DYNAMIC TREE-STRUCTURED SPARSE RPCA VIA COLUMN SUBSET SELECTION FOR BACKGROUND MODELING AND FOREGROUND DETECTION

Salehe Erfanian Ebadi [*], *Valia Guerra Ones* [†], *Ebroul Izquierdo* [*]

[*] Queen Mary University of London, [†] Delft University of Technology

ABSTRACT

Video analysis often begins with background subtraction, which consists of creation of a background model that allows distinguishing foreground pixels. Recent evaluation of background subtraction techniques demonstrated that there are still considerable challenges facing these methods. Processing per-pixel basis from the background is not only time-consuming but also can dramatically affect foreground region detection, if region cohesion and contiguity is not considered in the model. We present a new method in which we regard the image sequence to be made up of the sum of a low-rank background matrix and a dynamic tree-structured sparse matrix, and solve the decomposition using our approximated Robust Principal Component Analysis method extended to handle camera motion. Furthermore, to reduce the curse of dimensionality and scale, we introduce a low-rank background modeling via Column Subset Selection that reduces the order of complexity, decreases computation time, and eliminates the huge storage need for large videos.

***Index Terms*—** Approximated RPCA, structured-sparse, column subset selection, foreground detection, moving camera.

1. INTRODUCTION

Background subtraction can be defined as segmentation of a video stream into foreground, which appears at unique moments in time, and the background which is always present. It is a basic video processing task with manifold applications. The research in this paper addresses this fundamental task using an approximated Robust Principal Component Analysis (RPCA) based method for background modeling. Given a data matrix containing the frames of a video sequence stacked as its columns, $A \in \mathbb{R}^{m \times n}$, RPCA [1] solves the matrix decomposition problem

$$\min_{L,S} \|L\|_* + \|S\|_1 \quad s.t. \quad A = L + S, \tag{1}$$

as a surrogate for the actual problem

$$\min_{L,S} rank(L) + \|S\|_0 \quad s.t. \quad A = L + S, \tag{2}$$

where L is the low-rank component corresponding to the background and S is the sparse component containing the foreground outliers. We are interested in a case where we can decompose the matrix A into three components, namely a low-rank part L that can describe the background of the sequence, along with adaptivity to changes introduced to it, a sparse component S containing only the genuine rigid foreground regions, and a noise component E that collectively contains Gaussian noise and ambiguous pixels. We formulate the problem in equation (1) as

$$\min_{L,S,\tau} \|L\|_F + \lambda\|S\|_1 \quad s.t. \quad A \circ \tau = L + S + E, \tag{3}$$

meaning that the model does not seek the exact solution of decomposing a scene into background and foreground, but rather the approximate solution $A \approx L + S$ [2], [3] whereby the residual error E will have the desired properties described above. λ is a regularizing parameter ensuring no genuine foreground regions will be missed. We have assumed that the images in matrix $A \circ \tau$ are well aligned, where τ stands for some transformation in the image domain (e.g., 2D affine transformation for correcting misalignment, or 2D projective transformation for handling some perspective change).

Background modeling by the low-rank approximation has a number of benefits: firstly, that a robust estimation of the mostly static regions of the image is guaranteed; secondly, that this approximation can in part handle the variations in illumination in the background, such as a tree swaying backwards and forward, or water rippling in a lake, or traffic light changes that can be modeled by a few modes. Thirdly, a low-rank approximation of the background can help distinguish between general motion in the scene – which can be due to camera movement – and local varying motions caused by moving objects; since the background regions obey a single highly correlated motion pattern.

Despite the promising effects of using a low-rank approximation for obtaining the background model, a sparse constraint for foreground objects, is far too limited, if not useful in practice. The foreground regions are usually spatially coherent clusters. Thus, we prefer to detect contiguous regions of various sizes, and then lots of zero entries (regions) in the sparse matrix. With this objective in mind, we propose structured-sparsity inducing norms in the context of a novel dynamic group structure, by which the natural structure of foreground objects in the sparse matrix is preserved. The dynamicity of group structures is controlled via a patch-based group selection algorithm that preserves the natural shape of objects in the scene. The size and structure of these patches (or clusters) are dynamically refined in an iterative process. The matrix A can become humongous when processing large or long videos. To alleviate the curse of dimensionality and scale with an RPCA-based problem, we must leverage on the fact that such data have in fact low intrinsic dimensionality. We approach this problem as a *Column Subset Selection Problem* (CSSP) [4], [5] by which means it is possible to select a handful of the most representative and important columns of a matrix. Assuming that we have a long video of a scene at our disposal with hundreds or even thousands of frames, only a handful of these frames determine a model of the background; the rest will either contaminate the background or will be redundant to process. To this end, we propose to model the background of the sequence using a low-rank approximation by the output of the CSSP algorithm.

In a nutshell contributions of this paper are: inducing structured-sparsity in a novel group structure, namely a dynamic block structure; insensitivity to foreground object size, as a result of using within-patch normalized regularization; assumption of a Gaussian i.i.d. noise for discarding false positive pixels (false alarms); variable rank to accommodate illumination and small scene changes; a dimensionality reduction for RPCA problem via the column subset selection that reduces computational complexity and cost; and an exhaustive evaluation using four datasets that demonstrates top performance in comparison with the state-of-the-art alternatives.

2. RELATED WORK

Recent variants of low-dimensional models such as principal component analysis (PCA) have resolved part of the issues arising with this methodology, namely a non-SVD based fast solution [6]. However, still no considerations of the spatial distribution of outliers was considered. In an effort to incorporate such prior an MRF-based solution [7] has been proposed. But the result of imposing such smoothness constraint (even with the discontinuity preserving prior such as those based on Potts model) is that the foreground regions tend to be over-smoothed; as an example, the details in the silhouette of hands and legs of a moving person is sacrificed in favor of a more compact blob. Our idea is established in the so-called structured-sparsity or group-sparsity measures to incorporate the spatial prior. Structural information about nonzero patterns of variables have been developed and used in sparse signal recovery, and many approaches have been applied to these problems successfully, such as Lattice Matching Pursuit (LaMP) [8], Dynamic Group Sparsity (DGS) recovery [9], Bayesian Robust Matrix Factorization (BRMF) [10],and the Proximal Operator using Network Flow (ProxFlow) [11]. However, the majority of related methods [12], [13] typically assume that the block structure and its location is known or will suffer in *regularization* or *bootstrapping*. In contrast, our method does not assume a prior size or location or structure for sparsity, and dynamically updates these to best fit the natural object shape in the scene, without a separate training phase. The curse of dimensionality and scale remains an open problem with RPCA-based solutions with large input size. Different strategies for the dimensionality reduction for RPCA-based methods are the bilateral random projections [6] and [14]. In the next sections we present a number of solutions for the aforementioned critical issues with RPCA based solutions.

3. ROBUST FOREGROUND DETECTION VIA STRUCTURED SPARSITY

We propose sparsity-inducing norms that can incorporate prior structures on the support of the errors such as spatial continuity. We essentially consider a special case to the following problem

$$\min_{L,S,\tau} \|L\|_F + \lambda\psi(S) \quad s.t. \quad A \circ \tau \approx L + S, \tag{4}$$

with the regularizer $\psi(\cdot)$ on S chosen to be $\|S\|_{2,1}$. Clearly, the ℓ_1-norm regularization of (3) does not take into account any specific structures or possible relations among subsets of the entries, while in background subtraction, outliers (objects in the scene) normally have structural properties of spatial contiguity and locality. Therefore the $\ell_{2,1}$ (a group sparsity inducing norm) is used to induce more diverse and sophisticated sparse error patterns, that involves overlapping groups of variables. $\psi(\cdot)$ involves a hierarchical partition of the m variables in S into groups, called trees. A tree is defined in a way that leaf nodes are singleton groups corresponding to individual pixels, and internal nodes/groups correspond to local patches of varying size. Thus each parent node contains a hierarchy of child nodes that are spatially adjacent to each other and constitute a local part in the sparse image S. When a parent node goes to zero all its descendants in the tree must go to zero. Consequently, the nonzero or support patterns are formed by removing those nodes forced to zero. This is exactly the desired effect of structured error patterns of spatial locality and contiguity.

We can represent a scene using a tree structure by subdivision. In such a tree structure each child node is a subset of its parent node and the nodes of the same depth level do not overlap. Denote $\mathcal{G}$ as a set of groups from the power set of the index set $\{1,\ldots,\mathcal{M}\}$, with each group $G \in \mathcal{G}$ containing a subset of these indices. The aforementioned tree-structured groups used in this paper are formally defined as follows: A set of groups $\mathcal{G}$ is said to be *tree-structured* in $\{1,\ldots,\mathcal{M}\}$ if $\mathcal{G} = \{\ldots, G_1^i, G_2^i, \ldots, G_{b_i}^i, \ldots\}$ where $i = 0, 1, 2, \ldots, d$, d is the depth of the tree, $b_0 = 1$ and $G_1^0 = \{1, 2, \ldots, \mathcal{M}\}$, $b_d = \mathcal{M}$ and correspondingly $\{G_j^d\}_{j=1}^{\mathcal{M}}$ are singleton groups. Let G_j^i be the parent node of a node $G_{j'}^{i+1}$ in the tree, we have $G_{j'}^{i+1} \subseteq G_j^i$. We also have $G_j^i \cap G_k^i = \emptyset, \forall i = 1, \ldots, d, j \neq k, 1 \leq j, k \leq b_i$. Similar group structures are also considered in [15] with different sparsity-inducing norms. With the above notation, a general tree-structured sparsity-inducing norm can be written as

$$\psi(S) = \sum_{i=0}^{d} \sum_{j=1}^{b_i} w_j^i \|S_{G_j^i}\|_{2,1}, \tag{5}$$

where $S_{G_j^i}$ is a vector with entries equal to those of S for the indices in G_j^i and 0 otherwise. w_j^i are positive weights for groups G_j^i. Here it is chosen as $w_j^i = 1/max(A_{G_j^i})$ to overcome sensitivity of the regularization scheme to illumination variance across patches. Thus the objective function in the optimization program (3) is modified to

$$\min_{L,S,\tau} \|L\|_F + \lambda \sum_{i=0}^{d} \sum_{j=1}^{b_i} w_j^i \|S_{G_j^i}\|_{2,1} \quad s.t. \quad A \circ \tau \approx L + S, \tag{6}$$

where λ is a parameter controlling the trade-off between sparsity of $S + E$ and structured sparsity of S. To solve this problem we use an alternating minimization procedure. We proceed by minimizing the function for two parameters L and S one at a time until the solution reaches convergence; that means solving two reduced problems, each being minimized independently form one another

$$L^t = \underset{rank(L) \leq r}{\arg\min} \|A \circ \tau - L - S^{t-1}\|_F^2 \tag{7}$$

$$S^t = \arg\min_S \|A \circ \tau - L^t - S\|_F^2 + \lambda \sum_{i=0}^{d} \sum_{j=1}^{b_i} w_j^i \|S_{G_j^i}\|_{2,1} \tag{8}$$

We find a good initialization for τ by pre-aligning all frames in the sequences to the middle frame, before the main loops of minimization. The linearization of τ and pre-alignment is done by the robust multiresolution method proposed in [2], [16].

4. DIMENSIONALITY REDUCTION FOR DECOMPOSITION WITH CSSP

We propose a novel dimensionality reduction technique that calculates the background model from a handful of the "best" or "most representative" columns from a matrix. The theoretical computer science community has come up with randomized [17] and deterministic [14], [18] algorithms to solve this problem [5]. The CSSP is defined as: Let $A \in \mathbb{R}^{m \times n}$ and let $c < n$ be a sampling parameter. Find c columns of A – denoted as $C \in \mathbb{R}^{m \times c}$ – that minimize

$$\|A - CC^{\dagger}A\|_F \quad or \quad \|A - CC^{\dagger}A\|_2,$$

where $C^{\dagger}$ denotes the Moore-Penrose pseudo-inverse. We can equivalently write $C = A\mathcal{A}$, where the *sampling matrix* is $\mathcal{A} \in \mathbb{R}^{n \times c}$. A simple but extremely successful deterministic strategy is proposed [19] which is based on sampling columns of A that correspond to the largest leverage scores ℓ_i^{κ}, for some $\kappa < rank(A)$. As the number of columns to be selected is not known a priori, the algorithm selects the c columns of A that correspond to the largest c leverage scores ℓ_i^{κ} such that their sum $\sum_{i=1}^{c} \ell_i^{\kappa}$ is more than an "energy" parameter θ that essentially controls the quality of the approximation, with $c = \theta \times n$. The rank-κ leverage score of the i-th column of A is defined as

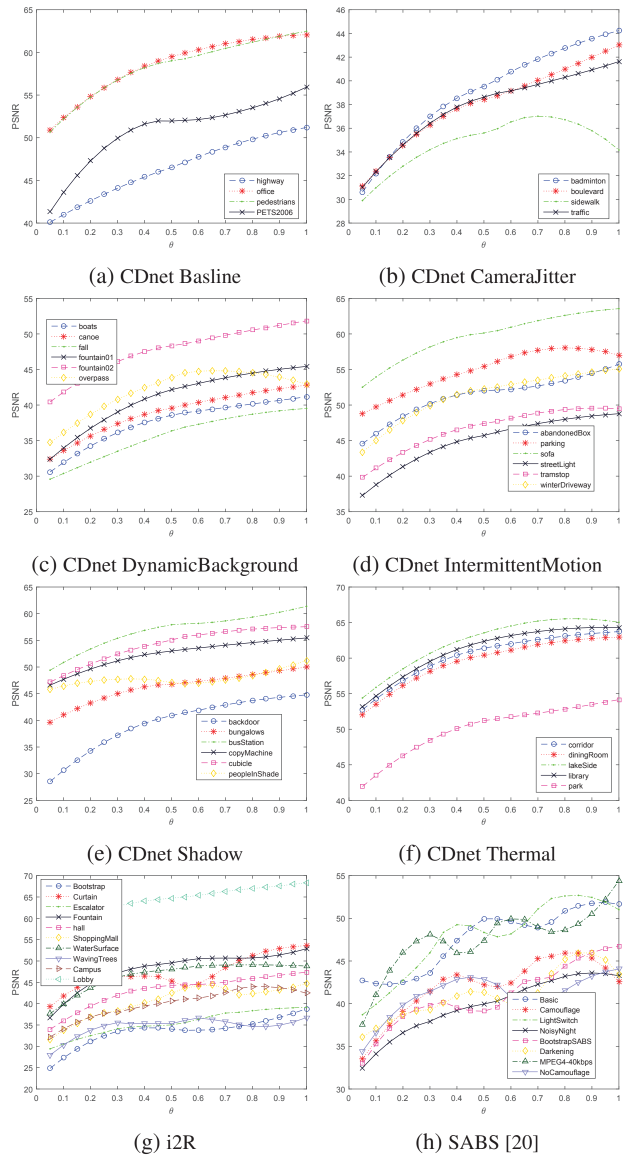
$$\ell_i^{\kappa} = \|V_{\kappa}(i,:)\|_2^2, \quad i = 1, 2, \ldots, n$$

Here, $V_\kappa(i,:)$ denotes the i-th row of V_κ, a matrix that contains the right singular vectors. A more sophisticated method that circumvents the lack of theoretical analysis of the above deterministic algorithm, uses randomization; the leverage scores are used to find a probability vector $\xi_i = \ell_i^\kappa/\kappa, \quad i = 1,\ldots,n$, where each i-th component is interpreted as the probability of the i-th column to be selected. Remarkably, the randomized algorithm above yields a matrix estimate that is "near-optimal", i.e., has error close to that of the best rank-κ approximation. Our algorithm's order of complexity with CSSP is significantly reduced from $\min(mn^2, m^2n)$ to $\min(mc^2, m^2, c)$

5. RESULTS

Our algorithm is implemented and tested in MATLAB on a desktop machine, single core on an Intel Core i7-4770 CPU and 32 GB of RAM. The average processing time on a sequence of 100 RGB frames with resolution 600×800 with image alignment and background motion estimation is about 665 seconds, which is decreased to 195 seconds with CSSP, meaning a time-saving of more than 3.4 times. We perform extensive tests using four datasets [20], [21], [22], [23], comprised of 49 videos with various challenges. This allows us to compare our method to a large number of alternative methods. For all the tests these same set of parameters are used: regularizing parameter $\lambda = 3/\sqrt{max(m,n)}$, depth of each tree $d = 3$, number of singleton groups $\mathcal{M} = 64$, and energy value for CSSP $\theta = .25$. We evaluate our method for foreground segmentation. The accuracy of foreground segmentation is measured by comparing calculated foreground support with the binary ground-truth images. Figure 1 shows the unrefined segmentation results for the i2R dataset without post-processing. Tables 1 and 2 show F-measure results for top competing methods in i2R, WallFlower, and CDnet datasets. The results guarantee superior performance for all datasets in segmentation accuracy. Our algorithm is insensitive to variability in object size as we have used a single λ for all the tests. This is corroborated with empirical evidence in segmentation accuracy measures, and also the qualitative results shown in Figure 1.

We also perform qualitative tests to illustrate the efficacy of the CSSP by comparing the calculated background via CSSP with low-rank modeling. Figure 2 shows the PSNR values obtained by using 20 values of θ linearly distributed in range $[.05, 1]$. According to this for all the our tests using 25% of the columns of A guarantees a very accurate model of the background, while a larger θ will not always result in significant increase in PSNR. An important observation here which demonstrates the advantage of using CSSP, is that, as we introduce more frames to the background (i.e., we use higher θ) we risk contaminating the background model by more foreground information; this is seen the fluctuations in Figure 2-(b), (c), (e), (g), and (h). That means an optimal θ is rather one that is smaller, that will select the most representative frames for the background of a sequence. Figure 3 demonstrates total time consumption for processing a 100 frame video in each of the datasets with varying θ in comparison with original low-rank modeling. Again our choice of θ lies in the elbow of these plots and provides time-saving guarantees.



(a) CDnet Basline (b) CDnet CameraJitter (c) CDnet DynamicBackground (d) CDnet IntermittentMotion (e) CDnet Shadow (f) CDnet Thermal (g) i2R (h) SABS [20]

Fig. 2: PSNR-θ plot of modeled background by CSSP vs. low-rank modeling. With energy value $\theta = .25$ the optimality of the quality of the modeled background is ensured.

6. CONCLUSION

We have presented a new background subtraction method and validated its efficacy and effectiveness with extensive testing. The method is based on an existing model, namely RPCA, but with new sparsity-inducing norms and group-structured sparsity constraints. Our simple model produces crisp and well-defined genuine foreground segmentation surpassing the performance of state-of-the-art methods. Moreover, our sparsity model dynamically evolves to best describe genuine foreground objects in the scene, which gives it a significant advantage when it comes to handling dynamic backgrounds, or foreground aperture. To make the problem computationally scalable we proposed using deterministic and randomized CSSP for low-rank matrix estimation. Our model proves itself to have excellent performance in dealing with heavy noise, thanks to the approximated RPCA model where the residual Gaussian noise is discarded into a third matrix in the decomposition. In addition, estimation of background motion induced by a jittering or moving camera is performed simultaneously with low-rank approximation. In future we would like to create a mechanism to handle shadows and thermal videos and obtain real-time performance for our model.

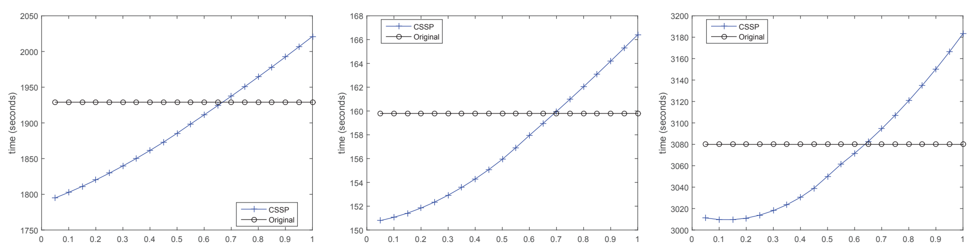
Fig. 1: *i2R* [22] and *WallFlower* [21] results: top row is the original image, second row is the ground truth, the third row is our unrefined results with no post-processing. We used the same frames as [24], [13], [25], [26], [27], and [28], for qualitative comparison.

Table 1: *i2R* [22] and *WallFlower* [21] dataset F-measure results.

method	cam	ft	ws	mc	lb	sm	ap	br	ss	*mean*
SemiSoftGoDec [29]	.0903 (10)	.2574 (9)	.4473 (10)	.4344 (10)	.3602 (10)	.6554 (8)	.5713 (7)	.3561 (10)	.2751 (9)	.3830 (10)
Stauffer [30]	.7570 (5)	.6854 (6)	.7948 (7)	.7580 (8)	.6519 (6)	.5363 (10)	.3335 (10)	.3838 (9)	.1388 (10)	.4842 (9)
Culibrk [27]	.5256 (7)	.4636 (8)	.7540 (8)	.7368 (9)	.6276 (9)	.5696 (9)	.3923 (9)	.4779 (8)	.4928 (8)	.5600 (8)
DECOLOR [7]	.3416 (9)	.2075 (10)	.9022 (5)	.8700 (4)	.646 (8)	.6822 (5)	.8169 (3)	.6589 (4)	.7480 (3)	.6525 (7)
Maddalena [28]	.6960 (6)	.6554 (7)	.8247 (6)	.8178 (7)	.6489 (7)	.6677 (7)	.5943 (5)	.6019 (6)	.5770 (6)	.6760 (6)
DP-GMM [26]	.7876 (3)	.7424 (5)	.9298 (3)	.8411 (5)	.6665 (5)	.6733 (6)	.5675 (8)	.6496 (5)	.5522 (7)	.7122 (5)
PCP [31]	.5226 (8)	.8650 (3)	.6082 (9)	.9014 (3)	.7245 (4)	.7785 (3)	.5879 (6)	.8322 (3)	.7374 (4)	.7286 (4)
LSD-GSRPCA [13]	.7613 (4)	.8371 (4)	.9050 (4)	.8357 (6)	.7313 (3)	.7362 (4)	.7222 (4)	.5842 (7)	.7214 (5)	.7594 (2)
SPGFL [24]	.8574 (2)	**.9322 (1)**	**.9856 (1)**	**.9744 (1)**	**.8840 (1)**	.8265 (2)	.7739 (2)	.8394 (2)	.8029 (2)	.8751 (3)
Ours	**.9277 (1)**	.8808 (2)	.9535 (2)	.9093 (2)	.7563 (2)	**.8950 (1)**	**.8343 (1)**	**.9196 (1)**	**.9377 (1)**	**.8904 (1)**

Table 2: *CDNet* [23] dataset: F-measure results for all the categories for the most competitive methods. Table accurate as of January 2016, with results from CDnet `http://changedetection.net/`. The online chart keeps updating.

method	bl	cj	db	im	sh	th	*mean*
DECOLOR [7]	.9215 (10)	.7776 (8)	.7084 (9)	.5945 (9)	.8317 (5)	.7081 (10)	.7570 (10)
SGMM-SOD [32]	.9223 (9)	.6988 (10)	.6826 (10)	.6957 (6)	.8613 (4)	.7081 (9)	.7624 (9)
DP-GMM [26]	.9286 (8)	.7477 (9)	.8137 (5)	.5418 (10)	.8127 (7)	.8134 (5)	.7763 (8)
2-pass RPCA [12]	.9281 (7)	.8152 (5)	.7818 (8)	.6826 (7)	.8063 (10)	.7597 (7)	.7956 (7)
MBS V0 [33]	.9287 (6)	.8367 (4)	.7904 (7)	.7092 (5)	.8063 (9)	.8115 (6)	.8092 (6)
MBS [34]	.9287 (5)	.8367 (3)	.7915 (6)	.7568 (4)	.8262 (6)	.8194 (3)	.8217 (5)
SuBSENSE [35]	**.9500 (1)**	.8150 (6)	.8180 (4)	.6570 (8)	.8990 (2)	.8170 (4)	.8260 (4)
PAWCS [36]	.9397 (4)	.8137 (7)	.8938 (3)	.7764 (3)	.8710 (3)	.8324 (2)	.8545 (3)
CDet [37]	.9458 (2)	.8367 (2)	.8991 (2)	**.8039 (1)**	.8122 (8)	**.8337 (1)**	.8552 (2)
Ours	.9430 (3)	**.8804 (1)**	**.9005 (1)**	.7837 (2)	**.9107 (1)**	.7195 (8)	**.8563 (1)**



(a) CDnet [23] (b) i2R [22] (c) SABS [20]

Fig. 3: Average time consumption vs. θ for processing a 100 frame video from each dataset using CSSP vs. original low-rank modeling. For our experiments we chose $\theta = .25$ that guarantees time-saving as well as near-optimal background modeling.

7. ACKNOWLEDGMENT

This work is supported in part by the LASIE project (`http://www.lasie-project.eu/`) with funding from the European Unions 7th Framework Program.

8. REFERENCES

[1] E. J. Candès, X. Li, Y. Ma, and J. Wright, "Robust principal component analysis?" *J. ACM*, vol. 58, no. 3, pp. 11:1–11:37, Jun. 2011.

[2] S. E. Ebadi, V. G. Ones, and E. Izquierdo, "Efficient background subtraction with low-rank and sparse matrix decomposition," in *Image Processing (ICIP), 2015 IEEE International Conference on*, Sept 2015, pp. 4863–4867.

[3] S. E. Ebadi, V. Guerra-Ones, and E. Izquierdo, "Approximated robust principal component analysis for improved general scene background subtraction," *CoRR*, vol. abs/1603.05875, 2016. [Online]. Available: http://arxiv.org/abs/1603.05875

[4] M. E. Broadbent, M. Brown, K. Penner, I. Ipsen, and R. Rehman, "Subset selection algorithms: Randomized vs. deterministic," *SIAM Undergraduate Research Online*, vol. 3, pp. 50–71, 2010.

[5] D. Papailiopoulos, A. Kyrillidis, and C. Boutsidis, "Provable deterministic leverage score sampling," in *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 2014, pp. 997–1006.

[6] J. He, L. Balzano, and A. Szlam, "Incremental gradient on the grassmannian for online foreground and background separation in subsampled video," in *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on*. IEEE, 2012, pp. 1568–1575.

[7] X. Zhou, C. Yang, and W. Yu, "DECOLOR: Moving object detection by detecting contiguous outliers in the low-rank representation." *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 3, pp. 597–610, 2013.

[8] V. Cevher, M. F. Duarte, C. Hegde, and R. Baraniuk, "Sparse signal recovery using markov random fields," in *Advances in Neural Information Processing Systems*, 2009, pp. 257–264.

[9] J. Huang, X. Huang, and D. Metaxas, "Learning with dynamic group sparsity," in *Computer Vision, 2009 IEEE 12th International Conference on*. IEEE, 2009, pp. 64–71.

[10] N. Wang and D.-Y. Yeung, "Bayesian robust matrix factorization for image and video processing," in *Computer Vision (ICCV), 2013 IEEE International Conference on*, Dec 2013, pp. 1785–1792.

[11] J. Mairal, R. Jenatton, F. R. Bach, and G. R. Obozinski, "Network flow algorithms for structured sparsity," in *Advances in Neural Information Processing Systems*, 2010, pp. 1558–1566.

[12] Z. Gao, L.-F. Cheong, and Y.-X. Wang, "Block-sparse RPCA for salient motion detection," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 36, no. 10, pp. 1975–1987, Oct 2014.

[13] X. Liu, G. Zhao, J. Yao, and C. Qi, "Background subtraction based on low-rank and structured sparse decomposition," 2015.

[14] Y. P. Hong and C.-T. Pan, "Rank-revealing QR factorizations and the singular value decomposition," *Mathematics of Computation*, vol. 58, no. 197, pp. 213–232, 1992.

[15] K. Jia, T.-H. Chan, and Y. Ma, "Robust and practical face recognition via structured sparsity," in *Computer Vision–ECCV 2012*. Springer, 2012, pp. 331–344.

[16] S. E. Ebadi, , and E. Izquierdo, "Approximated RPCA for fast and efficient recovery of corrupted and linearly correlated images and video frames," in *Systems, Signals and Image Processing (IWSSIP), 2015 International Conference on*, Sept 2015, pp. 49–52.

[17] M. W. Mahoney and P. Drineas, "CUR matrix decompositions for improved data analysis," *Proceedings of the National Academy of Sciences*, vol. 106, no. 3, pp. 697–702, 2009.

[18] P. Drineas, M. W. Mahoney, and S. Muthukrishnan, "Relative-error CUR matrix decompositions," *SIAM Journal on Matrix Analysis and Applications*, vol. 30, no. 2, pp. 844–881, 2008.

[19] I. T. Jolliffe, "Discarding variables in a principal component analysis. i: Artificial data," *Applied statistics*, pp. 160–173, 1972.

[20] S. Brutzer, B. Höferlin, and G. Heidemann, "Evaluation of background subtraction techniques for video surveillance," in *Computer Vision and Pattern Recognition (CVPR) IEEE*, 2011, pp. 1937–1944.

[21] K. Toyama, J. Krumm, B. Brumitt, and B. Meyers, "Wallflower: principles and practice of background maintenance," in *Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference on*, vol. 1, 1999, pp. 255–261 vol.1.

[22] L. Li, W. Huang, I. Y.-H. Gu, and Q. Tian, "Statistical modeling of complex backgrounds for foreground object detection," *IEEE Transactions on Image Processing*, vol. 13, no. 11, pp. 1459–1472, 2004.

[23] Y. Wang, P.-M. Jodoin, F. Porikli, J. Konrad, Y. Benezeth, and P. Ishwar, "CDnet 2014: An Expanded Change Detection Benchmark Dataset," in *IEEE CVPR Change Detection workshop*, United States, Jun. 2014, p. 8 p., https://hal-univ-bourgogne.archives-ouvertes.fr/hal-01018757.

[24] S. Javed, S. Oh, A. Sobral, T. Bouwmans, and S. Jung, "Background subtraction via superpixel-based online matrix decomposition with structured foreground constraints," in *Workshop on Robust Subspace Learning and Computer Vision, ICCV 2015*, 2015.

[25] B. Xin, Y. Tian, Y. Wang, and W. Gao, "Background subtraction via generalized fused lasso foreground modeling," *arXiv preprint arXiv:1504.03707*, 2015.

[26] T. Haines and T. Xiang, "Background subtraction with dirichletprocess mixture models," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 36, no. 4, pp. 670–683, April 2014.

[27] D. Culibrk, O. Marques, D. Socek, H. Kalva, and B. Furht, "Neural network approach to background modeling for video object segmentation," *Neural Networks, IEEE Transactions on*, vol. 18, no. 6, pp. 1614–1627, 2007.

[28] L. Maddalena and A. Petrosino, "A self-organizing approach to background subtraction for visual surveillance applications," *IEEE Transations on Image Processing*, vol. 17, no. 7, pp. 1168–1177, July 2008. [Online]. Available: http://dx.doi.org/10.1109/TIP.2008.924285

[29] T. Zhou and D. Tao, "GoDec: Randomized low-rank and sparse matrix decomposition in noisy case," in *Proceedings of the 28th International Conference on Machine Learning (ICML-11)*, ser. ICML '11, L. Getoor and T. Scheffer, Eds. ACM, June 2011, pp. 33–40.

[30] C. Stauffer and W. E. L. Grimson, "Adaptive background mixture models for real-time tracking," in *Computer Vision and Pattern Recognition, 1999. IEEE Computer Society Conference on.*, vol. 2. IEEE, 1999.

[31] Z. Zhou, X. Li, J. Wright, E. J. Candès, and Y. Ma, "Stable principal component pursuit," *CoRR*, vol. abs/1001.2363, 2010. [Online]. Available: http://arxiv.org/abs/1001.2363

[32] R. H. Evangelio and T. Sikora, "Complementary background models for the detection of static and moving objects in crowded environments," in *Advanced Video and Signal-Based Surveillance (AVSS), 2011 8th IEEE International Conference on*. IEEE, 2011, pp. 71–76.

[33] H. Sajid and S.-C. S. Cheung, "Background subtraction for static & moving camera," in *Image Processing (ICIP), 2015 IEEE International Conference on*.

[34] ——, "Universal multimode background subtraction," in *Image Processing (ICIP), Submitted to 2015 IEEE International Conference on*.

[35] P.-L. St-Charles, G.-A. Bilodeau, and R. Bergevin, "Subsense: A universal change detection method with local adaptive sensitivity," *Image Processing, IEEE Transactions on*, vol. 24, no. 1, pp. 359–373, 2015.

[36] ——, "A self-adjusting approach to change detection based on background word consensus," in *Applications of Computer Vision (WACV), 2015 IEEE Winter Conference on*. IEEE, 2015, pp. 990–997.

[37] Anonymous, "CDet," 2012. [Online]. Available: http://wordpress-jodoin.dmi.usherb.ca/method/146/