fine turned function parameters to be adaptive to various trajectory noises and local trajectory variations.

On aligned trajectory sequences, both unsupervised and supervised learning methods are widely used for trajectory

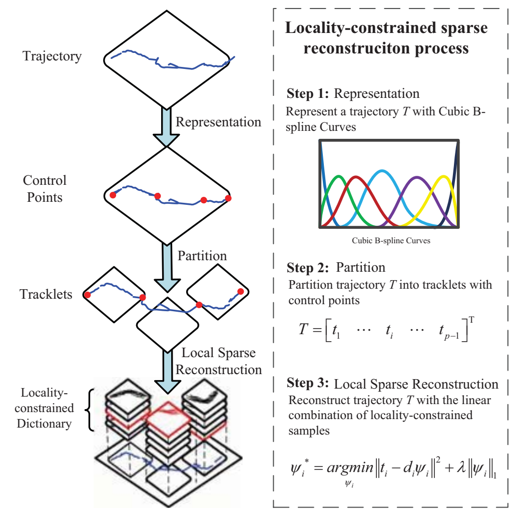


Fig. 1. Illustration of locality-constrained sparse recontruction

classification. In [9][12], self-organizing mapping and hierarchical clustering are used to learn prototypes. Test trajectories are classified by their distances, e.g. the Euclidean distance, the Hausdorff distance for Dynamic Time Warping (DTW) [4], to the learned prototypes. Supervised learning methods such as Gaussian Mixture Models (GMMs) [7], Bayesian Model [4][13], Hidden Markov Model (HMM) [14], One-Class Support Vector Machine (OC-SVM) [1], Hierarchical Hidden Markov Model [14] and hierarchical Bayesian models [13], are employed in trajectory classification. In general, these methods have good performance on large sample sets; however, the performance may drop significantly on small training sets.

Recently, the trajectory classification has been casted as a sparse reconstruction problem [15]. Intuition behind the sparse reconstruction lies in the fact that a test trajectory could be reconstructed by a sparse linear combination of a few samples in a trajectory dictionary, and reconstruction coefficients contains discriminative information that is be very effective to trajectory classification [16][17]. The sparse reconstruction based method improves the classification performance on small sample sets. However, the adopted global sparse reconstruction strategy is still challenged when incomplete trajectories of local variations.

In this paper, we propose a Locality-constrained Sparse Reconstruction (LSR) approach for robust trajectory classifi-

1051-4651/14 $31.00 © 2014 IEEE
DOI 10.1109/ICPR.2014.449

cation. It is based on the observation that locality is often as important as sparsity in trajectory representation [18]. Locality enables LSR be adaptive to trajectory noises, incompleteness and local variation, and sparsity guarantees that LSR has a good performance on small training sets. A trajectory can be seen as a union of partitioned tracklets (trajectory segments), as shown in Fig. 1, and the local tracklets from the similar global trajectories often hold the similar local shapes. Given a set of tracklets, we construct a dictionary, on which a discriminate code matrix is then constructed. The code-words in the matrix are assigned according to the LSR, which guarantees that the tracklets with local-nonzero coefficients in the dictionary are often similar to the tracklet for classification. With the code matrix, a loss weighted decoding strategy is employed to perform trajectory classification. In addition, by thresholding the reconstruction degree, the LSR approach is also used for anomalous trajectory detection. The contributions of this paper are summarized as follows:

- We propose Locality-constrained Sparse Reconstruction (LSR) approach;
- We apply the proposed approach to perform trajectory classification and anomalous trajectory detection.

The remainder of the paper is organized as follows. In section II, the LSR approach is introduced in detail. In section III, we describe the trajectory classification and anomalous trajectory detection procedures. Experimental results are presented in section IV and we conclude the paper in section V.

II. Locality-constrained Sparse Reconstruction

In this section, we first present the trajectory representation, on which we describe the trajectory partition and dictionary construction strategies. We then describe how to perform sparse linear reconstruction with a locality-constrained dictionary.

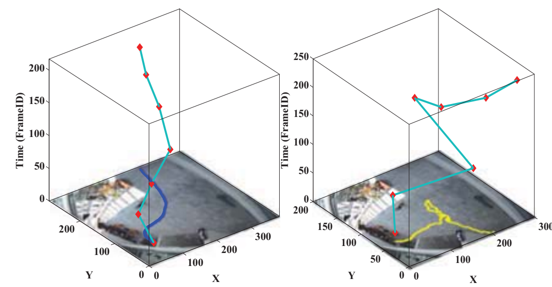
A. Trajectory Representation

For variable length trajectories, we use control points of cubic B-spline curves to extract a fixed-length parametric vector as trajectory representation [15]. This is done by approximating each spatial-temporal trajectory with a uniform cubic B-spline curve parameterized by time.

Given a trajectory in spatial and temporal space (x, y, t), we use B-spline control points to represent both its shape profile $T = \{(x_1, y_1), (x_2, y_2), \cdots, (x_{n-1}, y_{n-1}), (x_n, y_n)\}$ in a parametric way $F = \{C_1^x, C_2^x, \cdots, C_p^x, C_1^y, C_2^y, \cdots C_p^y\}$, where p is the number of control points and n is the length of trajectory, C_p^x is the normalized x-coordinate of p-th control point, and C_p^y is the normalized y-coordinate of p-th control point. Fig. 2 shows a normal and an anomalous trajectory both in spatial-temporal space and parametric space, respectively.

B. Trajectory Partition

The objective of trajectory partition is to divide a long trajectory into tracklets. Accordingly, a long feature vector is divided into a set of short sub-vectors, each of them will be more well represented by a linear model.



(a) (b)

Fig. 2. Trajectory representation with p=7. (a)control points (red dots) of a normal trajectory (blue curve). (b)control points (red dots) of the anomalous trajectory (yellow curve). (Better view in color version)

Supposing there are J trajectory classes (called routes) in a video scene, and each route $route_j = \{a_j^1, a_j^2, ..., a_j^K\}$ holds K trajectories, we have a sample set D in the scene as:

$$D = \cup\{route_j\} = \{a_1^1, a_1^2, ..., a_1^K, ..., a_J^K\}, j = 1, ..., J. \quad (1)$$

The trajectory partition is based on the fact that trajectories of a same category should often have similar shapes and share control points on the cubic B-spline curves. We partition the trajectories into tracklets (trajectory segments) based on the control points and then align the tracklets via the DTW algorithm to construct a local dictionary, as

$$D = [d_1 \cdots d_i \cdots d_{p-1}]^T, \quad i = 1, \ldots, p-1, \quad (2)$$

where $d_i = \{a_1^1(i), a_1^2(i), ..., a_1^K(i), ..., a_J^K(i)\}$ represents the i-th tracklets of all trajectories in a scene after partition. In Fig. 3, we show a class of similar trajectories and their partitioned tracklets.

C. Local Sparse Reconstruction

After the trajectory partition, a trajectory T is represented as

$$T = \left[t_1 \cdots t_i \cdots t_{p-1}\right]^T, \quad i = 1, \ldots, p-1, \quad (3)$$

where t_i is the feature vector of the i-th tracklet. Each tracklet can be approximately represented as a linear superposition of the d_i in the dictionary as follows:

$$t_i \approx d_i \psi_i, \quad (4)$$

where ψ_i represents a coefficient vector for superposition. Although the aforementioned model could be more complicated, we assume a linear function for efficiency and simplicity. For so many trajectory routes in a scene, the dictionary is often large, so the coefficient vector ψ_i should be sparse. This can be computed by optimizing the l_1 regularized least square problem in (5), which typically provides a sparse solution [15] as

$$\psi_i^* = \underset{\psi_i}{argmin} \left\|t_i - d_i\psi_i\right\|^2 + \lambda \left\|\psi_i\right\|_1, \quad (5)$$

where the regularization parameter λ is used to restrain the sparsity. Because most elements of the coefficient vector ψ_i are zero, a tracklet can be represented with as few as tracklets of local similarity, called local sparse reconstruction.

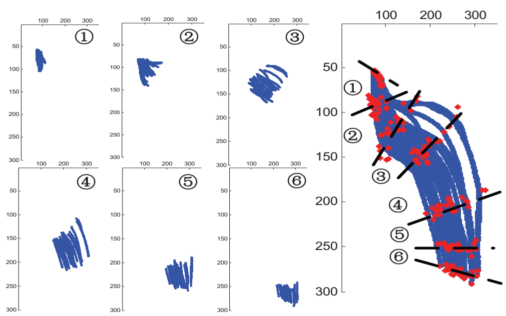


Fig. 3. Left: six partitioned tracklets from a class of similar trajectories (blue curves). Right: the whole trajectories with control points (red dots).

III. Trajectory Classification

With the proposed LSR, we further propose a discriminate encoding together with a loss weighted decoding strategy for classification. We also use a thresholding strategy for anomalous trajectory detection.

A. Discriminate Encoding

Tracklets represent local shapes of trajectories, however, they lose the global information of trajectories. Therefore, a combination of reconstruction results from multiple tracklets is required for accurate trajectory classification. We propose to use a discriminate code matrix $M \in \{0,1\}^{N_p \times N_c}$ to integrate local sparse reconstruction results and pose locality. N_p is the number of tracklets in a trajectory, and N_c is the length of the code-words (equal to J).

For classification of a test trajectory T, the key is to calculate M_{ij} with LSR under the hypothesis that the i-th tracklet can be approximately represented as a linear superposition of the i-th tracklet d_i in the dictionary. We define a characteristic function δ_j for each class, which keeps the sparse coefficients corresponding to j-th class in d_i and sets the coefficients corresponding to other classes to zero. And the sparse reconstruction residual for each tracklet can be calculated with:

$$\varepsilon_{ij}(t_i) = \left\| t_i - d_i \delta_j(\psi_i^*) \right\|_2, \; j = 1, \ldots, N_c. \tag{6}$$

Element M_{ij} that corresponds to assign the i-th tracklet to the j-th class is calculated as

$$M_{ij} = \begin{cases} 0, & if\ j \neq \underset{j}{argmin}(\varepsilon_{ij}) \\ 1, & if\ j = \underset{j}{argmin}(\varepsilon_{ij}) \end{cases}, \; j = 1, \ldots, N_c. \tag{7}$$

It should be noted that there is only one nonzero element in each row of M, that is, each tracklet can only belong to exactly one class. An example of M is shown in TABLE I.

B. Loss Weighted Decoding

With the code matrix, a trajectory is classified by using a loss-based decoding strategy [19]. The objective is to find a matrix W that weights a loss function and adjusts the decisions of the sparse reconstruction. The loss weighted decoding process is described in Algorithm 1, where the weighted matrix W is normalized in step 5, the linear loss function $L(\varepsilon_{ij}) = \varepsilon_{ij}$ is applied in step 8. The algorithm returns the class label of a test trajectory in step 10. It is effective to make a decision by assigning a label to a trajectory according to class with minimal decoding measure. The algorithm is detailed as follows.

TABLE I. An Example of Discriminate Code Matrix

Code	Class					
	Class 1	*Class 2*	*Class 3*	*Class 4*	…	*Class* N_c
Tracklet 1	0	0	1	0	…	0
Tracklet 2	1	0	0	0	…	0
Tracklet 3	0	0	0	1	…	0
Tracklet 4	0	0	0	0	…	1
…	…	…	…	…	…	…
Tracklet N_p	0	0	0	1	…	0

Algorithm 1 Trajectory classification

Input: Discriminate code matrix M
Output: Class label j^* of the test trajectory T

```
1:  Initialize each item w_ij = MAX of W
2:  for j = 1 to N_c do
3:    for i = 1 to N_p do
4:      if M_ij ≠ 0 and Σ_{j=1}^{N_p} M_ij ≠ 0 then
5:        w_ij ← M_ij / (Σ_{j=1}^{N_p} M_ij)
6:      end if
7:    end for
8:    e_j ← Σ_{i=1}^{N_p} w_ij L(ε_ij)
9:  end for
10: return j* ← argmin_j (e_j)
```

C. Anomalous Trajectory Detection

The LSR approach can also be used for anomalous trajectory detection using a dictionary constructed with only normal trajectories. In this case, for each tracklet of T, we calculate a reconstruction degree

$$H_i = \underset{j}{min} \frac{1/\varepsilon_{ij}}{\sum_{j=1}^{N_c} (1/\varepsilon_{ij})}, \tag{8}$$

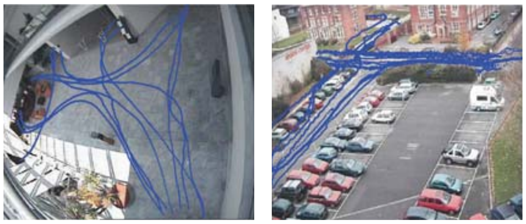
which represents how well T is reconstructed on the i-th tracklet, and classify T as a normal trajectory or an anomalous trajectory by

$$F = \begin{cases} 0, & if\ \sum_{i=1}^{N_p} sign(H_i - \theta) H_i < 0 \\ 1, & otherwise \end{cases}, \tag{9}$$

where θ is an empirically determined threshold (0.03). When the value of $\sum_{i=1}^{N_p} sign(H_i - \theta) H_i$ is smaller than zero, the test trajectory is classified as an outlier of the normal trajectories samples, and is detected as an anomalous trajectory with $F = 0$.

IV. Experimental Results

We performed experiments on two public benchmarks: the CAVIAR ("INRIA") dataset [20] and the Carpark dataset [21]. Lots of the trajectories in the datasets are of noise and local variation. In experiments, we follow [15] to use seven control points (p=7) for trajectory partition. 10% of the selected trajectories are cut manually to simulate incomplete trajectories. Trajectory examples are shown in Fig. 4.



(a) (b)

Fig. 4. Trajectory examples from (a) the CAVIAR dataset and (b) the Carpark dataset.

A. Trajectory Classification

The CAVIAR dataset [20] contains a series of trajectories in an entrance lobby. There are 22 classes of normal classes of trajectories, each of which has 100 simulated trajectories. So, the total number of trajectories that can be used to construct the dictionary is up to 2200. There are 21 test trajectories corresponds to people walking directly from one exit to another. In the Carpark dataset there are 269 trajectories in a car park scene from 16 trajectory classes. There are 27 test trajectories in this dataset. Correct classification rate (*CCR*) is used to evaluate the performance [15].

The LSR based approach is tested with respect to different percentages of training trajectories, and some classification results are shown in Fig. 5 and Fig. 6. In each sub-figure, the trajectory in deep color can be classified by LSR with those trajectories in light color. In Fig. 7, we compare our proposed locality-constrained sparse reconstruction approach with the sparse reconstruction approach [15]. It can be seen that the proposed LSR approach significantly outperforms the sparse reconstruction approach when using more than 60% trajectories samples to construct the dictionary. In particular, the LSR approach shows significant higher performance than the sparse reconstruction approach on incomplete data (red dashed line). On the CAVIAR dataset of incomplete trajectories, when using 60% trajectories samples to construct a dictionary, the *CCR* of the LSR approach is about 20% higher than that of sparse reconstruction approach. With the same settings on the Carpark dataset, the *CCR* of the LSR approach is about 10% higher than that of sparse reconstruction approach. This validates the effectiveness of the proposed locality-constrained sparse reconstruction, in particulary, on the incomplete trajectories.

B. Anomalous Trajectory Detection

Anomalous Correction Accuracy (*ACC*) is defined to measure the proportion of correctly classified normal and anomalous trajectories. In the experiment, we use 19 anomalous trajectories corresponding to people fighting, falling down, leaving or collecting packages in the CAVIAR dataset, and 7 anomalous trajectories corresponding to car turning at forbidden areas in the Carpark dataset. The experiment is performed on both complete trajectories and incomplete trajectories. A GMMs based approach [7] and a sparse reconstruction approach [15] are used for comparison.

In TABLE II, it can be seen that the performance of our LSR approach is very impressive. In the CAVIAR dataset, the our approach has the highest performance (*ACC*=93.21%) when using all samples on annotated trajectories. It also has the highest performance (*ACC*=89.5%) on incomplete trajectory. On the CAVIAR dataset, our approach has significant performance improvement over two recent approaches (the GMMs and sparse reconstruction approach). On the Carpark dataset, our approach also has performance improvement. Some detected anomalous trajectories are shown in Fig. 8.

(a) (b)

Fig. 8. Detected anomalous trajectory examples from (a) the CAVIAR dataset and (b) the Carpark dataset.

TABLE II. COMPARISONS OF ANOMALOUS DETECTION

Dataset		Method	Accuracy(%)
CAVIAR	Annotated	GMMs based [7]	84.75
		sparse reconstruction [15]	87.42
		Our Approach	**93.21**
	Incomplete	GMMs based [7]	74.00
		sparse reconstruction [15]	83.72
		Our Approach	**89.50**
Carpark	Annotated	GMMs based [7]	84.60
		sparse reconstruction [15]	87.32
		Our Approach	**89.75**
	Incomplete	GMMs based [7]	79.41
		sparse reconstruction [15]	85.29
		Our Approach	**88.24**

V. CONCLUSION

We have proposed a locality-constrained sparse reconstruction (LSR) approach for trajectory classification in surveillance video scenes. The proposed approach utilizes both the locality and sparsity to represent trajectories on a constructed dictionary set. A discriminative encoding strategy together with a loss weighted decoding process is used to classify the trajectories into different categories. We can also detect anomalous trajectories with a dictionary constructed on normal trajectories and a threshholding strategy. Experimental results on two public datasets show the good performance of our LSR approach. Comparisons with two recent approaches are also provided, which indicates that the LSR approach have significant advantages. In addition, this is achieved on small trajectory sample sets.

ACKNOLEGEMENT

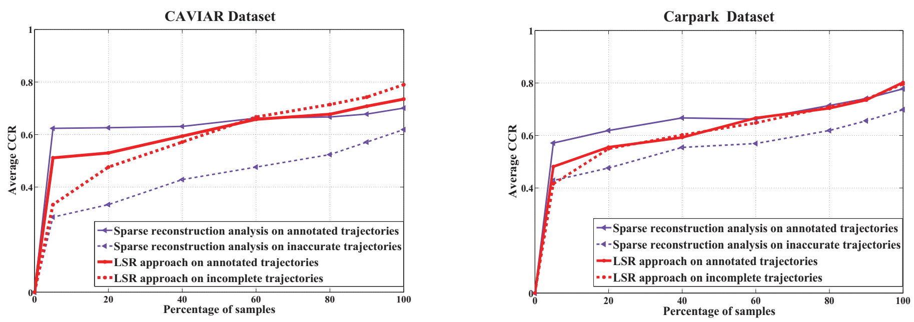
This work is supported in part by National Basic Research Program of China (973 Program) with Nos. 2011CB706900, 2010CB731800, and National Science Foundation of China with Nos. 61039003, 61271433 and 61202323.

REFERENCES

[1] B. T. Morris and M. M. Trivedi, “Trajectory learning for activity understanding: Unsupervised, multilevel, and long-term adaptive approach,” *IEEE Trans. PAMI*, vol. 33, no. 11, pp. 2287–2301, 2011.

Fig. 5. Examples of correctly classified testing trajectories in the CAVIAR dataset. Different colors represent different classes. (Better view in color version)

Fig. 6. Examples of correctly classified testing trajectories in the Carpark dataset. Different colors represent different classes. (Better view in color version))



(a) (b)

Fig. 7. Performance and comparision on two testsets of (a) the CAVIAR dataset, and (b) the Carpark dataset.

[2] T. Xiang and S. Gong, "Video behavior profiling for anomaly detection," *IEEE Trans. PAMI*, vol. 30, no. 5, pp. 893–908, 2008.

[3] V. Mahadevan, W. Li, V. Bhalodia, and N. Vasconcelos, "Anomaly detection in crowded scenes," *Proc. IEEE Conf. CVPR*, pp. 1975–1981, 2010.

[4] X. Wang, M. Keng Teck, G. Ng, and W. Grimson, "Trajectory analysis and semantic region modeling using a nonparametric bayesian models," *Proc. IEEE Conf. CVPR*, vol. 95, no. 3, pp. 287–312, 2011.

[5] J. Lee, J. Han, X. Li, and H. Gonzalez, "Traclass: trajectory classification using hierarchical region-based and trajectory-based clustering," *Proc. VLDB Endowment*, vol. 1, no. 1, pp. 1081–1094, 2008.

[6] D. Weinland, R. Ronfard, and E. Boyer, "A survey of vision-based methods for action representation, segmentation and recognition," *CVIU*, vol. 115, no. 2, pp. 224–241, 2011.

[7] R. Sillito and R. Fisher, "Semi-supervised learning for anomalous trajectory detection." *Proc. BMVC*, pp. 1035–1044, 2008.

[8] A. Adam, E. Rivlin, I. Shimshoni, and D. Reinitz, "Robust real-time unusual event detection using multiple fixed-location monitors," *IEEE Trans. PAMI*, vol. 30, no. 3, pp. 555–560, 2008.

[9] W. Hu, X. Xiao, Z. Fu, D. Xie, T. Tan, and S. Maybank, "A system for learning statistical motion patterns," *IEEE Trans. PAMI*, vol. 28, no. 9, pp. 1450–1464, 2006.

[10] A. Naftel and S. Khalid, "Classifying spatiotemporal object trajectories using unsupervised learning in the coefficient feature space," *Multimedia Systems*, vol. 12, no. 3, pp. 227–238, 2006.

[11] R. Sillito and R. Fisher, "Parametric trajectory representations for behaviour classification," *Multimedia Systems*, vol. 12, pp. 227–238, 2006.

[12] L. Kratz and K. Nishino, "Anomaly detection in extremely crowded scenes using spatio-temporal motion pattern models," *Proc. IEEE Conf. CVPR*, pp. 1446–1453, 2009.

[13] X. Wang, M. Keng Teck, G. Ng, and W. Grimson, "Trajectory analysis and semantic region modeling using nonparametric hierarchical bayesian models," *IJCV*, vol. 95, no. 3, pp. 287–312, 2011.

[14] D. Vasquez, T. Fraichard, and C. Laugier, "Growing hidden markov models: An incremental tool for learning and predicting human and vehicle motion," *International Journal of Robotics Research*, vol. 28, no. 11-12, pp. 1486–1506, 2009.

[15] C. Li, Z. Han, Q. Ye, and J. Jiao, "Visual abnormal behavior detection based on trajectory sparse reconstruction analysis," *Neurocomputing*, vol. 119, pp. 94–100, 2013.

[16] J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry, and Y. Ma, "Robust face recognition via sparse representation," *IEEE Trans. PAMI*, vol. 31, no. 2, pp. 210–227, 2009.

[17] Z. Han, J. Jiao, B. Zhang, Q. Ye, and J. Liu, "Visual object tracking via sample-based adaptive sparse representation (adasr)," *Pattern Recognition*, vol. 44, no. 9, pp. 2170–2183, 2011.

[18] J. Wang, J. Yang, K. Yu, F. Lv, T. Huang, and Y. Gong, "Locality-constrained linear coding for image classification," *Proc. IEEE Conf. CVPR*, pp. 3360–3367, 2010.

[19] S. Escalera, D. M. Tax, O. Pujol, P. Radeva, and R. P. Duin, "Subclass problem-dependent design for error-correcting output codes," *IEEE Trans. PAMI*, vol. 30, no. 6, pp. 1041–1054, 2008.

[20] CAVIAR, http://homepages.inf.ed.ac.uk/rbf/CAVIAR/.

[21] H. M. Dee and D. Hogg, "Detecting inexplicable behaviour." *Proc. BMVC*, pp. 1–10, 2004.