DP-Fusion: A generic framework of online multi sensor recognition

Ming Liu*, Lujia Wang†, Roland Siegwart*
* Autonomous Systems Lab, ETH Zurich, Switzerland
† Department of Electronic Engineering, The Chinese University of Hong Kong
[ming.liu,lujia.wang]@mavt.ethz.ch, rsiegwart@ethz.ch

***Abstract*—Multi sensor fusion has been widely used in recognition problems. Most existing work highly depend on the calibration between different sensor information, but less on modeling and reasoning of co-incidence of multiple hints. In this paper, we propose a generic framework for recognition and clustering problem using a non-parametric Dirichlet hierarchical model. It enables online labeling, clustering and recognition of sequential data simultaneously, while taking into account multiple types of sensor readings. The algorithm is data-driven, which does not depend on prior-knowledge of the data structure. The result shows the feasibility and its reliability against noise data.**

I. Introduction

Perception is the process that converts raw sensor readings to expedient information. As we know, human are good at perception. One important reason is that we use multiple sensors, such as eyes, nose and ears together, trying to gather information from different perspectives. Luo et al in [1] provided an interesting biological explanation of multi sensor integration for animals. Inspired by this fact, in past decades, multi sensor fusion [2] has shown its importance impact in different engineering fields, such as monitoring of complex structure, fault diagnosis and especially robotics. Most recent work treated multi sensor fusion problem in a decentralized fashion [3]. In brief words, they first considered multi sensor readings separately, reason/infer them, then fuse the conclusion of each sensor in the end. This pattern is potentially robust to failure of any one of sensors. Nevertheless, we observe the following drawbacks. First, it highly depends on the calibration precision between sensors; second, high believes on certain sensors may cause false positives in final results as well; last but not the least, coincidence of multiple information hints is simply ignored. However, it is the primary pattern how human recognize the world, i.e. through appearance, smell, audition, tactility etc. simultaneously.

In order to elevate this, we introduce a generic framework which allows recognition tasks to take multiple sensor readings simultaneously. It is proved to be low in computation. Meanwhile, the sensor measures are coherently linked together via clustering. As a primary feature, the proposed algorithm uses non-parametric statistics to discover the inner relations among data from different subjects. It starts from zero prior-knowledge and takes sequence of concurrent data from different sensors as input. No specific training is required during the process. It enables the fused data to autonomously build new clusters and recognize existing cluster in real-time.

The proposed model is stimulated by Dirichlet Process Mixture Model (DPMM) [4], which is nowadays widely used in texts classification and segmentation. The original algorithm takes only one type of input, such as words or letters. Moreover, the inference of a DPMM is computationally expensive, because sampling algorithms are usually required [5] from the large test set. We extend the model to multiple observation from different sensors and develop an online approximation algorithm which enables fast inference in real-time.

A. Pattern of Multi sensor Fusion

The taxonomy of data fusion algorithms varies. We only list several related elements that are generally used in surveys.

1) Decision Fusion: Decision making is the most critical problem for intelligent systems. It is a general concept and is usually embedded into specific paradigms, such as failure detection, object recognition, pedestrian detection etc. Several work regarding decision fusion have been proposed in the scope of decentralized multi sensor state representation. In [6], the authors introduced a decision fusion framework to fuse multi sensors by using confidence regions of the sensor model. Fauvel et al [7] uses fuzzy set theory to fuse the decision from multiple classifiers. [8] introduced a force aggregation and classification model by fusing information from sensors with different resolutions.

2) Sensory State: The purpose of multi sensor fusion is to obtain information more robustly than single sensor. In most cases, the target information can be considered as goal states. At early stage, Extended Kalmann Filter (EKF) was widely used, where the perception outcomes from multi sensors are taken as a unified state. This model is usually named as centralized state estimation. Several robotic applications are proposed, such as [9] fuses vision and haptic sensor for object recognition; [10] used neural networks to fuse the sensor information of intelligent vehicles etc. However, these early work do not treat the multi sensor fusion mathematically efficiently. Moreover, the robustness to sensor failure is a big problem. Durrant-Whyte et al proposed a decentralized architecture named Decentralized Kalmann Filter (DKF) in [3], which handles multi sensor data separately then fuse the conclusions from each filter. The obvious advantage of DFK lies in its robustness to single sensor failure. Some recent researches still follow the same concept, such as object recognition by [11], segmentation problem by [12], pose

This work was supported by the EU FP7 project NIFTi (contract # 247870) and EU project Robots@home (IST-6-045350).

estimation problem by [13], [14]. The proposed algorithm does not show explicitly decentralized characteristics. However, the joint probability given in section II depicts the independence of all sensor readings. It indicates that the confidence of each sensor is propagated to the posterior directly, which means sensor readings are not centralized as a single system state.

B. Clustering

In order to automate the classification and recognition process, an unsupervised learning algorithm is required. Sophisticated clustering algorithms usually depend on iterative calculation such as K-means, spectral clustering [15] or affinity-propagation [16]. A representative of online reasoning is chow-liu tree based segmentation [17] for static data and change point detection [18], [19] for sequential data. For extreme cases, the synchronization of multi sensor data need to be taken care of [20] or spatial and temporal hints must be jointly considered [21]. In this paper, an online naive change point detection algorithm is implemented, which is validated through simulation in section IV.

C. Recognition and inference

Recognition is the core of most robotic applications. For example, robot topological mapping requires detection and recognition of loop-closure; semantic mapping usually requires recognition of objects; human-machine interfaces require recognition of human behaviors etc. Researches targeting at these core problems attempt to seek the best algorithms to build efficient models which can represent this perception process efficiently.

Regarding inference approaches, hierarchical probabilistic methods based on statistical techniques won a great success in text mining and biological information processing [22], [23]. In this work, we alternate the classical mixture model to fit them with multiple types of observations. At the same time, we allow infinite increment of the number of labels. Furthermore, the model is to be learned, updated, inferred in real-time online.

In most of the related works, change-point detection [24], [19], [25] is the basis to segment a data sequence. In this work, as we are targeting at a lightweight method, the change-point detection is not feasible when using multiple hypothesis methods, such as particle filtering [19]. Instead, we use non-parametric statistic test to evaluate the labeling for each frame separately. This may cause instability in the output label. However, it relief the requirement of saving all the previous data of the sequence.

The theoretical advances in hierarchical probability frameworks, such as LDA [23] and HDP [22], provide a good support for our algorithm. The Dirichlet Process Mixture Model (DPMM) enables countable infinite clusters for the measures, which can be used to represent the process of state recognition.

D. Assumptions and Contributions

Not withdraw the generality, the proposed algorithm deals with data with the following assumptions.

- The multi sensor readings are synchronized, or they can be treated as a complete observation unit when they have different sampling rate;
- Features of the sensor readings are observable and computational feasible in near real-time;
- As an assumption of DPMM, multi sensor readings in the data set must be exchangeable, which indicates that the labeling of a reading does not depend on whether such reading appears earlier or later.

The objectives that we want to achieve in this paper are double folded.

- Modeling multi sensor recognition process using hierarchical probability model. The model of the recognition process depends on parameter set with small cardinality.
- A concise approach for on-line inference of the proposed Dirichlet Process Mixture Model;

The remainder of this paper is organized as follows. We will start with proposing the hierarchical model for online recognition using multi sensor data. The full inference of the model will also be introduced. Then we introduce an approximate method for fast inference of the model. We explain the evaluation of the model using simulation in section IV. The conclusion and future steps of this work are given in the end.

II. Model Formulation

We propose a DPMM model as shown in figure 1, where the parameters are depicted in rectangles, and random variables are in circles. Especially, the following components are designed in the proposed model.

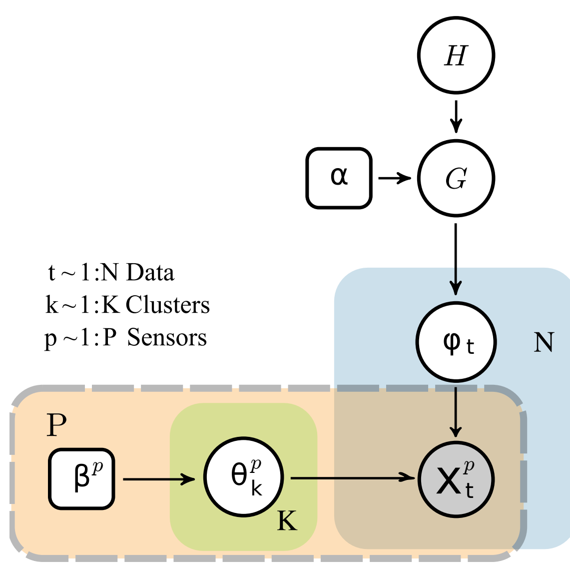

Fig. 1. Directed Acyclic Graph (DAG) of the proposed model for recognition by multi sensor fusion

A. Chinese Restaurant Process (CRP)

G is a Dirichlet process distributed with base distribution H and concentration parameter α. The base distribution is the mean of the DP and the concentration parameter α is as an inverse variance. The distribution G itself has point masses, and the draw from G will be repeated by sequential draws considering the case of an infinite sequence. Additionally, ϕ_t is an indicator of the cluster identity which the current data set at time t belongs to. We could see that ϕ_t is the target variable of inference. If the process is considered as a partition problem, a CRP model is usually used. It uses a prior obtained from a stick-breaking process [26]. By integrating over G, the drawing of ϕ_t 's can be depicted as:

$$\phi_t \mid \phi_{1:t-1} \sim \frac{\sum_{n=1}^{t-1} \delta_{\phi_n} + \alpha H}{t-1+\alpha}$$

where δ_{ϕ_n} is an indicator of a certain frame n is labeled as ϕ_n, i.e. a mass point function locates at ϕ_n. We must notice that this assumption implies that the more we see a certain cluster of data, the high a prior that data from such cluster may be observed again. The target problem is then converted to an estimation of

$$P(\phi_t \mid \phi_{\setminus t}, G, \boldsymbol{x}, \boldsymbol{\theta}; \alpha, \beta)$$

where $\phi_{\setminus t}$ is the full set of indicators excluding the current one, namely the history labels. Sets of random variables and sets parameters are shown in bold.

B. Multi Sensor Data Perception

The multi sensor data of P different types of readings are modeled as the orange plate (encircled by dashed line) shown in figure 1. For all N readings in the sequence, x_t^p represents the perceived information acquired at time-stamp t from sensor p. Taking discretized readings as an example, perceived information from raw sensor data can be represented as histograms [27]. Assuming there are K different clusters, θ_k is then a matrix of $K \times Z^p$, where Z^p is the number of possible histograms for sensor p. x_t's of dimension Z^p are drawn from θ_k. In general cases, Z^p represent the number of possible readings from sensor p.

On one hand, x_t^p is inherently determined by its label ϕ_t, as defined in section II-A; on the other hand, we can also consider the sensor reading as a draw (sample) from a sensor model θ_k^p for cluster k, with a sensor model prior β^k. So far, we build a model of two sub-processes, namely the sensoring process and perception process, which serves as a basis to build data-driven inference model of the recognition problem.

C. Model Inference

As a summary of the proposed model,

$$\begin{aligned} G &\sim Dir(\alpha H) \\ \phi_t \mid G &\sim G \\ x_t^p &\sim F(\phi_t, \theta_{\phi_t}^p) \end{aligned}$$

F represent the generation function of the measurements from the base models, regarding label ϕ_t. The joint probability can be written directly as,

$$p(\phi\ G\ \theta\ \boldsymbol{x};\ \boldsymbol{\beta}) = \prod_{p=1}^{P} \prod_{k=1}^{K} p(\theta_k^p\ ;\ \beta^p)$$
$$\prod_{t=1}^{N} p(G\ ;\ H, \alpha)\ p(\phi_t \mid G)\ \prod_{p=1}^{P} p(x_t^p \mid \theta_{\phi_t}^p)$$

In order to factorize it to independent components, we integrate the joint probability over $\theta^1, \theta^2 \dots \theta^P$ and G,

$$\begin{aligned} p(\phi\ \boldsymbol{x}\ ;\ \boldsymbol{\beta}) \ &= \ \int_{\theta^1} \dots \int_{\theta^P} \int_G p(\phi\ G\ \theta\ \boldsymbol{x}\ ;\ \beta)\ dG\ d\theta^1 \dots d\theta^P \\ &= \ \int_{\theta^1} \prod_{r=1}^{K} p(\theta_k^1\ ;\ \beta^1) \prod_{t=1}^{N} p(x_t^1 \mid \theta_{\phi_t}^1)\ d\theta^1 \\ &\dots \\ &\int_{\theta^P} \prod_{r=1}^{K} p(\theta_k^P\ ;\ \beta^P) \prod_{t=1}^{N} p(x_t^P \mid \theta_{\phi_t}^P)\ d\theta^P \\ &\int_G \int_H \prod_{t=1}^{N} p(\phi_t \mid G) p(G\ ;\ H\alpha)\ dH\ dG \end{aligned} \tag{1}$$

The last component is an exception of G, i.e. $E_G\left[p(\phi_1\ \phi_2\ \phi_3\ \phi_4 \cdots\ \phi_N \mid G)\right]$. According to the features of the Dirichlet process, it is proportional to the product $\prod_{t=1}^{N} p(\phi_t \mid \phi_{\setminus t}) \propto p(\phi_N \mid \phi_{\setminus N})$. Therefore,

$$\int_G \int_H \prod_{t=1}^{N} p(\phi_t \mid G) p(G\ ;\ H\alpha)\ dH dG \propto \frac{\sum_{t=1}^{N-1} \delta_{\phi_t} + \alpha \delta_{\phi_{\bar{k}}}}{N-1+\alpha} \tag{2}$$

where δ_{ϕ_n} is a mass point function located at ϕ_n. $\bar{k}$ is the indicator for a new cluster.

The first parts can be treated in a similar manner. Take the integral of θ_p for an instance, using n_v^k representing the number of measures who is the v-th element in θ_p within cluster k.

$$\begin{aligned} &\int_{\theta^p} \prod_{k=1}^{K} p(\theta_k^p\ ;\ \lambda) \prod_{t=1}^{N} p(w_t \mid \theta_{\phi_t})\ d\theta^p \\ &= \prod_{k=1}^{K} \int_{\theta_k^p} \frac{\Gamma(\sum_{v=1}^{Z^p} \beta_v^p)}{\prod_{v=1}^{Z^p} \Gamma(\beta_v^p)} \prod_{v=1}^{Z^p} \theta_{k,v}^{\beta_v^p - 1} \prod_{v=1}^{Z^p} \theta_{k,v}^{n_v^k}\ d\theta_k^p \\ &= \prod_{k=1}^{K} \int_{\theta_k^p} \frac{\Gamma(\sum_{v=1}^{Z^p} \beta_v^p)}{\prod_{v=1}^{Z^p} \Gamma(\beta_v^p)} \prod_{v=1}^{Z^p} \theta_{k,v}^{\beta_v^p + n_v^k - 1}\ d\theta_k^p \end{aligned} \tag{3}$$

since from the integral of Dirichlet distribution,

$$\int_{\theta_k^p} \frac{\Gamma(\sum_{v=1}^{Z^p} \beta_v^p + n_v^k)}{\prod_{v=1}^{Z^p} \Gamma(\beta_v^p + n_v^k)} \prod_{v=1}^{Z^p} \theta_{k,v}^{\beta_v^p + n_v^k - 1}\ d\theta_k^p = 1 \tag{4}$$

The joint probability is represented as follows.

$$\begin{aligned} &p(\phi\ \boldsymbol{x}\ ;\ \boldsymbol{\beta}) \\ &\propto \prod_{p=1}^{P} \prod_{k=1}^{K} \frac{\Gamma(\sum_{v=1}^{Z^p} \beta_v^p)}{\prod_{v=1}^{Z^p} \Gamma(\beta_v^p)} \frac{\prod_{v=1}^{Z^p} \Gamma(\beta_v^p + n_v^k)}{\Gamma(\sum_{v=1}^{Z^p} \beta_v^p + n_v^k)} \\ &\left(\frac{\sum_{t=1}^{N-1} \delta_{\phi_t} + \alpha \delta_{\phi_{\bar{k}}}}{N-1+\alpha} \right) \end{aligned} \tag{5}$$

When we consider a collapsed Gibbs sampling process on the cluster indicator ϕ_t at time t, we have

$$p(\phi_t \mid \phi_{\setminus t}\ \boldsymbol{x}\ ;\ \boldsymbol{\beta}) \ \propto\ p(\phi_t\ \phi_{\setminus t}\ \boldsymbol{x}\ ;\ \boldsymbol{\beta}) \tag{6}$$

However, the huge size of Z^p makes the direct inference not possible. Usually sampling methods [5] is used to estimate the posterior. Nevertheless, the sampling based algorithm usually is computational expensive as well. It is required to find an online approximation algorithm, in order to make the algorithm work in real-time.

III. Approximation

In this section, we introduce the approximation algorithm to infer the proposed DPMM. For the case where measurement $\phi_t = k$, for simplicity, we rewrite equation 5 as follows.

$$\begin{aligned} & p(\phi_t = k \mid \phi_{\ \setminus t}\ \boldsymbol{x}) \\ & \propto \prod_{p=1}^{P} \frac{\prod_{v=1}^{Z^p} \Gamma(\beta_v^p + n_v^k)}{\Gamma(\sum_{v=1}^{Z^p} \beta_v^p + n_v^p)} \left(\frac{\sum_{t=1}^{N-1} \delta_k + \alpha\delta_{\phi_{\bar{k}}}}{N-1+\alpha} \right) \\ & = \prod_{p=1}^{P} \xi^p(x_t^p \mid \theta_{\phi_t}^p) p(\phi_t \mid \phi_{\ \setminus t}) \end{aligned} \tag{7}$$

We could see from equation 7 that the first P components $\xi^p()$s calculate the gamma function of the count of a certain observation over all possibilities. In another word, they represent the probability of a certain measure showing up in a sequence of observations. Therefore, it can also be considered as a measure of the similarity of current observation to all the predefined models. As a result, we don't need sampling methods to estimate this measure if we can approximate the underlying similarity between current observation and reference models. This conclusion leads to very flexible means to recognition problems, since the similarity between observation and model can be obtained by various criteria, e.g. number of matched key-point features, result of spectrum analysis, dot product of observation vectors etc. In the end, a scalar will be used to indicate this similarity. The resulting scalar s can further represent the observation as a sample from a distribution of exponential family, such as zero-mean Gaussian distributions $[e.g.\ C \cdot e^{-s^2}]$ or Beta distribution $[e.g.\ Be(1, S)\ where\ S > 1]$.

However, another factor much be considered. It is the weighting factor among all sensors. As for equation 5, this factor is modeled by prior $\boldsymbol{\beta}$. Joining with the approximation by exponential family distribution,

$$\xi^p(x_t^p \mid \theta_{\phi_t}^p) \equiv e^{-(\omega_p s^2(x_t^p, \hat{\theta}_k^p))}$$

A set of weights for sensors can be used as follows.

$$\begin{cases} p(\phi_t = k \mid \phi_{\ \setminus t}\ \boldsymbol{x}) \\ \propto \left(\dfrac{\sum_{t=1}^{N-1} \delta_k + \alpha\delta_{\phi_{\bar{k}}}}{N-1+\alpha} \right) e^{-(\sum_p \omega_p s^2(x_t^p, \hat{\theta}_k^p))} \\ \displaystyle\sum_{p=1}^{P} \omega_p = 1. \end{cases} \tag{8}$$

where $\hat{\theta}_k^p$ is the incrementally estimated model, and $s()$ depicts the matching result between the current observation and the model. One example of the incremental estimation of the model is given in section V.B of [27].

IV. Simulation

The simulation with multi sensor inputs for online clustering and recognition is introduced in this section. We simulate three synchronized sensor readings, which are observed from a system with change states. The ground truth of the changing state is shown in first block of figure 2. Subplot A,B and C show the readings from three different sensors. Please note that sensor reading C provides only noisy signal, independently to the state change. We use the case C to simulate that low information sensor readings could be successfully omitted by the proposed decentralized framework. The sensor models are zero-biased Gaussian distributions. We first check the

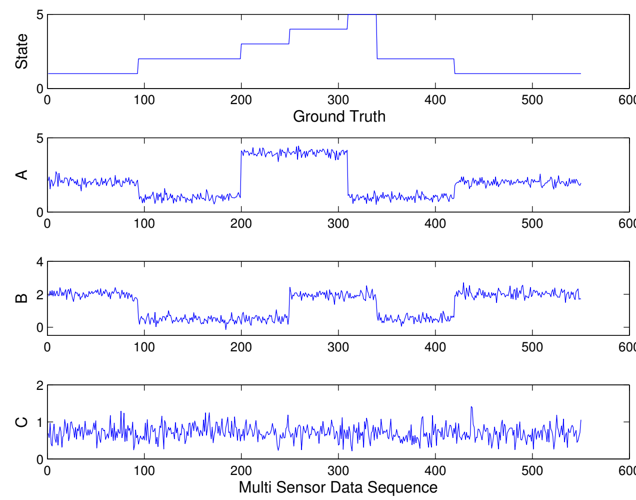

Fig. 2. Simulated multi sensor readings

likelihoods of the joint probability regarding state 1,2 and 3 without considering change-point detection, in order to validate the distinctness drawn from equation 7. The joint likelihood for three sensors are shown in figure 3. It shows

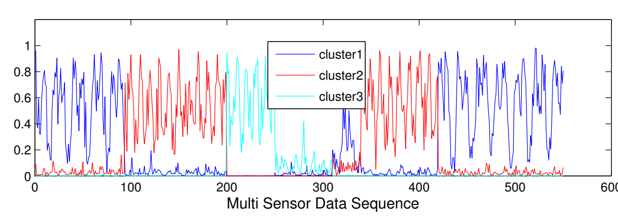

Fig. 3. Likelihood of readings

that if the cluster of the data is given, the model could validate sensor readings as samples from each cluster. The remaining problem is that the clusters of date (respecting each state of the system) need to be automatically detected and incrementally generated. To this end, we use a naive change-point detection algorithm, since the noise level of the simulated data is low. A time-stamp is considered as change-point when the posterior of the observations is lower than a threshold in conjugated 5 readings. For sophisticated change-point detection algorithm such as particle filter, please refer to [25], [19], [28]. The result is shown in figure 4. The first

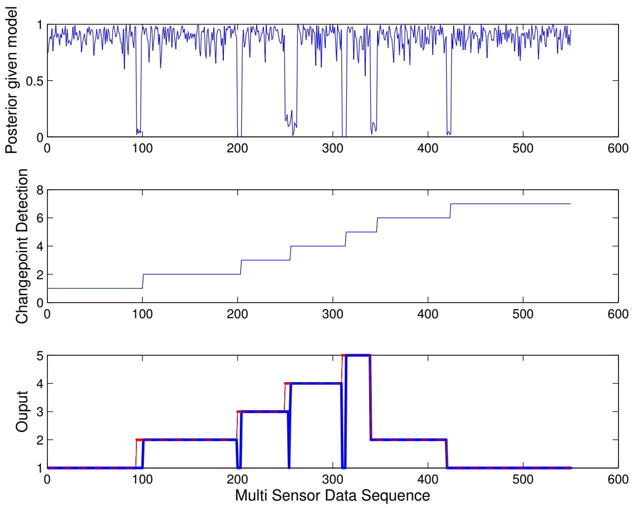

Fig. 4. Simulation of change point detection and clustering result

subplot shows the posterior of MAP (Maximize-a-Posterior) result. The change-point detection is shown in the second part. In the end, we present the resulting labeling (in blue) against the ground truth (in red).

The results indicate that the proposed DPMM model is able to detect and register new clusters of data online, while performing recognition task simultaneously. Experiment results on a real-time scene recognition problem, by fusing two different types of readings, can be obtained from our previous report [27].

V. Discussion and Reasoning

A. *Independence and decentralized state*

We observe from equation 8 that the inference of the DPMM model falls back to a product of the likelihood of each sensor reading and a CRP process. It shows that given the observation x_t^p, all the sensor perception model are independent. This is consistent with the original model design of figure 1. Therefore the system state can be easily written as a decentralized way. It means that a DFK filter is also applicable as post-processing.

B. *Complexity*

We instantiated a similar model as scene recognition problem in [27] for dual-sensor perception. Based on further study, we draw the following properties of the model.

1) Cardinality: Recognition algorithm usually leads to a big set of parameters. The choice of parameters especially thresholds will lead to dramatically change in the final result. Equation 8 shows that the proposed algorithm depends on the weighting factors for sensors and prior of the CRP process. The influence of prior α for the CRP process can be ignored when the number of measurements N goes large. The weighting factors can either be chosen empirically (e.g. vision usually plays a more important role in object recognition than laser) or enable them to be adaptively tuned by variance analysis (e.g. greater variance among different measures leads to higher weight) etc.

2) Computational complexity: As the number of measurement grows, new models are automatically detected and updated. The complexity of the algorithm only rises linearly along with the number of models. Based on the dual-sensor model in [27], we record the computation time over the test. The result is shown in figure 5. This result implies the potential of the proposed method can be extended to large scale dataset without jeopardizing the real-time ability.

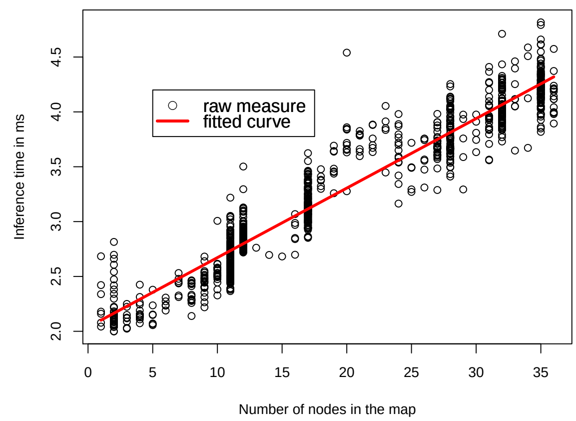

Fig. 5. Inference time vs number of nodes

VI. Conclusion and Future Work

In this paper, we present *DP-Fusion*, an on-line information fusion framework for multi sensor data based on Dirichlet Process Mixture Model. It combines synchronized sensor readings to automatically cluster data into models, while recognizing data from existing models simultaneously. Results show its advantage of on-line computing mode and low computational cost. This study also implies that the inference of a DPMM can be approximated by the product of the conditional probability. We envision that similar concept can be borrowed to solve other inference problem as well. For further study, we will explore automatic learning of the parameter sets required by the model using kernel density estimation. Extended experiments on large dataset will also be carried out.

References

[1] R. Luo and M. Kay, "Multisensor integration and fusion in intelligent systems," *Systems, Man and Cybernetics, IEEE Transactions on*, vol. 19, no. 5, pp. 901–931, 1989.

[2] D. Hall and J. Llinas, "An introduction to multisensor data fusion," *Proceedings of the IEEE*, vol. 85, no. 1, pp. 6–23, 1997.

[3] H. Durrant-Whyte, B. Rao, and H. Hu, "Toward a fully decentralized architecture for multi-sensor data fusion," in *Robotics and Automation, 1990. Proceedings., 1990 IEEE International Conference on*. IEEE, 1990, pp. 1331–1336.

[4] S. MacEachern and P. Müller, "Estimating mixture of dirichlet process models," *Journal of Computational and Graphical Statistics*, pp. 223–238, 1998.

[5] R. Neal, "Markov chain sampling methods for dirichlet process mixture models," *Journal of computational and graphical statistics*, pp. 249–265, 2000.

[6] S. Thomopoulos, R. Viswanathan, and D. Bougoulias, "Optimal decision fusion in multiple sensor systems," *Aerospace and Electronic Systems, IEEE Transactions on*, no. 5, pp. 644–653, 1987.

[7] M. Fauvel, J. Chanussot, and J. Benediktsson, "Decision fusion for the classification of urban remote sensing images," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 44, no. 10, pp. 2828–2838, 2006.

[8] B. Yu, J. Giampapa, S. Owens, and K. Sycara, "An evidential model of multisensor decision fusion for force aggregation and classification," in *Information Fusion, 2005 8th International Conference on*, vol. 2. IEEE, 2005, pp. 8–pp.

[9] P. Allen, "Integrating vision and touch for object recognition tasks," *Multisensor integration and fusion for intelligent machines and systems*, pp. 407–440, 1995.

[10] S. Nagata, M. Sekiguchi, and K. Asakawa, "Mobile robot control by a structured hierarchical neural network," *Control Systems Magazine, IEEE*, vol. 10, no. 3, pp. 69–76, 1990.

[11] D. Klimentjew, N. Hendrich, and J. Zhang, "Multi sensor fusion of camera and 3d laser range finder for object recognition," in *Multisensor Fusion and Integration for Intelligent Systems (MFI), 2010 IEEE Conference on*. IEEE, 2010, pp. 236–241.

[12] N. Hendrich, D. Klimentjew, and J. Zhang, "Multi-sensor based segmentation of human manipulation tasks," in *Multisensor Fusion and Integration for Intelligent Systems (MFI), 2010 IEEE Conference on*. IEEE, 2010, pp. 223–229.

[13] G. Nuetzi, S. Weiss, D. Scaramuzza, and R. Siegwart, "Fusion of imu and vision for absolute scale estimation in monocular slam," *Journal of Intelligent and Robotic Systems*, vol. 61, p. 287299, 2011.

[14] A. Martinelli, "Vision and imu data fusion: Closed-form solutions for attitude, speed, absolute scale, and bias determination," *Robotics, IEEE Transactions on*, vol. 28, no. 1, pp. 44 –60, feb. 2012.

[15] I. S. Dhillon, Y. Guan, and B. Kulis, "Kernel k-means , Spectral Clustering and Normalized Cuts," in *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 2004, pp. 551–556.

[16] B. Frey and D. Dueck, "Clustering by passing messages between data points," *science*, vol. 315, no. 5814, p. 972, 2007.

[17] M. Liu, F. Colas, and R. Siegwart, "Regional topological segmentation based on mutual information graphs," in *Proc. of the IEEE International Conference on Robotics and Automation (ICRA)*, 2011.

[18] J. Pillow, Y. Ahmadian, and L. Paninski, "Model-based decoding, information estimation, and change-point detection techniques for multineuron spike trains," *Neural Computation*, vol. 23, no. 1, pp. 1–45, 2011.

[19] A. Ranganathan, "PLISS: labeling places using online changepoint detection," *Autonomous Robots*, no. 32, pp. 351–368, 2012.

[20] N. Kaempchen and K. Dietmayer, "Data synchronization strategies for multi-sensor fusion," in *Proceedings of the IEEE Conference on Intelligent Transportation Systems*. Citeseer, 2003.

[21] B. Douillard, D. Fox, and F. Ramos, "A spatio-temporal probabilistic model for multi-sensor multi-class object recognition," *Robotics Research*, pp. 123–134, 2011.

[22] Y. Teh, M. Jordan, M. Beal, and D. Blei, "Hierarchical dirichlet processes," *Journal of the American Statistical Association*, vol. 101, no. 476, pp. 1566–1581, 2006.

[23] D. Blei, A. Ng, and M. Jordan, "Latent dirichlet allocation," *The Journal of Machine Learning Research*, vol. 3, pp. 993–1022, 2003.

[24] L. Orváth and P. Kokoszka, "Change-point detection with nonparametric regression," *Statistics: A Journal of Theoretical and Applied Statistics*, vol. 36, no. 1, pp. 9–31, 2002.

[25] B. Ray and R. Tsay, "Bayesian methods for change-point detection in long-range dependent processes," *Journal of Time Series Analysis*, vol. 23, no. 6, pp. 687–705, 2002.

[26] J. Sethuraman, "A constructive definition of dirichlet priors," *Statistica Sinica*, vol. 4, pp. 639–650, 1994.

[27] M. Liu and R. Siegwart, "DP-FACT: Towards topological mapping and scene recognition with color for omnidirectional camera," in *Proc. of the IEEE International Conference on Robotics and Automation (ICRA)*, 2012.

[28] V. Moskvina and A. Zhigljavsky, "An algorithm based on singular spectrum analysis for change-point detection," *Communications in Statistics-Simulation and Computation*, vol. 32, no. 2, pp. 319–352, 2003.