# Visual Scene Reconstruction Using a Bayesian Learning Framework

Sami Bourouis[1,2](✉), Nizar Bouguila[3], Yexing Li[3], and Muhammad Azam[3]

[1] Department of Information Technology, College of Computers and Information Technology, Taif University, Taif, Saudi Arabia
s.bourouis@tu.edu.sa

[2] Université de Tunis El Manar, ENIT, LR-SITI Lab, 1002 Tunis, Tunisia

[3] The Concordia Institute for Information Systems Engineering (CIISE) Concordia University, Montreal, QC H3G 1T7, Canada
nizar.bouguila@concordia.ca, mu_azam@encs.concordia.ca

**Abstract.** In this paper, we focus on constructing new flexible and powerful parametric framework for visual data modeling and reconstruction. In particular, we propose a Bayesian density estimation method based upon mixtures of scaled Dirichlet distributions. The consideration of Bayesian learning is interesting in several respects. It allows simultaneous parameters estimation and model selection, it permits also taking uncertainty into account by introducing prior information about the parameters and it allows overcoming learning problems related to over- or under-fitting. In this work, three key issues related to the Bayesian mixture learning are addressed which are the choice of prior distributions, the estimation of the parameters, and the selection of the number of components. Moreover, a principled Metropolis-within-Gibbs sampler algorithm for scaled Dirichlet mixtures is developed. Finally, the proposed Bayesian framework is tested on a challenging real-life application namely visual scene reconstruction.

**Keywords:** Mixture of scaled Dirichlet distribution
Bayesian inference · Markov chain Monte Carlo algorithm
Scene reconstruction

## 1 Introduction

A recent convergence of computer graphics and computer vision has produced a set of techniques known as "data-based modeling and rendering" (DBMR). This thematic refers to methods that use pre-existing data (image and video) in order to generate new scenes and therefore gain in productivity and in realism [15]. Reconstruction of scenes has been the topic of extensive research [19,23]. It has been motivated by the exponentially growing number of available photo collections that can be used to automatically reconstruct 3D geometry and scene models; and by many potential applications such as security (e.g. crime scene


A. Mansouri et al. (Eds.): ICISP 2018, LNCS 10884, pp. 225–232, 2018.
https://doi.org/10.1007/978-3-319-94211-7_25

reconstruction) and broadcast production (e.g. movies). In the past, image synthesis was limited by the fact that all generated images did not use real data or real images to be computed. Recent studies have shown that it is possible to reconstruct the full geometry and photometry of a scene using real world digital images captured with a camera [15,16]. Some works have been proposed to generate visual scenes by analyzing and modeling captured images and video. For example, in [22], authors extended the paradigm of image-based rendering into video-based rendering, generating novel animations from video. Instead of using the image as a whole, they can also record an object and separate it from the background using background-substraction. They called this special type of video texture a video sprite. This approach has been extended in [12] where the authors introduced several learning techniques mainly to perform accurate visual data representation and then generation. In [15], authors proposed a method for synthesizing a given image that would be seen from a new viewpoint. Despite the great effort and potential done in the past decade, several challenges still need to be overcome. To deal with such problem, statistical learning approaches have been used in this context to estimate a most likely pixel value from different input views of the same scene. For instance, a two-component Gaussian mixture model has been proposed in [16] for scenes reconstruction from multiple views. To achieve acceptable quality for the reconstructed image, the design of such statistical approach normally needs a quite important number of input images. Our research here is inspired by the successful application of machine learning techniques in this area of research. Its main goal is to use image-domain features to develop new probabilistic models and to generate from these models new visual scenes from different viewpoints. To this end, there have existed many techniques to tackle this problem. Among them, finite mixture models have received a lot of attention in several domains such as pattern recognition, computer vision, data mining, and machine learning [18]. In this paper, we focus on this very topic and our main purpose is to take into account the complexity of input real data by modeling them with non-Gaussian distributions given that the Gaussian assumption is not appropriate in several applications where the data partitions are non-Gaussian. Thus, we propose to develop a new flexible mixture model based on the so-called scaled Dirichlet distribution (SDMM) [21] which is proposed as a powerful alternative to the well-known Dirichlet mixture model [5,14] and could provide better results [21]. An important problem when using mixture models is the learning of the parameters [1,7,18,20]. The parameters are usually obtained by the method of maximum likelihood (ML) performed within the expectation-maximization algorithm as done in [8,9,21]. Unfortunately this learning approach has several drawbacks such as dependency on initialization and convergence to saddle points. Bayesian learning has been proposed to overcome problems related to frequentist approaches in general and ML techniques in particular. Having appropriate prior distributions, Bayesian estimation is feasible now thanks to the development of simulation-based numerical integration techniques such as Markov chain Monte Carlo (MCMC) [4,17]. MCMC methods simulate parameters estimates by running appropriate Markov Chains using specific algorithms such as Gibbs sampler and the Metropolis algorithm [11,17]. They allow to estimate the posterior distribution of the model without needing to know the normalizing constant in Bayes' theorem. In this paper we aim to propose a Bayesian learning approach to scaled Dirichlet mixture models. To the best of our knowledge the learning of finite scaled Dirichlet mixture models have never been tackled using Bayesian inference. We propose in this work to apply our complete learning algorithm to synthesise new images from some input views. The paper is organized as follows. The next section describes the mixture model and the Bayesian approach in details. The complete estimation algorithm is given, too. Section 3 is devoted to experimental results. We end the paper with some concluding remarks.

## 2 Scaled Dirichlet-Based Bayesian Learning Framework

Mixture model is a well established approach to unsupervised learning for complex applications involving data defined in high-dimensional heterogenous (non homogenous) spaces. In this section, we introduce our Bayesian approach for visual data modeling.

### 2.1 The Finite Scaled Dirichlet Mixture Model

Let $\mathcal{X} = \{\boldsymbol{X}_1, \boldsymbol{X}_2, \ldots, \boldsymbol{X}_N\}$, a set of proportional vectors which are independent identically distributed, be a realization from a $K$-component mixture distribution. The corresponding likelihood is:

$$p(\mathcal{X}|\Theta) = \prod_{n=1}^{N} \sum_{k=1}^{K} p_k p(\boldsymbol{X}_n|\theta_k) \tag{1}$$

where $\{p_k\}$'s are the mixing parameters that are positive and sum to one, $\Theta = (\boldsymbol{p}, \theta)$, $\boldsymbol{p} = (p_1, \ldots, p_K)$, and the $\theta = \{\theta_k\}$'s are component-specific parameter vectors. Each $\boldsymbol{X}_n$ is supposed to arise from one of the $K$ components, but the cluster memberships are unknown and must be estimated. In our mixture model, $p(\boldsymbol{X}_n|\theta_k)$ is a scaled Dirichlet distribution denoted by

$$p(\boldsymbol{X}_n|\theta_k) = \frac{\Gamma(\alpha_{k+})}{\prod_{d=1}^{D} \Gamma(\alpha_{kd})} \frac{\prod_{d=1}^{D} \beta_{kd}^{\alpha_{kd}} X_{nd}^{\alpha_{kd}-1}}{(\sum_{d=1}^{D} \beta_{kd} X_{nd})^{\alpha_{k+}}} \tag{2}$$

where $\Gamma$ denotes the Gamma function, $\alpha_{k+} = \sum_{d=1}^{D} \alpha_{kd}$ and $\theta_k = (\boldsymbol{\alpha}_k, \boldsymbol{\beta}_k)$ is our model parameter. $\boldsymbol{\alpha}_k = (\alpha_{k1}, \ldots, \alpha_{kD})$ is the shape parameter that describes the form of the SDMM which is important in finding patterns inherent in a dataset, and $\boldsymbol{\beta}_k = (\beta_{k1}, \ldots, \beta_{kD})$ is the scale parameter that simply controls how the density plot is spread out.

The estimation of the parameters $\Theta$ and the selection of the appropriate number of components are determined through the learning of our finite mixture model SDMM. It is noteworthy that a frequentist approach based on ML estimation was developed previously in [21] and in this paper, we go a step further and we investigate both learning issues from a purely Bayesian perspective.

### 2.2 MCMC-Based Scaled Dirichlet Mixture Learning

The goal of Bayesian inference is to infer the model's parameters. To this end, we need to set the prior distribution on the mixture parameters and then to compute the posterior distribution from the data and selected prior. The prior can be viewed as our prior belief about the parameter before looking at the data and the posterior distribution describes our belief about the parameters after we have observed and analyzed the data. The choice of an appropriate prior distribution $p(\Theta)$ is very important for Bayesian analysis. In this case, the posterior distribution is expressed as: $p(\Theta|\mathcal{X}) \propto p(\mathcal{X}|\Theta)p(\Theta)$. In Bayesian inference, The introduction of the $\boldsymbol{Z}_n = (Z_{n1}, \ldots, Z_{nK})$ membership vectors simplifies the Bayesian analysis, where $Z_{nk} = 1$ if $\boldsymbol{X}_i$ belongs to class $k$, and $Z_{nk} = 0$ otherwise. This is done by associating with each observation $\boldsymbol{X}_n$ a missing multinomial variable $\boldsymbol{Z}_n \sim \mathcal{M}(1; \hat{Z}_{n1}, \ldots, \hat{Z}_{nK})$, where

$$\hat{Z}_{nk} = \frac{p_k p(\boldsymbol{X}_n|\theta_k)}{\sum_{k=1}^{K} p_k p(\boldsymbol{X}_n|\theta_k)} \tag{3}$$

Let's $p(\boldsymbol{p}|\mathcal{Z})$ a distribution which is given by: $p(\boldsymbol{p}|\mathcal{Z}) \propto p(\boldsymbol{p})p(\mathcal{Z}|\boldsymbol{p})$. We need then to determine $p(\boldsymbol{p})$ and $p(\mathcal{Z}|\boldsymbol{p})$. It is known that the vector $\boldsymbol{p}$ is proportional, thus a natural choice, as a prior, for this vector would be the Dirichlet distribution [17]

$$p(\boldsymbol{p}) = \frac{\Gamma(\sum_{k=1}^{K} \eta_k)}{\prod_{k=1}^{K} \Gamma(\eta_1)} \prod_{k=1}^{K} p_k^{\eta_k - 1} \tag{4}$$

where $\boldsymbol{\eta} = (\eta_1, \ldots, \eta_M)$ is the parameter vector of the Dirichlet distribution. Moreover, we have

$$p(\mathcal{Z}|\boldsymbol{p}) = \prod_{n=1}^{N} p(\boldsymbol{Z}_n|\boldsymbol{p}) = \prod_{k=1}^{K} p_k^{n_k}.$$

Hence

$$\begin{aligned} p(\boldsymbol{p}|\mathcal{Z}) \propto p(\boldsymbol{p})p(\mathcal{Z}|\boldsymbol{p}) &= \frac{\Gamma(\sum_{k=1}^{K} \eta_k)}{\prod_{k=1}^{K} \Gamma(\eta_k)} \prod_{k=1}^{K} p_k{}^{\eta_k + n_k - 1} \\ &\propto \mathcal{D}(\eta_1 + n_1, \ldots, \eta_K + n_K) \end{aligned} \tag{5}$$

where $\mathcal{D}$ is a Dirichlet distribution with parameters $(\eta_1 + n_1, \ldots, \eta_K + n_K)$. We also need to place prior distributions over parameters $\boldsymbol{\alpha}_k$ and $\boldsymbol{\beta}_k$. Formal conjugate priors do not exist for both parameters. Thus, Gamma distributions $\mathcal{G}(.)$ are adopted here with the assumption that these parameters are statistically independent:

$$p(\alpha_{kd}) = \mathcal{G}(\alpha_{kd}|u_{kd}, v_{kd}) \tag{6}$$

$$p(\beta_{kd}) = \mathcal{G}(\beta_{kd}|g_{kd}, h_{kd}) \tag{7}$$

Having these priors, the posterior distributions are given by

$$p(\boldsymbol{\alpha}_k|\mathcal{Z}, \mathcal{X}) \propto p(\boldsymbol{\alpha}_k) \prod_{Z_{ik}=1} p(\boldsymbol{X}_i|\boldsymbol{\alpha}_k) \tag{8}$$

$$p(\boldsymbol{\beta}_k|\mathcal{Z}, \mathcal{X}) \propto p(\boldsymbol{\beta}_k) \prod_{Z_{ik}=1} p(\boldsymbol{X}_i|\boldsymbol{\beta}_k) \tag{9}$$

Having all these posterior probabilities in hand, the steps of the Gibbs sampler are as follows

1. Initialization
2. Step t: For t = 1,...
   (a) Generate $Z_i^{(t)} \sim \mathcal{M}(1; \hat{Z}_{i1}^{(t-1)}, \dots, \hat{Z}_{iK}^{(t-1)})$
   (b) Compute $n_k^{(t)} = \sum_{i=1}^{N} \mathbb{I}_{Z_{ik}^{(t)}=j}$
   (c) Generate $\boldsymbol{p}^{(t)}$ from Eq. (5)
   (d) Generate $\boldsymbol{\alpha}_k^{(t)}$ and $\boldsymbol{\beta}_k^{(t)}$ $(k = 1, \dots, K)$ and from Eqs. 8 and 9, respectively, using random-walk Metropolis-Hastings (M-H) algorithm [10].

## 3 Experiments: Scenes Reconstruction

In this section, we validate our method using a challenging application namely scene reconstruction. The hyperparameters of the model $\eta_k$ are fixed at 1 which is a classical and a reasonable choice. Based on our experiments, an optimal choice of the initial values of the hyperparameters $u_{kd}$, $g_{kd}$, $v_{kd}$, $h_{kd}$ is to set them as 1, 0.01, 1, and 0.01, respectively. The goal of this section is to apply the reconstruction approach proposed in [16] by deploying our scaled Dirichlet mixture model. Scene reconstruction application has typically three steps: First, features are extracted from input images and then these features are matched between input images and finally, the resulting correspondences are deployed to estimate the final 3D geometry. Using this approach, the synthesized pixels are given as Bayesian generated estimates from the mixture model given a set of different images representing different views of the same object. In this experiment, we limited ourselves to a qualitative assessment of results. Unfortunately, we were not able to compare the obtained results with previous studies because of the lack of published works with a complete quantitative results on the same images. We are mainly motivated here in investigating the ability of the proposed Bayesian framework to synthesized pixels and to reconstruct new images from a few views (we used only 8 views for each case) of the same object. Future work will be devoted to quantitative evaluation. Figures 1, 2, 3 and 4 show examples of some objects views that we have used to validate the reconstruction approach and also the obtained reconstructed results when using the framework based on the scaled Dirichlet model. The results show some ghost effect, yet are very encouraging taking into account the limited number of views used and the difficulty of the problem. Future works could be devoted to the improvement of the reconstruction approach and to the handling of more complex scenes.

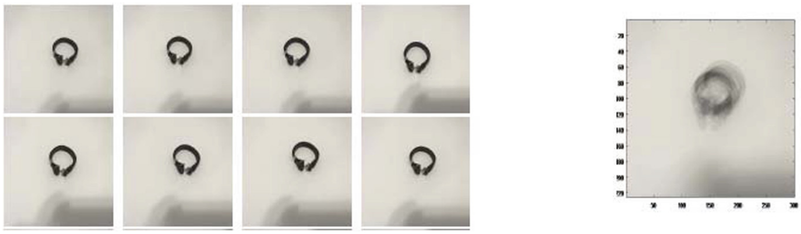

**Fig. 1.** Example 1: Representative objects views used for reconstruction (left side) and the obtained reconstructed image using the proposed framework (right side)

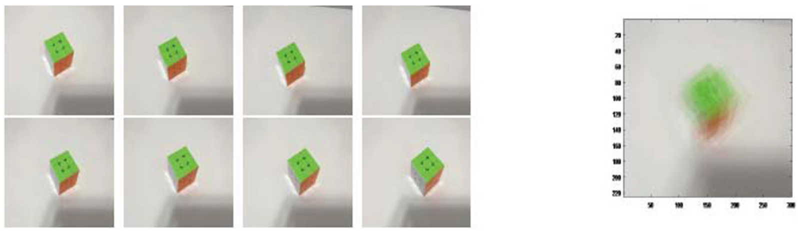

**Fig. 2.** Example 2: Representative objects views used for reconstruction (left side) and the obtained reconstructed image using the proposed framework (right side)

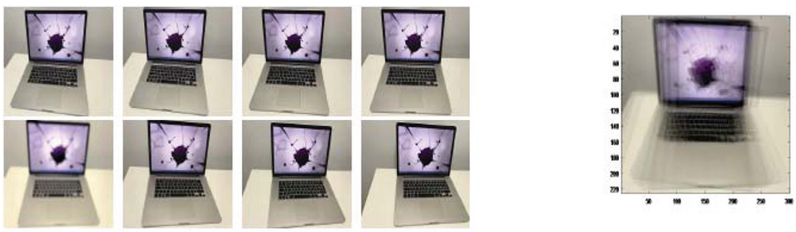

**Fig. 3.** Example 3: Representative objects views used for reconstruction (left side) and the obtained reconstructed image using the proposed framework (right side)

**Fig. 4.** Example 4: Representative objects views used for reconstruction (left side) and the obtained reconstructed image using the proposed framework (right side)

## 4 Conclusion

In this paper we have introduced a new statistical framework based on the scaled Dirichlet mixture model. The proposed framework has been learned via Bayesian inference by developing a principled MCMC-based algorithm. Experimental results have involved a challenging application namely scenes reconstruction. The obtained results are promising taking into account the complexity of such application. Future works could be devoted to the improvement of obtained results by introducing other post-processing steps. Another promising future work concerns the automatic selection of relevant features when dealing with online data modeling via the proposed framework which can be performed, for instance,, using an approach similar to the one developed in [2,3,6,13].

**Acknowledgment.** The first author would like to thank the Deanship of Scientific Research at Taif University for the continuous support. This work was supported under the grant number 1-437-5047.

## References

1. Allili, M.S., Bouguila, N., Ziou, D.: Finite generalized Gaussian mixture modeling and applications to image and video foreground segmentation. In: Proceedings of the Fourth Canadian Conference on Computer and Robot Vision (CRV), pp. 183–190 (2007)
2. Amayri, O., Bouguila, N.: On online high-dimensional spherical data clustering and feature selection. Eng. Appl. of Artif. Intell. **26**(4), 1386–1398 (2013)
3. Bouguila, N.: A model-based approach for discrete data clustering and feature weighting using MAP and stochastic complexity. IEEE Trans. Knowl. Data Eng. **21**(12), 1649–1664 (2009)
4. Bouguila, N.: Bayesian hybrid generative discriminative learning based on finite Liouville mixture models. Pattern Recogn. **44**(6), 1183–1200 (2011)
5. Bouguila, N., Ziou, D.: MML-based approach for finite dirichlet mixture estimation and selection. In: Perner, P., Imiya, A. (eds.) MLDM 2005. LNCS (LNAI), vol. 3587, pp. 42–51. Springer, Heidelberg (2005). https://doi.org/10.1007/11510888_5
6. Bouguila, N., Ziou, D.: A countably infinite mixture model for clustering and feature selection. Knowl. Inf. Syst. **33**(2), 351–370 (2012)
7. Bourouis, S., Mashrgy, M.A., Bouguila, N.: Bayesian learning of finite generalized inverted Dirichlet mixtures: application to object classification and forgery detection. Expert Syst. Appl. **41**(5), 2329–2336 (2014)
8. Channoufi, I., Bourouis, S., Bouguila, N., Hamrouni, K.: Color image segmentation with bounded generalized Gaussian mixture model and feature selection. In: 4th International Conference on Advanced Technologies for Signal and Image Processing (ATSIP 2018) (2018)
9. Channoufi, I., Bourouis, S., Bouguila, N., Hamrouni, K.: Image and video denoising by combining unsupervised bounded generalized Gaussian mixture modeling and spatial information. Multimedia Tools Appl. 1–16 (2008). https://doi.org/10.1007/s11042-018-5808-9
10. Congdon, P.: Applied Bayesian Modelling. John Wiley and Sons, Hoboken (2003)

11. Elguebaly, T., Bouguila, N.: Bayesian learning of generalized Gaussian mixture models on biomedical images. In: Schwenker, F., El Gayar, N. (eds.) ANNPR 2010. LNCS (LNAI), vol. 5998, pp. 207–218. Springer, Heidelberg (2010). https://doi.org/10.1007/978-3-642-12159-3_19
12. Fan, W., Bouguila, N.: Novel approaches for synthesizing video textures. Expert Syst. Appl. **39**(1), 828–839 (2012)
13. Fan, W., Bouguila, N.: Variational learning of a Dirichlet process of generalized Dirichlet distributions for simultaneous clustering and feature selection. Pattern Recogn. **46**(10), 2754–2769 (2013)
14. Fan, W., Sallay, H., Bouguila, N., Bourouis, S.: A hierarchical Dirichlet process mixture of generalized Dirichlet distributions for feature selection. Comput. Electr. Eng. **43**, 48–65 (2015)
15. Fitzgibbon, A., Wexler, Y., Zisserman, A.: Image-based rendering using image-based priors. Int. J. Comput. Vision **63**(2), 141–151 (2005)
16. Li, W., Li, B.: Probabilistic image-based rendering with Gaussian mixture model. In: 18th International Conference on Pattern Recognition (ICPR 2006), vol. 1, pp. 179–182 (2006)
17. Marin, J., Mengersen, K., Robert, C.: Bayesian modeling and inference on mixtures of distributions. In: Dey, D., Rao, C. (eds.) Handbook of Statistics, vol. 25, pp. 459–507. Elsevier Sciences, Amsterdam (2004)
18. McLachlan, G., Peel, D.: Finite Mixture Models. John Wiley and Sons, Hoboken (2004)
19. Mustafa, A., Kim, H., Guillemaut, J.Y., Hilton, A.: General dynamic scene reconstruction from multiple view video. In: 2015 IEEE International Conference on Computer Vision (ICCV), pp. 900–908, December 2015
20. Najar, F., Bourouis, S., Bouguila, N., Belguith, S.: A comparison between different Gaussian-based mixture models. In: 14th IEEE International Conference on Computer Systems and Applications, Tunisia. IEEE (2017)
21. Oboh, B.S., Bouguila, N.: Unsupervised learning of finite mixtures using scaled Dirichlet distribution and its application to software modules categorization. In: 2017 IEEE International Conference on Industrial Technology (ICIT), pp. 1085–1090, March 2017
22. Schödl, A., Essa, I.A.: Machine learning for video-based rendering. In: Advances in Neural Information Processing Systems. pp. 1002–1008 (2001)
23. Snavely, N., Simon, I., Goesele, M., Szeliski, R., Seitz, S.M.: Scene reconstruction and visualization from community photo collections. Proc. IEEE **98**(8), 1370–1390 (2010)