

# Spatial Activity Recognition in a Smart Home Environment using a Chemotactic Model

#Daniel E. Riedel, Svetha Venkatesh, and Wanquan Liu  
Department of Computing, Curtin University of Technology,  
GPO Box U1987 Perth, Western Australia 6845  
{riedelde,svetha,wanquan}@cs.curtin.edu.au

## Abstract

*Spatial activity recognition is challenging due to the amount of noise incorporated during video tracking in everyday environments. We address the spatial recognition problem with a biologically-inspired chemotactic model that is capable of handling noisy data. The model is based on bacterial chemotaxis, a process that allows bacteria to change motile behaviour in relation to environmental gradients. Through adoption of chemotactic principles, we propose the chemotactic model and evaluate its performance in a smart house environment. The model exhibits greater than 99% recognition performance with a diverse six class dataset and outperforms the Hidden Markov Model (HMM). The approach also maintains high accuracy (90-99%) with small training sets of one to five sequences. Importantly, unlike other low-level spatial activity recognition models, we show that the chemotactic model is capable of recognising simple interwoven activities.*

## 1. INTRODUCTION

Non-invasive activity recognition is an inherently difficult problem. Many issues arise due to the infinite number of spatial and temporal variations within and between activities, intrinsic noise from visual tracking systems and activity interweaving. Finding viable solutions to the activity recognition problem is becoming increasingly important to develop systems capable of supporting our aging population and minimising our reliance on carer and nursing facilities. Unfortunately, existing approaches to the activity recognition problem only address subsets of the overall problem, requiring the need for new models.

Current activity recognition approaches can be broadly classified into two categories: bottom-up and top-down. Bottom-up approaches use low-level data to develop simple models of activities and involve template matching or probabilistic models (HMMs [1] and extensions to the HMM). Template matching techniques are slow and sensitive to noise in observations and variation in patterns of the same activity [2]. Probabilistic models used in bottom-up approaches are able to better deal with uncertainty, yet do not scale well with large training sets and long activity sequences. Overall, bottom-up models do not allow complex activity recognition as the techniques employ flat models of activities. In contrast, top-down approaches recognise complex, semantically rich

activities. This is achieved through the use of plan recognition techniques including dynamic bayesian networks (DBNs) [3] and stochastic grammars [4], [5]. These techniques have an average non-polynomial time complexity, preventing their use in practical situations or requiring the need for simplified and less expressive models for tractability. The significant disadvantages presented by current bottom-up and top-down approaches further demonstrates the need for exploration of alternate model formulations.

Chemotaxis is a process that allows motile bacteria such as *Escherichia coli* and *Salmonella typhimurium* to directionally swim in response to chemical or other physical gradients [6]. The process acts by either attracting cells in the direction of an increasing gradient and possibly towards nutrients or repelling organisms from harmful regions by moving in the direction of decreasing gradients [7]. In a uniform and static environment, cells carry out a random walk by alternating tumbles (changing direction) and running (going forward in a straight line). The duration of tumbles and runs are exponentially distributed, with the mean duration of runs being approximately ten times longer than that of tumbles, allowing cells to “walk” [8]. In the presence of an increasing favourable gradient, bacteria decrease the tumbling frequency and increase the run length allowing organisms to move towards an attractant source [9].

Bacterial cells interact with the environment through receptors present on the cell surface. Binding of matching molecules to the receptors produces a signal that allows cells to respond to an input stimulus [6]. The mapping process, termed signal transduction, integrates extracellular signals, translating them into a series of intracellular structural or chemical changes, which through a series of enzyme catalysed reactions, changes cellular production levels or function [7]. In the case of *Escherichia coli* bacterial chemotaxis, the resulting change of function is an increase in the duration of clockwise flagella<sup>1</sup> rotation, which results in the observed longer run movement [8].

Studies on *Escherichia coli* have shown that bacteria sense spatial gradients as temporal changes in attractant or repellent concentration. The finding indicates the presence of an intracellular short term memory, that allows cells to remember previous spatial concentrations for comparison to current levels [10]. It is believed that chemotaxis confers an evolutionary

<sup>1</sup>Flagella are whip-like appendages that provide locomotion

advantage to bacterial species that possess the characteristic, allowing them to better survive and respond to changes in dynamic environments.

This paper focuses on recognising activities of similar length captured by a video tracking system and within the constraints of a smart house environment. Consequently, activities are assumed to be spatially similar and carried out in a periodic manner. To perform spatial activity recognition a biologically-inspired paradigm is explored. Biological systems have proven to be a useful basis for solving many real world problems as a consequence of their innate robustness, adaptiveness, diversity and error tolerance [11]. Our research uses the biological concept of chemotaxis to address the uncertainty and interweaving issues of the activity recognition problem.

Development and use of a bacterial chemotactic activity recognition model is novel. Bacteria, that exist in competitive and nutrient poor environments, evolved the chemotactic capability to sense and respond to dynamic environments. Through use of the chemotactic paradigm, we have developed a model that is capable of dealing with noise from video tracking systems, resulting in a higher recognition performance than the HMM. Unlike traditional HMM-based techniques, the model does not adhere to the Markovian assumption which reflects the true nature of real world activities. In addition, our solution can function adequately with small numbers of training sequences, can easily incorporate long activity sequences and has a lower training complexity compared to most geometric, plan-based and HMM-based approaches.

Current solutions to the activity recognition problem including template matching, probabilistic inferencing, stochastic planning or combinations are disadvantaged with rigid environmental constraints, poor robustness and/or poor scalability. Our chemotactic model is robust to spatial variations in activity sequences of a constant length, unlike template-based techniques. Due to the lower computational complexity of the approach, the chemotactic technique is also more scalable to larger environments than HMM-based methods. An additional, yet significant characteristic of the chemotactic model is its ability to cater for the interweaving of activities. This important and complex issue has not been adequately addressed in other activity recognition models.

The paper is organised as follows. An overview of the chemotactic and HMM models are provided in sections 2 and 3 respectively. Experimental methodology, including collection of activity sequences, training procedures and evaluation techniques are presented in section 4. Section 5 describes results from recognition performance analyses of the model in comparison to a HMM-based approach and evaluation of the model's ability for recognition of interweaved activities. A summary is presented in section 6.

## 2. CELLULAR CHEMOTACTIC MODEL

The cellular chemotactic model uses an environmental and cellular abstraction of chemotaxis to represent and recognise activities. An activity consists of a group of cells, where the cells model the movement of individual bacteria in response

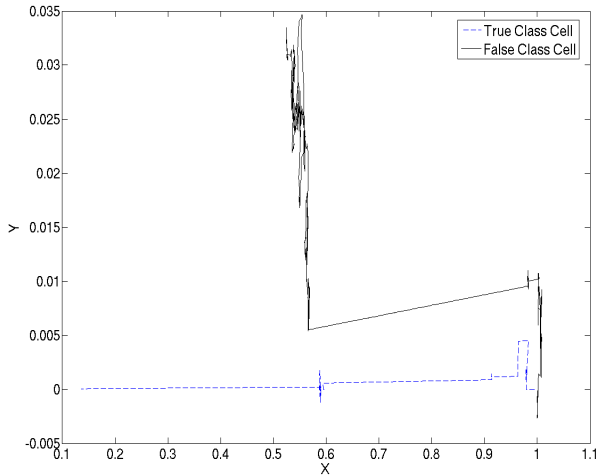
to environmental dynamics. A cell is composed of receptor types  $\{R_i\}_{i=1}^n$  that function to match *molecules* from the *environment*. A *molecule* is a symbol  $u$ , where  $u \in U$  and  $U$  is the set of all possible spatial symbols, for example  $U = \{1, 2, 3, \dots, 156\}$ . Each receptor type has a specified number of receptors denoted by  $|R_i|$ . The total number of receptors of a cell is given by  $p$ , where  $p = \sum_i |R_i|$ . Cells have an  $x, y$  tuple that determines the cell position within the two dimensional *environmental* space  $E$ . These values are initially set to 1.0 and 0.0, respectively and represent the starting position of the cells. In the model, the attractant source or the place where *molecules* are conceptually released, is set to the origin of the coordinate space. Cellular running times for cells with and without environmental gradients, explained later, are represented by exponential distributions with means of  $\mu_{LR}$  and  $\mu_R$ , where  $\mu_{LR} > \mu_R$ .  $\mu_{LR}$  is the exponential mean of the distribution for a long running movement, whilst  $\mu_R$  is the exponential mean for the normal running motion. Movement of the cell within the coordinate space is determined by the velocity  $v$  of the cell (1). Velocity is normalised against the sequence length and  $\mu_{LR}$ , to account for variation in intraclass sequence length.

$$v = p \times \mu_{LR} . \quad (1)$$

In normal biological settings, release of molecules into a fluidic environment results in formation of a gradient that dissipates over time. This gradient is traversed by bacterial cells in order to locate an attractant source, thus increasing the fitness and survivability of the organism. In our model, addition of  $u$  molecules to  $E$ , results in an increase in the environmental concentration of  $u$ . The increase in environmental concentration of  $u$  is detected by cells with a free receptor of the same receptor type. Typically, chemotactic bacteria would then make a series of random moves with a bias towards the attractant source, re-evaluating environmental concentrations at each move. In our model, we know the area of highest concentration of molecules or the attractant origin, therefore, cells can easily determine the direction of travel and move towards the attractant if and only if receptor types match, free receptors are available and the cell is not in a region of high concentration.

Chemotactic cells in the model detect increasing environmental concentrations via a memory associated with the irreversible binding of molecules to receptors. We represent the cellular memory through a histogram approach with selected histogram bins representing the receptor types. Each histogram bin or receptor type has a corresponding fixed maximum bin frequency. The maximum bin frequency describes the number of receptors  $|R_i|$  of a given receptor type  $R_i$  that a cell possesses. When a molecule  $u$  is released into the environment  $E$ , the environmental concentration of that chemical increases. Cells that have a matching receptor type for the molecule then check if any of the particular receptors are free. If so, the molecule binds and the cells behaviour is modified by changing direction towards the attractant origin and increasing the length of the running movement from  $\mu_R$  to  $\mu_{LR}$ . If

a cell does not have a receptor for that particular molecule or the cell does have a corresponding receptor type but no free receptors, then the cell performs a random walk with a randomly selected direction over a distance of  $\mu_R$ . When cells move close to the attractant source and the euclidean distance  $d$  between the cell and origin is less than the high concentration threshold  $\theta_{high}$ , the cells perform random walks irrespective of increasing environmental concentrations. If the cells move outside the high concentration area  $d > \theta_{high}$ , then the cells return to normal behaviour. Motile bacteria use the corresponding change in behaviour at high environmental concentrations to prevent being restricted to local regions of high attractant concentrations. We instead utilise the characteristic as a tolerance mechanism for sequence expansion. Figure 1 illustrates the behaviour of true and false class cells in response to a test sequence. From Fig. 1 it is obvious that true class cells with similar patterns to the test sequence exhibit more straight line movement towards the attractant source compared to false class cells. Therefore, cells with higher degrees of similarity to test sequences will end up closer to the attractant source. From a qualitative perspective, areas of sequence similarity may also be heuristically identified through visualisation of stretches of straight line movement, that is before, after or between regions of random walking.



**Fig. 1:** Chemotactic Cell Movements of true and false class cells. Movement is in the direction of the origin 0,0.

After all molecules representing an activity sequence are “released” into the chemotactic environment consisting of  $\beta$  classes with  $m$  cells per class, we find the cell  $\phi$  in  $Z$ , where  $Z$  is the set of all activity cells, which has minimum euclidean distance to the attractant source  $\delta$  of  $E$  according to (2).

$$\phi = \underset{g \in Z}{\operatorname{argmin}} d(g, \delta) . \quad (2)$$

The minimum distance cell  $\phi$  is then used in the classification decision.

### 3. HIDDEN MARKOV MODEL

A Hidden Markov Model (HMM) is a form of stochastic state transition model that is capable of dealing with time sequential data [1]. It was first applied to the activity recognition domain, through the work of [12] and since has been utilised extensively, particularly through variants of the basic HMM. The HMM is popular in automatic speech recognition, bioinformatics and activity recognition approaches due to its innate ability to deal with noisy observations and its high discrimination properties.

In this study, we utilise the discrete HMM to recognise spatial activity patterns. The smart house environment is discretised into  $1 \times 1$  metre states, with each activity sequence being mapped to a sequence of these states, comprising a total of 156 possible spatial states. A discrete HMM is characterised by the number of hidden states in the model  $N$ , the number of distinct observation symbols per state  $M$ , the state transition probability matrix  $A$  ( $A = \{a_{ij}\}$ ), the observation symbol probability distribution matrix  $B$  ( $B = \{b_j(k)\}$ ) and the initial state distribution vector  $\pi$ . A derived HMM  $\lambda$  is typically represented by the tri-tuple of parameters  $\{\pi, A, B\}$  which are defined as follows:

$$\begin{aligned} \pi &= Pr(q_1 = S_j), \quad 1 \leq i \leq N \\ a_{ij} &= Pr(q_{t+1} = S_j | q_t = S_i), \quad 1 \leq i, j \leq N \\ b_j(k) &= Pr(v_k \text{ at } t | q_t = S_j), \quad 1 \leq j \leq N, 1 \leq k \leq M \end{aligned}$$

where  $q_t$  is the state at time  $t$ ,  $S$  is the individual states such that  $S = \{S_1, S_2, \dots, S_N\}$  and  $V$  denotes the individual symbols  $V = \{v_1, v_2, \dots, v_M\}$ . Parameters  $\pi, A, B$  are derived using the Baum-Welch (Forward-Backward) algorithm, however scaling [1] is used in both the model estimation and inferencing, due to the use of lengthy observation sequences. The training complexity for the HMM with Baum-Welch parameter estimation is  $O(hTN^2)$ , where  $T$  is the length of the  $h$  observation sequences, whilst the complexity of determining the probability of an observed sequence of length  $T$  with a HMM  $\lambda$  and using the scaled forward procedure is  $O(TN)$ .

### 4. METHODOLOGY

Activity sequences were collected in the Smart House environment using the multiple camera tracking system of [13]. The data set comprises six single person activities: *get home-watch TV*, *have a snack-watch TV*, *at home-watch TV*, *reading newspaper*, *having breakfast-toast* and *having breakfast-eggs*. Each of the activity sequences were approximately ninety seconds in length, with subjects entering the environment from the north/south doors of Room 1 or the north door of Room 2 and carrying out the specified movements. Attempts were made to minimise the intraclass spatial variations during the capture procedure, however, the tracking system did incorporate some spatial variability. Once obtaining two dimensional tracking data from the video tracking system, we converted the representation into a one dimensional symbolic form. To do this the smart house environment was separated into one metre

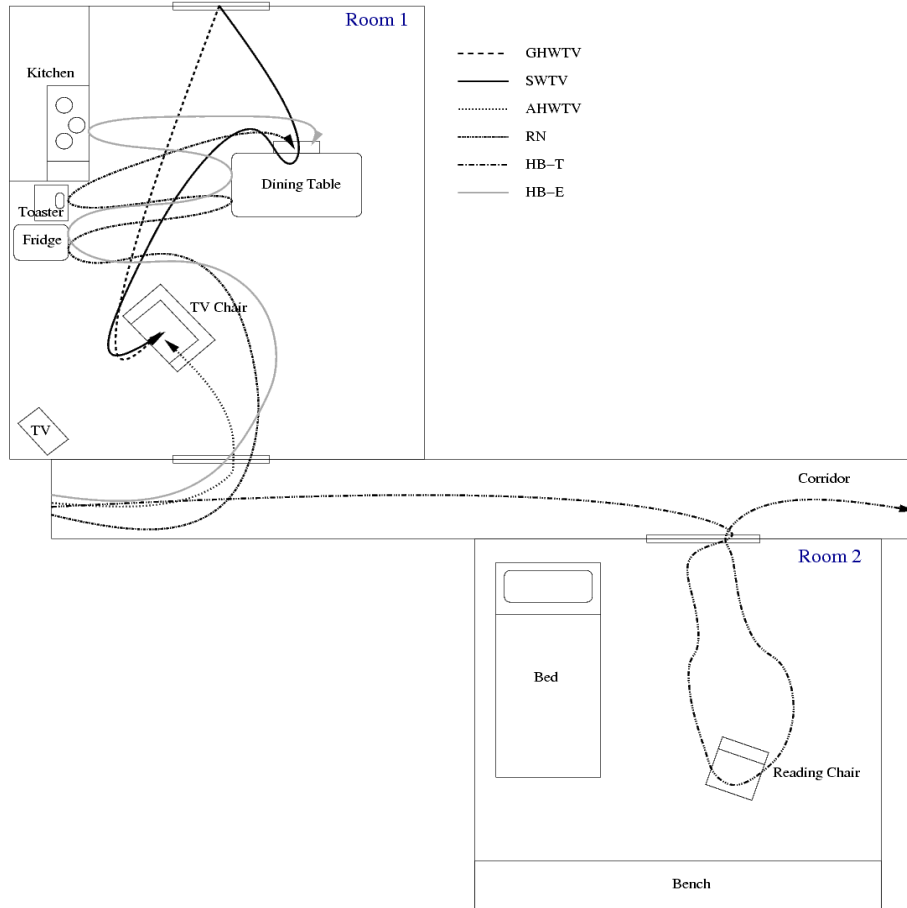


Fig. 2: Spatial paths for each of the six classes of activity sequences.

grids with unique integers  $u$  (where  $u \in U$ ) being assigned to each grid position. Each raw tracking coordinate consisting of an  $x, y$  position was then translated into an integer symbol  $u$ , for the whole activity sequence. The characteristics of the activity classes can be found in Table 1. Figure 2. provides a graphical layout of the Smart House area and the paths taken for the six classes of activities.

Following encoding of the activity sequences into symbolic form, supervised classification was performed. The six classes of activities were separated into training and testing sets of size ten. For each of the training class sets, histograms were generated with bin sizes of one. From the histograms, those bins with frequencies greater than one were represented as receptors  $R_i$ , with the number of receptors  $|R_i|$  being the histogram bin frequency. Cell velocity was derived for each cell using (1). To quantify the recognition performance of the model with the testing sets, simulations coupled with cross-validation were performed obtaining precision and recall statistics. Precision measures the ability of the technique to correctly classify, whilst recall measures the completeness of a technique's classification, that is the proportion of the true class test cases that were identified [14].

To determine the effectiveness of the approach we compared the chemotactic model to the HMM (where  $N = 5$ ,  $M =$

156 and  $\iota = 20$ ) and built consequent HMMs for each activity class. The ability of our chemotactic model to function adequately with few training sequences was also evaluated. In addition we show how the model can be used to recognise simple interwoven activities.

## 5. EXPERIMENTAL RESULTS

All experiments were performed using the following cell parameters:  $\mu_R=1.0$ ,  $\mu_{LR}=1.5$ ,  $\theta_{high}=0.05$  and  $U = \{1, 2, 3, \dots, 156\}$ , which is the set of spatial integers used for the encoding.

### A. Evaluation of Recognition Performance

To evaluate the performance of the chemotactic approach, cells were developed using the approach of Section 4. To compare the performance of the model, we compare the maximum recognition rates to that of the HMM, which is also constructed from encoded activity sequences. The maximum resulting precision and recall rates for chemotactic and HMM models were obtained with ten training sequences and are shown in Table. 2. Both precision and recall rates for the chemotactic model showed an improvement over the HMM. The observed higher recognition performance of our model is likely the result of the chemotaxis process better accounting for noise

**TABLE 1:** COMPOSITION OF THE SMART HOUSE ACTIVITY RECOGNITION DATA SET

Activity Name	Activity Description
Get Home-Watch TV (GHWTV)	Walk through Room1 North door and sit on TV chair
Have a Snack-Watch TV (SWTV)	Walk through Room1 North door, sit down at dining table to eat snack and then move to sit on TV chair
At Home-Watch TV (AHWTV)	Walk through Room1 South door and sit down on TV chair
Reading Newspaper (RN)	Walk down corridor, enter Room2, sit down on bed chair, read for a while, then leave
Have Breakfast-Toast (HB-T)	Walk though Room1 South door, go to fridge for OJ, put OJ on dining table, toast bread, eat at dining table
Have Breakfast-Eggs (HB-E)	Walk though Room1 South door, go to fridge for OJ, put OJ on dining table, cook eggs, eat at dining table

**TABLE 2:** MAXIMUM PRECISION AND RECALL RATES FOR THE CHEMOTACTIC AND HMM ACTIVITY RECOGNITION MODELS.

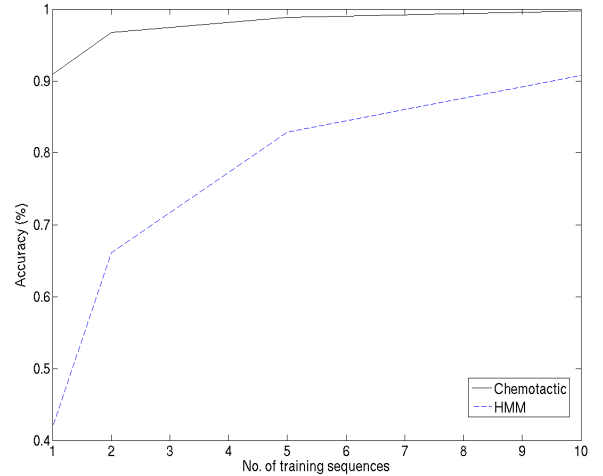
Technique	Precision (%)	Recall (%)
Chemotactic Model	99.75	100.00
HMM	90.75	90.75

through its random walk motions in the absence of sequence correspondence. Even though our model does not enforce sequential consistency, it still exhibits good discrimination properties. We also analysed the effect of manipulating the means of the exponential distributions governing the differing running actions. When  $\mu_{LR} \gg \mu_R$ , no discernible change in recognition performance was noted. In addition, the random walks performed by cells in the absence of gradients was negligible and cells only moved significantly in relation to receptor matches. When  $\mu_{LR}$  approached  $\mu_R$  a small decrease in recognition was observed. We attribute the decrease to a lack of disparity between the matching and non-matching states of the cell.

### B. Number of training sequences versus recognition performance

To analyse the effect of increasing numbers of training sequences with respect to recognition performance, we trained chemotactic and HMM models with 1,2,5 and 10 training sequences from the six activity classes. Figure 3. demon-

strates the effect of increasing training sequence numbers in relation to accuracy. As noted from Fig. 3, the chemotactic



**Fig. 3:** No. of training sequences versus accuracy for chemotactic and HMM models.

model is able to recognise activities with significantly higher accuracy than HMM, especially with smaller numbers of training sequences. We attribute the chemotactic models good performance to the generalisation capability of the underlying histogram approach. The observed poor performance of the HMM with few training sequences is expected as the discrete HMM is unable to accurately determine the probabilities of state transitions with small training sets, thus resulting in poor recognition performance.

### C. Recognition of interwoven activities

During the course of experimentation, we identified a unique capability of the chemotactic model; its ability to recognise interwoven activities. To test our hypothesis we constructed an alternate data set that consisted of two classes of activities as shown in Fig. 4. The first activity is synonymous to SWTV in Section 4., whilst the second activity starts at the dining table, involves some cooking and then returning to the dining table. We train chemotactic cells with both classes of activities and use an interwoven sequence that involves entering from the north door, having a snack, cooking, returning to the dining table to eat and then proceeding to watch tv. The interwoven sequence thus incorporates the beginning section of activity 1, transfers to activity 2 and then returns to complete activity 1. The overall length of the interwoven sequence is the sum of the lengths of activity 1 and 2. Figure 5. shows the resulting cellular movements of both classes in response to the interwoven sequence. In Fig. 5a, a long random walk at approximately  $x = 0.85$  was evident, indicating interruption of activity 1. At the same time ( $x \approx 1.0$ ), the other activity cell in Fig. 5b was observed to change from a random walk to straight line movement towards the attractant origin (0,0) indicating similarity to the test sequence. The cell in Fig. 5b then reverted

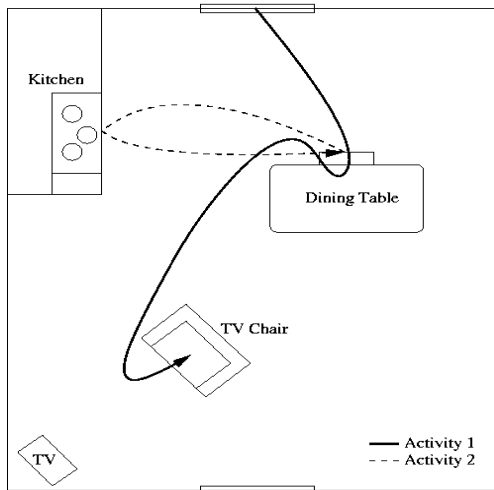


Fig. 4: Snack-Watch TV and Cooking spatial paths.

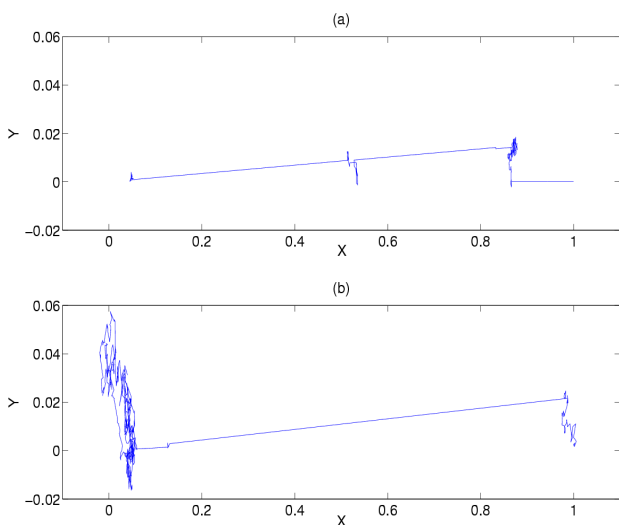


Fig. 5: Interwoven cell movements (a) Snack-Watch TV Cell (b) Cooking Cell.

back to a random walk at  $x = 0.05$ , with the opposite cell in Fig. 5a returning to straight line movement towards the attractant origin. As demonstrated in Figures 5a and 5b, the final positions of both the SWTV and Cooking cells were close to the attractant origin indicating that both activities had been recognised. We can therefore conclude that our chemotactic approach is able to recognise simple interwoven activities. The exhibited tolerance to interweaving is the result of the chemotactic cells performing smaller random walks in the absence of sequence similarity. Therefore, if an activity is in progress and a corresponding cell is moving in the direction of the attractant source, interruption of that activity will result in the cell performing a random walk. The random walk allows the cell to remain in the vicinity of where the activity was disrupted, until the activity is resumed. If an activity is disrupted for a long period of time, it is possible that cells

may wander away from the area of disruption thus reducing recognition performance. If one trains separate HMMs with the same two classes of activities and then uses the above interwoven sequence  $O$  for inferencing, the resulting  $Pr(O|\lambda)$  for both models will be very small. This occurs as the models do not observe particular states, found in the other class during training. Therefore, when calculating the sequence probability with a scaled forward procedure [1], some of the derived symbol probabilities for observed symbols are zero or very small, resulting in an overall low probability for the sequence.

## 6. CONCLUSION

This paper has presented a novel chemotactic model for online spatial activity recognition, that achieves higher recognition performance than the HMM. Evaluation of recognition performance over different training set sizes showed that the chemotactic approach was superior to the HMM, obtaining greater than 90% accuracy with the given dataset. Through relaxation of the sequential constraint that is inherent in most activity recognition techniques, we have been able to develop a model with low training and inferencing complexity, yet high recognition performance. The research also identified and evaluated a unique aspect of the chemotactic model, that is its ability to recognise simple interweaved activities.

## REFERENCES

- [1] Rabiner, L.R.: A tutorial on hidden markov models and selected applications in speech recognition. *Proceedings of the IEEE* **77** (1989) 257–286
- [2] Aggarwal, J., Cai, Q.: Human motion analysis: a review. *Computer Vision and Image Understanding* **73** (1999) 428–440
- [3] Intille, S., Bobick, A.: Representation and visual recognition of complex, multi-agent actions using belief networks. Technical report, M.I.T Media Laboratory Perceptual Computing Section (1998) No. 454.
- [4] Bobick, A., Ivanov, Y.: Action recognition using probabilistic parsing. In: *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. (1998) 196–202
- [5] Pynadath, D., Wellman, M.: Probabilistic state-dependent grammars for plan recognition. In: *Proceedings of the Sixteenth Annual Conference on Uncertainty in Artificial Intelligence*. (2000) 507–514
- [6] Adler, J.: Chemotaxis in bacteria. *Annu. Rev. Biochem.* **44** (1975) 341–356
- [7] Bourret, R., Stocks, A.: Molecular information processing: lessons from bacterial chemotaxis. *The Journal of Biological Chemistry* **277** (2002) 9625–9628
- [8] Berg, H.: Bacterial microprocessing. In: *Cold Spring Harbor Symposia on Quantitative Biology*. Volume 55. (1990) 539–545
- [9] MacNab, R., Koshland, D.: The gradient-sensing mechanism in bacterial chemotaxis. *Proc. Nat. Acad. Sci.* **69** (1972) 2509–2512
- [10] Segall, J., Block, S., Berg, H.: Temporal comparisons in bacterial chemotaxis. *Proc. Natl. Acad. Sci.* **83** (1986) 8987–8991
- [11] Paton, R., ed.: *Computing with biological metaphors*. Chapman & Hall, UK (1994)
- [12] Yamato, J., Ohya, J., Ishii, K.: Recognizing human action in time-sequential images using hidden markov model. In: *Computer Vision and Pattern Recognition*. (1992) 379–385
- [13] Nguyen, N., Bui, H., Venkatesh, S., West, G.: Recognising and monitoring high-level behaviours in complex spatial environments. In: *Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition (CVPR-03)*. (2003)
- [14] Vijay, V., Bollmann, B., Jung, G.: A critical investigation of recall and precision as measures of retrieval system performance. *ACM Trans. Info. Sys.* **7** (2003) 205–229