VIDEO BACKGROUND SUBTRACTION USING ONLINE INFINITE DIRICHLET MIXTURE MODELS

Wentao Fan

Concordia University
Electrical and Computer Engineering
Montreal, Canada
email: wenta_fa@encs.concordia.ca

Nizar Bouguila

Concordia University
Institute for Information Systems Engineering
Montreal, Canada
email:nizar.bouguila@concordia.ca

ABSTRACT

Video background subtraction is an essential task in computer vision for detecting moving objects in video sequences. In this paper, we propose a novel Bayesian nonparametric statistical approach to subtract video background. The proposed approach is based on a mixture of Dirichlet processes with Dirichlet distributions, which can be considered as an infinite Dirichlet mixture model. Compared to other background subtraction approaches, the proposed one has the advantages that it is more robust and adaptive to dynamic background, and it has the ability to handel multi-modal background distributions. Moreover, thanks to the nature of nonparametric Bayesian models, the determination of the correct number of components is sidestepped by assuming that there is an infinite number of components. Our results demonstrate the merits of the proposed approach.

Index Terms— Background subtraction, Dirichlet process, mixture models, Dirichlet distribution, variational Bayes.

1. INTRODUCTION

Video background subtraction is the process of identifying moving foreground objects from the background in a sequence of video frames. It is an important task in computer vision and has been applied in many applications involving video surveillance, traffic monitoring, human motion analysis and object tracking. Video background subtraction is a considerable challenging problem due to the dynamic nature of video backgrounds such as lighting changes, rain, moving leaves and shadows cast by moving objects. In recent years, much research has been devoted to the study of video background subtraction and many techniques have been proposed, such as: high level region analysis [1], kernel density estimation [2], Markov random fields [3], and hidden Markov models [4]. A brief review of background subtraction techniques can be found in [5].

Among various approaches to video background subtraction, finite mixture models have the advantage to handel multi-modal background distributions and have shown promising results in several recent works [6, 7, 8]. However, all the aforementioned mixture modeling approaches have to address the problem of determining the correct number of mixture components (either manually selected or determined by adopting some selection criteria). This difficulty can be solved in an elegant way by assuming that the number of components is countably infinite using a probabilistic structure known as Dirichlet process [9]. The Dirichlet process belongs to Bayesian nonparametric models in which the sizes of models are allowed to grow with data size. Another issue regarding background subtraction techniques using finite mixtures is that most of the works make the Gaussian assumption [6, 8], which means that each pixel in a frame is represented as a mixture of Gaussians. Unfortunately, this assumption is not realistic in practice and recent works have shown that other distributions, such as the Dirichlet distribution [10, 11] may provide a better modeling performance in the case of non-Gaussian data, and in particular normalized count data (i.e., proportion vectors) which arise in a wide variety of applications such as text, image and video modeling.

In this work, we attempt to propose a novel approach to video background subtraction through an online infinite Dirichlet mixture model. The main contributions of the paper are threefold. Firstly, we extend the finite Dirichlet mixture model to the infinite case using Dirichlet process mixture framework with a stick-breaking construction. Secondly, an online (or incremental) variational learning algorithm is developed to learn the proposed model. Lastly, we apply the proposed model to address the problem of video background subtraction and compare our approach with other existing approaches. The rest of this paper is organized as follows: Section 2 presents our infinite Dirichlet mixture model. Section 3 describes our online variational Bayes framework for learning the proposed model. Section 4 presents the methodology

The completion of this research was made possible thanks to the Natural Sciences and Engineering Research Council of Canada (NSERC).

that we have adopted for subtracting video backgrounds. Experimental results are demonstrated in Section 5. Section 6 closes this paper with conclusions.

2. THE INFINITE DIRICHLET MIXTURE MODEL

In this work, the finite Dirichlet mixture model is extended to the infinite case through a Dirichlet process. Suppose that a given D-dimensional random vector $\vec{X} = (X_1, \ldots, X_D)$ is drawn from a mixture of Dirichlet distributions with infinite components as

$$p(\vec{X}|\vec{\pi}, \vec{\alpha}) = \sum_{j=1}^{\infty} \pi_j \mathrm{Dir}(\vec{X}|\vec{\alpha}_j) \tag{1}$$

where $\vec{\pi}$ denotes the mixing coefficients which are positive and sum to one. $\vec{\alpha}_j = (\alpha_{j1}, \ldots, \alpha_{jD})$ are the positive parameters of the Dirichlet distribution $\mathrm{Dir}(\vec{X}|\vec{\alpha}_j)$ which is associated with component j, where $\mathrm{Dir}(\vec{X}|\vec{\alpha}_j)$ is defined by

$$\mathrm{Dir}(\vec{X}|\vec{\alpha}_j) = \frac{\Gamma(\sum_{l=1}^{D} \alpha_{jl})}{\prod_{l=1}^{D} \Gamma(\alpha_{jl})} \prod_{l=1}^{D} X_l^{\alpha_{jl}-1} \tag{2}$$

where $\sum_{l=1}^{D} X_l = 1, X_l > 0$ for $l = 1, \ldots, D$.
In our work, the Dirichlet process is represented using a stick-breaking construction [12], such that the mixing weights π_j are constructed by recursively breaking a unit length stick into an infinite number of pieces as $\pi_j = \lambda_j \prod_{s=1}^{j-1}(1-\lambda_s)$. The stick breaking variable λ_j is distributed according to $\lambda_j \sim \mathrm{Beta}(1, \beta)$, where β is a positive real number and is the concentration parameter of the Dirichlet process.
Given an observed data set $\mathcal{X} = (\vec{X}_1, \ldots, \vec{X}_N)$, we introduce a vector $\vec{Z} = (Z_1, \ldots, Z_N)$ as the mixture component assignment variable, in which each element Z_i takes an integer value j denoting the component from which $\vec{X}_i$ is drawn. The marginal distribution over $\vec{Z}$ can be specified as

$$p(\vec{Z}|\vec{\lambda}) = \prod_{i=1}^{N} \prod_{j=1}^{\infty} \left[\lambda_j \prod_{s=1}^{j-1} (1-\lambda_s) \right]^{\mathbf{1}[Z_i=j]} \tag{3}$$

where $\mathbf{1}[\cdot]$ is an indicator function which has the value 1 when $Z_i = j$ and 0 otherwise. The following step is to introduce prior distributions over random variables $\vec{\lambda}$ and $\vec{\alpha}$ in our Bayesian framework. According to the stick breaking construction, the prior distribution of $\vec{\lambda}$ is a specific Beta distribution as $p(\vec{\lambda}|\vec{\beta}) = \prod_{j=1}^{\infty} \mathrm{Beta}(1, \beta_j)$. Furthermore, we assume that the Dirichlet parameters $\vec{\alpha}$ are statistically independent and for each parameter α_{jl}, the Gamma distribution $\mathcal{G}(\cdot)$ is adopted to approximate the conjugate prior of $\vec{\alpha}$ as $p(\vec{\alpha}) = \prod_{j=1}^{\infty} \prod_{l=1}^{D} \mathcal{G}(\alpha_{jl}|u_{jl}, v_{jl})$, where u_{jl} and v_{jl} are hyperparameters, subject to the constraints $u_{jl} > 0$ and $v_{jl} > 0$.

3. MODEL LEARNING

In this work, we adopt an online learning algorithm proposed in [13] to learn the proposed infinite Dirichlet mixture model through variational Bayes [14]. It contains two phases: building phase and compression phase. The goal of the model building phase is to inference the current optimal mixture model, while the target of the compression phase is to determine which mixture component that groups of data points should be assigned to. In this algorithm, data points can be sequentially processed in small batches where each one may contain one or more data points.

3.1. Model Building Phase

Given an observed data set $\mathcal{X}$, we define $\Theta = \{\vec{Z}, \vec{\alpha}, \vec{\lambda}\}$ as the set of random variables. The main goal of variational Bayes is to find a proper approximation $q(\Theta)$ for the posterior distribution $p(\Theta|\mathcal{X})$, which is achieved by maximizing the free energy $\mathcal{F}(\mathcal{X}, q)$, where $\mathcal{F}(\mathcal{X}, q) = \int q(\Theta) \ln[p(\mathcal{X}, \Theta)/q(\Theta)] d\Theta$. Motivated from [15], we truncate the variational distribution $q(\Theta)$ at a value M, such that $\lambda_M = 1$, $\pi_j = 0$ when $j > M$, and $\sum_{j=1}^{M} \pi_j = 1$, where the truncation level M is a variational parameter which can be freely initialized and will be optimized automatically during the learning process [15]. Moreover, in order to achieve tractability, we assume that the approximated posterior distribution $q(\Theta)$ can be factorized into disjoint tractable factors as: $q(\Theta) = [\prod_{i=1}^{N} q(Z_i)][\prod_{j=1}^{M} \prod_{l=1}^{D} q(\alpha_{jl})][\prod_{j=1}^{M} q(\lambda_j)]$. Then, the free energy can be maximized with respect to $q(\vec{Z})$, $q(\vec{\alpha})$ and $q(\vec{\lambda})$ with a guaranteed convergence. Thus, we can obtain the following solutions for the variational factors:

$$q(\vec{Z}) = \prod_{i=1}^{N} \prod_{j=1}^{M} r_{ij}^{\mathbf{1}[Z_i=j]}, \qquad q(\vec{\alpha}) = \prod_{j=1}^{M} \prod_{l=1}^{D} \mathcal{G}(\alpha_{jl}|u_{jl}^*, v_{jl}^*) \tag{4}$$

$$q(\vec{\lambda}) = \prod_{j=1}^{M} \mathrm{Beta}(\lambda_j|a_j, b_j) \tag{5}$$

where we have defined

$$r_{ij} = \frac{\exp(\rho_{ij})}{\sum_{j=1}^{M} \exp(\rho_{ij})} \tag{6}$$

$$\rho_{ij} = \widetilde{\mathcal{R}}_j + \sum_{l=1}^{D} (\bar{\alpha}_{jl} - 1) \ln X_{il} + \langle \ln \lambda_j \rangle + \sum_{s=1}^{j-1} \langle \ln(1-\lambda_s) \rangle \tag{7}$$

$$u_{jl}^* = u_{jl} + \sum_{i=1}^{N} r_{ij} \bar{\alpha}_{jl} \Big[\Psi(\sum_{l=1}^{D} \bar{\alpha}_{jl}) - \Psi(\bar{\alpha}_{jl}) + \Psi'(\sum_{l=1}^{D} \bar{\alpha}_{jl}) \times \sum_{s \neq l}^{D} \bar{\alpha}_{js} (\langle \ln \alpha_{js} \rangle - \ln \bar{\alpha}_{js}) \Big] \tag{8}$$

$$v_{jl}^* = v_{jl} - \sum_{i=1}^{N} r_{ij} \ln X_{il} \tag{9}$$

$$a_j = 1 + \sum_{i=1}^{N} \langle Z_i = j \rangle, \qquad b_j = \beta_j + \sum_{i=1}^{N} \sum_{s=j+1}^{M} \langle Z_i = s \rangle \tag{10}$$

where $\Psi(\cdot)$ is the digamma function, and $\langle \cdot \rangle$ is the expectation evaluation. Note that, $\widetilde{\mathcal{R}}_j$ is the approximated lower

bound of $\mathcal{R}_j$, where $\mathcal{R}_j = \langle \ln[\Gamma(\sum_{l=1}^{D}\alpha_{jl})/\prod_{l=1}^{D}\Gamma(\alpha_{jl})]\rangle$. Since a closed-form expression cannot be found for $\mathcal{R}_j$, The second-order Taylor series expansion is applied to find a lower bound approximation $\widetilde{\mathcal{R}}_j$. The expected values in the above formulas are given by $\langle Z_i = j\rangle = r_{ij}$, $\bar{\alpha}_{jl} = \langle\alpha_{jl}\rangle = u^*_{jl}/v^*_{jl}$, $\langle \ln\lambda_j\rangle = \Psi(a_j) - \Psi(a_j + b_j)$, $\langle \ln(1-\lambda_j)\rangle = \Psi(b_j) - \Psi(a_j + b_j)$, and $\langle \ln\alpha_{jl}\rangle = \Psi(u^*_{jl}) - \ln v^*_{jl}$.
After convergence, the observed data points are clustered into M groups according to corresponding responsibilities r_{ij} through (6). These newly formed groups of data points are also denoted as "clumps" according to [13]. Following [13], these clumps are subject to the constraint that all data points $\vec{X}_i$ in the clump c share the same $q(Z_i) \equiv q(Z_c)$ which is a key factor in the following compression phase.

3.2. Compression Phase

In the compression phase, we attempt to determine clumps that possibly belong to the same mixture component while taking into account future arriving data. Suppose that we have already observed N data points, and our goal is to make an inference at some target time T where $T \geq N$. This is fulfilled by scaling the current observed data to the target size T, which is equivalent to using the variational posterior distribution of the observed data N as a predictive model of the future data [13]. Therefore, we can obtain the modified free energy for the compression phase as the following

$$\mathcal{F} = \sum_{j=1}^{M}\sum_{l=1}^{D}\Big\langle \ln\frac{p(\alpha_{jl}|u_{jl}, v_{jl})}{q(\alpha_{jl})}\Big\rangle + \sum_{j=1}^{M}\Big\langle \ln\frac{p(\lambda_j|\beta_j)}{q(\lambda_j)}\Big\rangle + \frac{T}{N}\sum_{c}|n_c|\ln\sum_{j=1}^{M}\exp(\rho_{cj}) \quad (11)$$

where $\frac{T}{N}$ is the data magnification factor and $|n_c|$ denotes the number of data points in clump c. The corresponding update equations for maximizing this free energy function are

$$r_{cj} = \frac{\exp(\rho_{cj})}{\sum_{j=1}^{M}\exp(\rho_{cj})} \quad (12)$$

$$\rho_{cj} = \widetilde{\mathcal{R}}_j + \sum_{l=1}^{D}(\bar{\alpha}_{jl} - 1)\ln\langle X_{cl}\rangle + \langle\ln\lambda_j\rangle + \sum_{s=1}^{j-1}\langle\ln(1-\lambda_s)\rangle \quad (13)$$

$$u^*_{jl} = u_{jl} + \frac{T}{N}\sum_{c}|n_c|r_{cj}\Big[\Psi(\sum_{l=1}^{D}\bar{\alpha}_{jl}) + \Psi'(\sum_{l=1}^{D}\bar{\alpha}_{jl}) \times \sum_{s\neq l}^{D}\bar{\alpha}_{js}(\langle\ln\alpha_{js}\rangle - \ln\bar{\alpha}_{js}) - \Psi(\bar{\alpha}_{jl})\Big]\bar{\alpha}_{jl} \quad (14)$$

$$v^*_{jl} = v_{jl} - \frac{T}{N}\sum_{c}|n_c|r_{cj}\ln\langle X_{cl}\rangle \quad (15)$$

$$a_j = 1 + \frac{T}{N}\sum_{c}|n_c|\langle Z_c = j\rangle \quad (16)$$

$$b_j = \beta_j + \frac{T}{N}\sum_{c}|n_c|\sum_{s=j+1}^{M}\langle Z_c = s\rangle \quad (17)$$

where $\langle X_{cl}\rangle$ denotes average over all data points contained in clump c. In the compression phase, we first hard assign each clump or data point to the component with the highest responsibility r_{cj} obtained from the model building phase as

$$I_c = \arg\max_{j} r_{cj} \quad (18)$$

where $\{I_c\}$ denote which component the clump (or data point) c belongs to in the compression phase. Next, we cycle through each component and split it along its principal component. Then, we refine this split by updating Eqs. (12)∼(17). After convergence criterion is reached for refining the split, the clumps are then hard assigned to one of the two candidate components. Among all the potential splits, we choose the one that results in the largest change in the free energy (Eq. (11)). The splitting process repeats itself until a stopping criterion is met. Based on [13], the stoping criterion for the splitting process can be expressed as a limit on the amount of memory required to store the components. In our work, the component memory cost for the mixture model is $\mathcal{MC} = (D-1)N_c$, where N_c is the number of components. We then define an upper limit on the component memory cost $\mathcal{C}$, and the compression phase stops when $\mathcal{MC} \geq \mathcal{C}$. Therefore, the computational time and the space requirement is bounded in each learning round. After the compression phase, the currently observed data points are discarded while the resulting components can be treated in the same way as data points in the next round of leaning. The incremental variational inference for infinite Dirichlet mixture model is summarized in Algorithm 1.

Algorithm 1

```
Choose the initial truncation level M.
Initialize the values for hyper-parameters u_jl, v_jl and β_j.
Initialize the values of r_ij by K-Means algorithm.
while More data to be observed do
    Perform the model building phase through Eqs. (4)∼(5).
    Initialize the compression phase using Eq. (18).
    while MC ≥ C do
        for j = 1 to M do
            if evaluated(j) = false then
                Split component j and refine this split using Eqs (12)∼(17).
                ΔF(j) = change in Eq. (11).
                evaluated(j) = true.
            end if
        end for
        Split component j with the largest value of ΔF(j).
        M = M + 1.
    end while
    Discard the current observed data points.
    Save resulting components into next learning round.
end while
```

4. VIDEO BACKGROUND SUBTRACTION

We tackle the problem of video background subtraction as pixel-level evaluations by following the idea proposed by [6], in which each pixel is represented by a mixture of density functions. The pixel-based approaches for subtracting background contains decision whether the pixel belongs to the background or some foreground object. In this section, we propose a novel video background subtraction approach based on the proposed online infinite Dirichlet mixture model (*OnInDMM*). Assume we have observed a frame $\mathcal{X}$ which contains N pixels, such that $\mathcal{X} = \{\vec{X}_1, \dots, \vec{X}_N\}$. Each pixel $\vec{X}_i$ is modeled as a mixture of infinite Dirichlet distributions: $p(\vec{X}_i|\vec{\pi}, \vec{\alpha}) = \sum_{j=1}^{\infty} \pi_j \mathrm{Dir}(\vec{X}_i|\vec{\alpha}_j)$, where $\vec{X}_i$ is the RGB color (three-dimensional) or the intensity value (one-dimensional) of the pixel.

Our background subtraction methodology can be summarized as follows: First, the pixel values in an observed frame are normalized to the unit sum as a preprocessing step. Next, the background model is learnt using the proposed *OnInDMM* via variational Bayes. Notice that, within this mixture model, some of the mixture components model the scene background while others model the foreground objects. Therefore, the final step is to determine whether $\vec{X}_i$ is a foreground or background pixel. In our work, we adopt the assumption that a mixture component is considered to belong to background if it occurs frequently (high π_j) and does not vary much (low standard deviation σ_j). So, after ordering all the estimated components based on the ratio $\pi_j/\|\sigma_j\|$, the first B components are chosen as background components, where

$$B = \arg\min_b \sum_{j=1}^{b} \pi_j > \Lambda \tag{19}$$

where Λ is a threshold that represents the minimum portion of the data that should be accounted for by the background, and the rest of the components are defined as foreground objects. Therefore, we can perform background subtraction for an observed frame by determining if the testing pixel $\vec{X}_i$ belongs to any of the components B.

5. EXPERIMENTAL RESULTS

Four public available video sequences with different characteristics have been adopted for evaluating the performance of the proposed background subtraction approach. These video sequences were selected to test the efficiency of our algorithm under diverse scenarios. These sequences are: 1) **Box**: This sequence shows a red box moving around by a hand with a static background; 2) **Plastic drum**: The sequence is about a plastic drum floating on the surface of sea; 3) **Person**: A person walks in front a swaying tree; 4) **Cafeteria**: The sequence consists of several minutes of an overhead view of a cafeteria. In our experiments, we initialize the truncation level M as 15. The initial values of the hyperparameters u_{jl}, v_{jl}, and β_j are set to 1, 0.01 and 0.1, respectively. Since our approach is pixel-based, foreground objects detection is thus considered as binary classification of each pixel, resulting in a segmentation mask. The *ground truth* frames are generated by manually highlighting all the moving objects in sequences. The results of representative foreground mask generated by the proposed *OnInDMM* for each video sequence can be seen in Fig. 1, where the threshold Λ is set to 0.75. The performance of the classification is measured by recall and precision, and are defined as

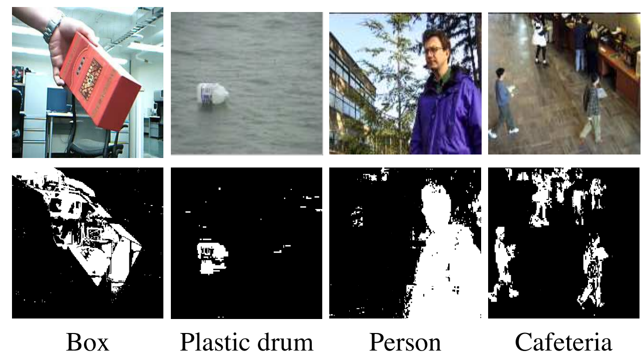

Fig. 1. Foreground masks generated by *OnInDMM* for each video sequence.

$$\text{Recall} = \frac{\text{number of correctly identified foreground pixels}}{\text{number of foreground pixels in ground truth}}$$

$$\text{Precision} = \frac{\text{number of correctly identified foreground pixels}}{\text{number of foreground pixels detected}}$$

In our case, the results of recall and precision are based on the averages over all the measured frames. A trade off needs to be considered between recall and precision: an increase in recall by detecting more foreground pixels causes a decrease in precision. Thus, we should attempt to maintain as high recall value as possible without sacrificing too much precision.

For comparison, we have also applied two other well-established background subtraction algorithms on the same data sets namely the online finite Gaussian mixture model (*OnGMM*) proposed in ([6]), and the online finite Dirichlet mixture model (*OnDMM*) proposed in [11]. Since the number of components has to be specified manfully for the *OnGMM* approach, we set it to 5 in our experiment. Figure 2 shows the comparison results in terms of precision-recall graph by varying the threshold Λ for each data set. As we can see clearly, the proposed *OnInDMM* provides the best precision and recall results for each video sequence. This is because infinite mixture models can provide a more accurate and adaptive background model than finite ones.

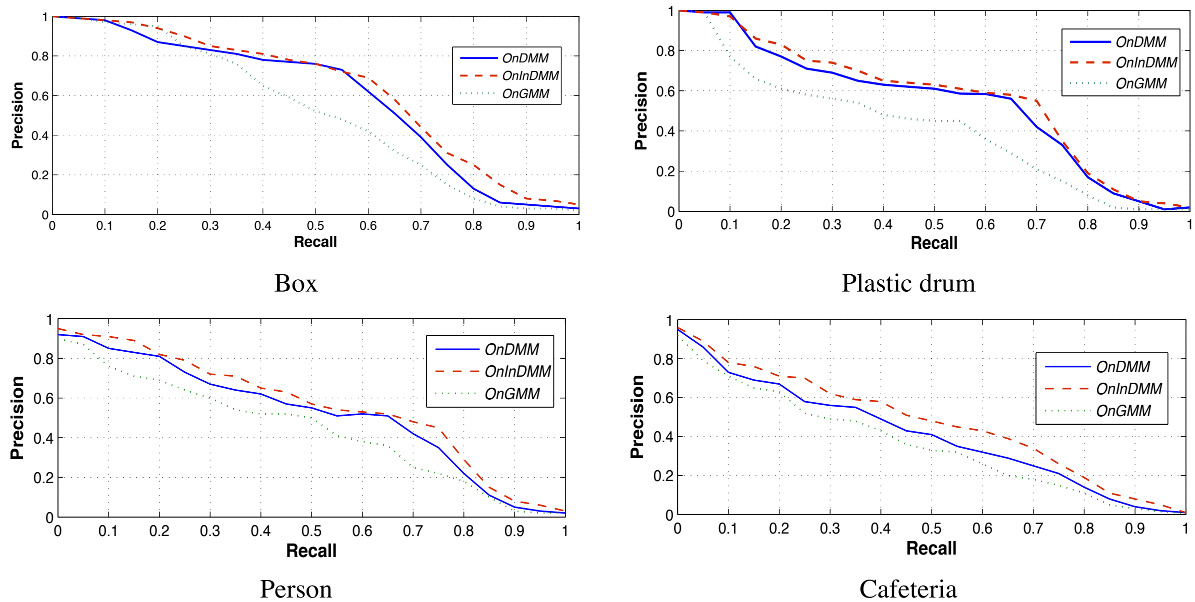

Fig. 2. Precision-recall graph for each video sequence.

6. CONCLUSION

In this paper, we developed a new approach to the problem of video background subtraction based on an online infinite Dirichlet mixture model. The model is learned in a variational way. The proposed approach is more robust to background changes (such as the variation of illumination or weather conditions) and has the merit that the difficulty of determining the number of components is avoided. The experimental results have shown the effectiveness of the proposed approach.

7. REFERENCES

[1] K. Toyama, J. Krumm, B. Brumitt, and B. Meyers, "Wallflower: Principles and Practice Of Background Maintenance," in *Proc. of the IEEE International Conference on Computer Vision (ICCV)*, 1999, pp. 255–261.

[2] A. M. Elgammal, D. Harwood, and L. S. Davis, "Non-parametric Model for Background Subtraction," in *Proc. of the European Conference on Computer Vision (ECCV)*, 2000, pp. 751–767.

[3] N. Paragios and V. Ramesh, "A MRF-based Approach for Real-Time Subway Monitoring," in *Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2001, pp. 1034–1040.

[4] B. Stenger, V. Ramesh, N. Paragios, F. Coetzee, and J.M. Buhmann, "Topology free hidden markov models: Application to background modeling," in *Proc. of the IEEE International Conference on Computer Vision (ICCV)*, 2001, pp. 294–301.

[5] M. Piccardi, "Background Subtraction Techniques: A Review," in *Proc. of the IEEE International Conference on Systems, Man and Cybernetics (SMC)*, 2004, pp. 3099– 3104.

[6] C. Stauffer and W.E.L. Grimson, "Adaptive Background Mixture Models for Real-Time Tracking," in *Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 1999, pp. 246–252.

[7] D. Lee, "Effective Gaussian Mixture Learning for Video Background Subtraction," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 5, pp. 827–832, 2005.

[8] Z. Zivkovic and F. van der Heijden, "Efficient Adaptive Density Estimation Per Image Pixel for The Task Of Background Subtraction," *Pattern Recognition Letters*, vol. 27, pp. 773–780, 2006.

[9] R. M. Korwar and M. Hollander, "Contributions to the theory of Dirichlet processes," *The Annals of Probability*, vol. 1, pp. 705–711, 1973.

[10] N. Bouguila and D. Ziou, "Unsupervised Selection of a Finite Dirichlet Mixture Model: An MML-Based Approach," *IEEE Transactions on Knowledge and Data Engineering*, vol. 18, no. 8, pp. 993–1009, 2006.

[11] W. Fan and N. Bouguila, "Online Variational Learning of Finite Dirichlet Mixture Models," *Evolving Systems*, vol. 3, no. 3, pp. 153–165, 2012.

[12] J. Sethuraman, "A Constructive Definition of Dirichlet Priors," *Statistica Sinica*, vol. 4, pp. 639–650, 1994.

[13] R. Gomes, M. Welling, and P. Perona, "Incremental Learning of Nonparametric Bayesian Mixture Models," in *Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June, pp. 1–8.

[14] H. Attias, "A Variational Bayes Framework for Graphical Models," in *Proc. of Advances in Neural Information Processing Systems (NIPS)*, 1999, pp. 209–215.

[15] D. M. Blei and M. I. Jordan, "Variational Inference for Dirichlet Process Mixtures," *Bayesian Analysis*, vol. 1, pp. 121–144, 2005.