PAM-based Behavior Modelling

Thien Huynh-The[1], Ba-Vui Le[1], Muhammad Fahim[1], Sungyoung Lee[1], Yongik Yoon[2], and Byeong Ho Kang[3]

[1] Department of Computer Engineering,
Kyung Hee University, Gyeonggi-do, 446-701, Korea
{thienht,lebavui,fahim,sylee}@oslab.khu.ac.kr
[2] Department of Multimedia Science,
Sookmyung Women University, Korea
yiyoon@sookmyung.ac.kr
[3] Department of Computing and Information Systems,
University of Tasmania, Australia
byeong.kang@utas.edu.au

Abstract. A novel approach for human behavior modelling is represented in this paper based on the Pachinko Allocation Model (PAM) algorithm for the video-based road surveillance. In particular, the authors focus on the behavior analysis and modelling for learning and training as the main distribution of this research. Sparse object features in sequence of frames are modelled into activities and behaviors with full topic correlations to avoid omissions of small activities.

Keywords: human behavior modelling, CCTV system, pachinko allocation model, video-based road surveillance.

1 Introduction

In recent years, Human Behavior Analysis via the CCTV systems has become an interesting field, however, achieving high performance of recognition is not an easy task, especially in the real-time environment. The recognition performance by classifying the new behavior based on existing models depends on the modelling. Therefore, the correlation of a behavior and its class would be described via a probabilistic model. As a simple Dynamic Bayesian Network (DBN) [1], [2], the Hidden Markov Models (HMMs) [3], [4], have become the powerful tool for activity modelling. However, they are usually sensitive to noise or input errors which are the reasons for low recognition rate. Therefore, these shortcomings have motivated recent approaches to apply topic models as the effective and novel solution. These approaches such as Latent Dirichlet Allocation (LDA) [5], Dual Hierarchical Dirichlet Process (DHDP) [6], can define the relationship between each atomic activity with its corresponding behavior through the probabilistic model. However, they can not represent relationships fully, especially topic-topic and topic-word, thus missing or incorrect classification can occur in both the modelling and recognition stage. In this paper, the authors propose a modelling approach using the Pachinko Allocation Model [7] to solve existing problem of previous approaches.

2 A PAM-based behavior modelling

After achieving the object features as position and direction, they need to be modelled into sparse activities. The behavior is the collection of atomic activities which are considered in the sequence of frames. In this paper, PAM is proposed to capture not only correlations among activities but also correlations among behaviors themselves. As a special structure, a four-level hierarchy PAM consists one root behavior, s_1 behaviors at the second level $T = \{t_1, t_2, \ldots, t_{s_1}\}$, s_2 activity groups at the third level $T' = \{t'_1, t'_2, \ldots, t'_{s_2}\}$ and N features at the bottom. In PAM, behaviors are fully associated to activity groups which are then connected to features. The Fig. 1 shows the hierarchical topic model (a) and the graphic model (b) of PAM for behavior modelling. The multinomial of the root θ_r^d and behaviors $\theta_{t_i}^d$ are sampled from the Dirichlet distribution $g_r(\alpha_r)$ and $g_i(\alpha_i)|_{i=1}^{s_1}$, respectively, where d is a matrix containing features of a number of frames. A long clip D presenting for a certain behavior will be divided into n small clips $D = \{d_1, d_2, \ldots, d_n\}$. Meanwhile the activity group is modelled with fixed multinomial distributions $\phi_{t'_j}\big|_{j=1}^{s_2}$ and $\psi_{t'_j}\big|_{j=1}^{s_2}$ which are sampled from Dirichlet distributions $g(\beta)$ and $g(\gamma)$, respectively. For each small clip:

1. Derive a multinomial distributions θ_r from α_r.
2. For each behavior, derive s_1 multinomial distributions θ_{t_i} from α_i.
3. Derive s_2 multinomial distributions $\phi_{z'}$ from β and $\psi_{z'}$ from γ for each activity group z'. For kth feature with location p_k and direction q_k in d:
   (a) Derive a behavior z_k and a activity group z'_k from θ_r and θ_z
   (b) Derive a location p_k and direction feature q_k from $\phi_{z'}$ and $\psi_{z'}$.

The hyper-parameters as Dirichlet priors α, β, and γ can be estimated via the Gibbs sampling [7]. The marginal probability of a small clip as:

$$P(d|\,\alpha,\beta,\gamma) = \int P(\theta_r|\,\alpha_r)\prod_{i=1}^{s_1} P(\theta_{t_1}|\,\alpha_i) \prod_k \sum_{z_k, z'_k} (P(z_k|\,\theta_r)\,P(z'_k|\,\theta_z)\,P(f_k|\,\phi_{z'},\psi_{z'}))d\theta^{(d)} \quad (1)$$

Finally, the probability of generating D is computed as:

$$P(D|\,\alpha,\beta,\gamma) = \int \prod_{j=1}^{s_2} \left(P\left(\phi_{t'_j}\middle|\,\beta\right) + P\left(\psi_{t'_j}\middle|\,\gamma\right)\right) \prod_d P(d|\,\alpha,\beta,\gamma)\,d\phi d\psi \quad (2)$$

In order to recognize, the Support Vector Machines (SVMs) is chosen to train the model derived from the behaviors and activity groups with labels.

3 Experimental results

The QMUL data set [8] was used for evaluation with non-overlapping clips which run at 30fps in frame rate and 360×288 in frame resolution. The QMUL

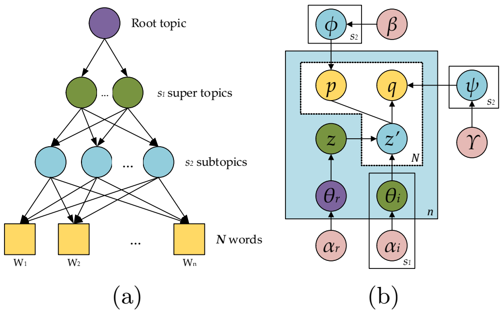

Fig. 1. Pachinko Allocation Model: (a) Hierarchical topic model (b) Graphic model.

dataset contains 108 clips (72 for training and 36 for testing) presenting 2 behaviors: vertical and horizontal traffic. The 4-second clips will be generated from a long clip for assessment. The Fig. 2 represents the detected activities in some samples as small clips which will be modelled for different probabilistic models. Some activities can appear in clips of both behaviors, therefore, the decision of class is employed based on a trained structure of SVM with the highest correlation. Moreover, the evaluation is performed though recall and precision value and compared with an approach using the LDA. The detail results have been shown in the Table. 1. In both recall and precision results, PAM is better than LDA due to subtopic as activity group layer. However, taking more computation time in learning and recognition is the trade-off of this approach. Thus, this limitation can be reduced by discarding reduplicated models for the training stage to ensure the real-time results.

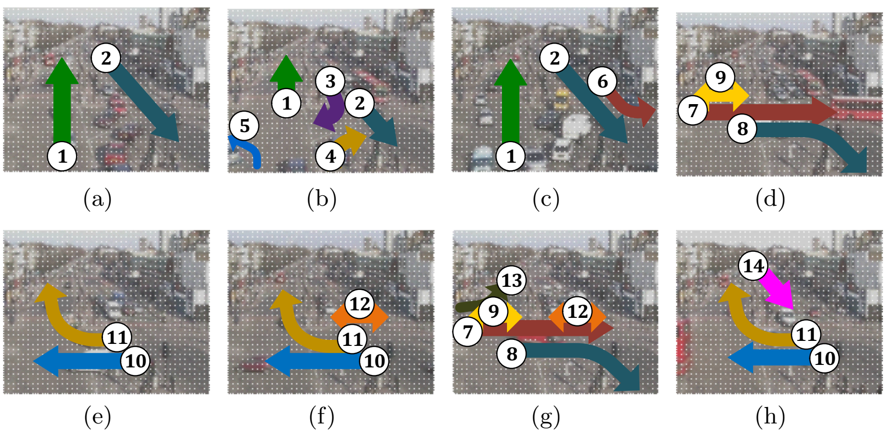

Fig. 2. Activities have been detected in PAM-modelling for training. (a)-(c): 3 models of the vertical traffic behavior. (d)-(h): 5 models of the horizontal traffic behavior.

Table 1. Confusion matrix of the SVM classifier

	PAM			LDA		
Behavior	Vertical	Horizontal	Recall	Vertical	Horizonta	Recall
Vertical	53	7	0.883	49	11	0.817
Horizontal	7	48	0.873	8	47	0.855
Precision	0.883	0.873		0.860	0.810	

4 Conclusion

In this paper, we proposed the human behavior modelling based on the four-level hierarchy PAM, in which, the probabilistic data presenting the relationship of features and behaviors will be generated for training stage. The PAM captures correlations among words as features and among topics as activities or behaviors, thus the errors in modelling will be limited to improve the recognition rate. The model was also evaluated and compared with the LDA approach. In the future, we need to improve the recognition performance by filtering highest characteristic features for training and optimizing the SVM classifier.

Acknowledgments. This research was support by the MSIP (Ministry of Science, ICT&Future Planning), Korea, under the ITRC (Information Technology Research Center) support program supervised by the NIPA (National IT Industry Promotion Agency) (NIPA-2014-(H0301-14-1003)).

References

1. Oliver, N., Rosario, B., Pentland, A.: A bayesian computer vision system for modeling human interactions. IEEE Trans. Pattern Anal. Mach. Intell., vol. 22, no. 8, pp. 831–843, Aug 2000.
2. Xiang, T., Gong, S.: Video behavior profiling for anomaly detection. IEEE Trans. Pattern Anal. Mach. Intell., vol. 30, no. 5, pp. 893908, May 2008.
3. Brand, M., Kettnaker, V.: Discovery and segmentation of activities in video. IEEE Trans. Pattern Anal. Mach. Intell., vol. 22, no. 8, pp. 844–851, Aug 2008.
4. Xiang, T., Gong, S.: Beyond tracking: Modelling activity and understanding behaviour. Int J. Comput. Vis., vol. 67, no. 1, pp. 21–51, Apr 2006.
5. Zhao, L., Shang, L., Gao, Y., Yang, Y., Jia, X.: Video behavior analysis using topic models and rough sets [applications notes]. IEEE Comput. Intell. Mag., vol. 8, no. 1, pp. 56–67, Feb 2013.
6. Wang, X., Ma, X., Grimson, W.: Unsupervised activity perception in crowded and complicated scenes using hierarchical bayesian models. IEEE Trans. Pattern Anal. Mach. Intell., vol. 31, no. 3, pp. 539–555, March 2009.
7. Li, W., McCallum, A.:Pachinko allocation dag-structured mixture models of topic correlations. In: 23th International Conference on Machine Learning, pp. 557–584, ACM Press, New York (2006).
8. QMUL Junction Dataset, `http://www.eecs.qmul.ac.uk/~ccloy/downloads_qmul_junction.html`