

Behavior Recognition Using any Feature Space Representation of Motion Trajectories

Shehzad Khalid
Bahria University
Pakistan

1. Introduction

In recent years, there has been a growth of research activity aimed at the development of sophisticated content-based video data management techniques. This development is now especially timely given an increasing number of systems that are able to capture and store data about object motion such as those of humans and vehicles. This has acted as a spur to the development of content-based visual data management techniques for tasks such as behavior classification and recognition, detection of anomalous behavior and object motion prediction. Behavior can obviously be categorized at different levels of granularity. In far-field surveillance, we are primarily interested in trajectory-based coarse motion description involving movement direction (right/left or up/down) and motion type (walking, running or stopping). These techniques are essential for the development of next generation 'actionable intelligence' surveillance systems.

Processing of trajectory data for activity classification and recognition has gained significant interest quite recently. Various techniques have been proposed for modeling of trajectory-based motion activity patterns and using the modeled patterns for classification and anomaly detection. Much of the earlier research focus in motion analysis has been on high-level object trajectory representation schemes that are able to produce compressed forms of motion data (Aghbari et al., 2003; Chang et al., 1998; Dagtas et al., 2000; Hsu & Teng, 2002; Jin & Mokhtarian, 2004; Khalid & Naftel, 2005; Shim & Chang, 2004). This work presupposes the existence of some low-level visual tracking scheme for reliably extracting object-based trajectories (Hu, Tan, Wang & Maybank, 2004; Vlachos et al., 2002). The literature on trajectory-based motion understanding and pattern discovery is less mature but advances using Learning Vector Quantization (LVQ) (Johnson & Hogg, 1995), Self-Organising Maps (SOMs) (Hu, Xiao, Xie, Tan & Maybank, 2004; Owens & Hunter, 2000), Hidden Markov Models (HMMs) (Bashir et al., 2006; 2005b), and fuzzy neural networks (Hu, Xie, Tan & Maybank, 2004) have all been reported. These approaches are broadly categorized into statistical and neural network based approaches.

In a development of trajectory-based motion event recognition systems, there are different questions that we need to answer before proposing or selecting a pattern modeling and recognition technique. These includes:

1. What is the feature space representation of trajectories?

2. What is the distribution of trajectories in a given feature space representation? Do we need to cater for complex shape distributions that may exist in a given motion pattern?
3. Do we expect to have a multimodal distribution of trajectories within a given pattern?

Most of the trajectory-based motion recognition system, as proposed in relevant literature (Hu, Xiao, Xie, Tan & Maybank, 2004; Hu, Xie, Tan & Maybank, 2004; Khalid & Naftel, 2005; 2006; Owens & Hunter, 2000) can operate only on feature space representation of trajectories that lies in Euclidean space with a computable mean. However, a survey of recent literature in the areas of motion feature computation for trajectory representation shows that most of the feature space representation are complex and do not lie in the Euclidean space (Bashir et al., 2006; 2005a;b; 2007; Hamid et al., 2005; Keogh et al., 2001; Xiang & Gong, 2006; Zhong et al., 2004). It is not possible to compute a mean representation of different trajectories using such complex feature spaces. They can therefore not be applied to complex feature spaces with incalculable mean. These approaches expect that the trajectories in a given motion pattern follow certain standard distribution such as Gaussian. They can not cater for multimodal complex shape distribution of trajectories within a given motion pattern which is expected in the presence of complex feature space representation of trajectories. The research presented in this chapter focuses on presenting a trajectory-based behavior recognition and anomaly detection system that have an answer to all of the above raised questions. The proposed approach does not impose any limitation on the representation of trajectories. It can operate using any trajectory representation in any feature space with a given distance function. The proposed approach can perform modeling, classification and anomaly detection in the presence of multimodal distribution of trajectories within a given motion pattern.

The remainder of the chapter is organized as follows: We review some relevant background material in section 2. In section 3, we present a framework of multimodal modeling of activity patterns using any feature space with a computable similarity function. A soft classification and anomaly detection techniques using multimodal m -Medoids model is presented in section 4. Comparative evaluation of currently proposed multimodal m -Medoids and previously proposed localized m -Medoids (Khalid, 2010a) based approach for activity classification and anomaly detection is presented in section 5. Experiments have been performed to show the effectiveness of proposed system for trajectory-based modeling, classification and anomaly detection in the presence of multimodal distribution of trajectories within a pattern, as compared to competitors. These experiments are reported in section 6. The last section summarizes the paper.

2. Background and related work

Motion trajectory descriptors are known to be useful candidates for video indexing and retrieval schemes. Variety of trajectory modeling techniques have been proposed to compute the feature for trajectory representation. Most of the techniques for learning motion behaviour patterns and recognition from trajectories use discrete point sequence vectors as input to a machine learning algorithm. Related work within the data mining community on representation schemes for indexing time series data is also relevant to the parameterisation of object trajectories. An object trajectory can be defined as a set of points representing the ordered observations of the location of a moving object made at different points in time. A trajectory can therefore be represented as a time series implying that indexing techniques

for time series are also applicable to motion data. For example, Discrete Fourier Transforms (DFT) (Faloutsos et al., 1994), Discrete Wavelet Transforms (DWT) (Chan & Fu, 1999), Adaptive Piecewise Constant Approximations (APCA) (Keogh et al., 2001), and Chebyshev polynomials (Cai & Ng, 2004) have been used to conduct similarity search in time series data. Previous work has also sought to represent moving object trajectories through piecewise linear or quadratic interpolation functions (Chang et al., 1998; Jeanin & Divakaran, 2001), motion histograms (Aghbari et al., 2003) or discretised direction-based schemes (Dagtas et al., 2000; Shim & Chang, 2001; 2004). Spatiotemporal representations using piecewise-defined polynomials were proposed by (Hsu & Teng, 2002), although consistency in applying a trajectory-splitting scheme across query and searched trajectories can be problematic. Affine and more general spatiotemporally invariant schemes for trajectory retrieval have also been presented (Bashir et al., 2003; 2004; Jin & Mokhtarian, 2004). The importance of selecting the most appropriate trajectory model and similarity search metric has received relatively scant attention (Khalid & Naftel, 2005).

Modeling of motion patterns using trajectory data to perform motion based behavior recognition and anomaly detection has gained significant interest recently. Various techniques have been proposed for modeling of motion activity patterns and using the modeled patterns for classification and anomaly detection. These approaches are broadly categorized into statistical and neural network based approaches. Almost all statistical approaches dealing with anomaly detection are based on modelling the density of training data and rejecting test patterns that fall in regions of low density. There are various approaches that use Gaussian mixture models to estimate the probability density of data (Brotherton et al., 1998; Roberts & Tarassenko, 1994; Yeung & Chow, 2002). Various techniques based on hidden Markov models (HMM) have also been proposed (Xiang & Gong, 2005; 2006; Zhang et al., 2005). (Yacoub & Black, 1999) and (Bashir et al., 2005b; 2007) have presented a framework for modeling and recognition of human motion based on a trajectory segmentation scheme. A framework is presented to estimate the multimodal probability density function (PDF), based on PCA coefficients of the sub-trajectories, using GMM. Different classes of object motion are modelled by a continuous HMM per class where the state PDFs are represented by GMMs. The proposed technique has been shown to work for sign language recognition. The proposed classification system can not handle anomalies in test data and can only classify samples from normal patterns. (Xiang & Gong, 2005; 2006) propose a framework for behavior classification and anomaly detection in video sequences. Natural grouping of behaviour patterns is learnt through unsupervised model selection and feature selection on the eigenvectors of a normalized affinity matrix. A Multi-Observation Hidden Markov Model is used for modelling the behaviour pattern. (Hu et al., 2006; 2007) and (Khalid & Naftel, 2006) models normal motion patterns by estimating single multimodal gaussian for each class. For anomaly detection in (Hu et al., 2006), the probability of a trajectory belonging to each motion pattern is calculated. If the probability of association of trajectory to the closest motion pattern is less than a threshold, the trajectory is treated as anomalous. In (Rea et al., 2004), a semantic event detection technique based on discrete HMMs is applied to snooker videos. (Zhang et al., 2005) propose a semi-supervised model using HMMs for anomaly detection. Temporal dependencies are modelled using HMMs. The probability density function of each HMM state is assumed to be a GMM. (Owens & Hunter, 2000) uses Self Organizing Feature Maps (SOFM) to learn normal trajectory patterns. While classifying trajectories, if the distance of

the trajectory to its allocated class exceeds a threshold value, the trajectory is identified as anomalous.

In our previous work (Khalid, 2010b), we have proposed m -Medoids based activity Modeling and Classification approach using low dimensional feature vector representation of trajectories in Euclidean Space (MC-ES). m -Medoids based approach models a pattern by a set of cluster centres of mutually disjunctive sub-classes (referred to as medoids) within the pattern. Once the m -Medoids model for all the classes have been learnt, the MC-ES approach performs classification of new trajectories and anomaly detection by checking the closeness of said trajectories to the models of different classes using hierarchical classifier. The anomaly detection module required specification of threshold which is used globally for all the patterns. However, this approach had unaddressed issues like manual specification of threshold for anomaly detection, identification of appropriate value of threshold for anomaly detection and anomaly detection of motion patterns with different scale and orientation which is used globally for all the patterns. These issues are addressed by a localized m -Medoids based approach (LMC-ES) as proposed in (Khalid, 2010a) which enables us to automatically select a local significance parameter for each pattern taking into consideration the distribution of individual patterns. LMC-ES can effectively handle patterns with different orientation and scale and has been shown to give superior performance than competitors including GMM, HMM and SVM based classifiers. However, there are still open issues (i) Modeling, classification and anomaly detection in the presence of multimodal distribution of trajectories within a pattern (ii) Soft classification in the presence of multimodal pattern distribution to minimize misclassification (iii) Modeling and classification in feature spaces for which we can not compute mean.

The contribution of this work is to present an extension of m -Medoids based modeling approach, wherein the multimodal distribution of samples in each pattern is represented using multimodal m -Medoids. An approach for multimodal model-based classification and anomaly detection is also presented. The presented mechanism is based on a soft classification approach which enables the proposed multimodal classifier to adapt to the multimodal distribution of samples within different patterns. The multimodal m -Medoids based modeling and classification is applicable to any feature spaces with a computable pairwise similarity measure.

3. Multimodal m -Medoids based modeling

Given a representation of trajectories in any feature space for a given motion pattern, we wish to model the underlying distribution of trajectories within a pattern using training data. A pattern is modeled by a set of cluster centers of mutually disjunctive sub-classes (referred to as medoids) within the pattern. The proposed modeling technique referred to as m -Medoids modeling, models the class containing n members with m medoids known *a-priori*. Modeling of pattern using multimodal m -Medoids approach in general feature space is a three step process, (i) identification of m medoids, (ii) computation of set of possible normality ranges for the pattern and (iii) selection of customized normality range for each medoid. The resulting models of identified patterns can then be used to classify new unseen trajectory data to one of the modeled classes or identify it as anomalous if it is significantly distant from all of the modeled pattern.

3.1 Step 1: Identification of m-Medoids

The algorithm for identification of medoids using finite dimensional features in general feature space with a computable similarity matrix is based on the affinity propagation based clustering algorithm (Frey & Dueck, 2007). Let $DB^{(i)}$ be the classified training samples associated to pattern i , the modeling algorithm comprises the following steps:

1. Form the affinity matrix $A \in R^{n \times n}$ defined by

$$A(a, b) = \begin{cases} \exp\left(\frac{-\text{dist}(s_a, s_b)}{2\sigma^2}\right) & \text{if } a \neq b \\ P(a) & \text{otherwise} \end{cases} \quad (1)$$

Here $s_a, s_b \in DB^{(i)}$, σ is the scaling parameter and $P(a)$ is the preference parameter indicating the suitability of sample a to be selected as an exemplar (medoid). We set $P(a)$ to the median of affinities of sample a with n samples. We use a dynamic value of σ which is set to be the 6th nearest neighbor of s_a to cater for variation in local distribution of trajectory samples.

2. Initialize availability matrix $\mathfrak{S}(a, b) = 0 \quad \forall a, b$. The entry $\mathfrak{S}(a, b)$ in availability matrix stores the suitability of trajectory s_b to be selected by trajectory s_a as its exemplar.
3. Update responsibility matrix \mathfrak{R} as

$$\mathfrak{R}(a, b) = A(a, b) - \max_{\forall c \text{ s.t. } b \neq c} \{\mathfrak{S}(a, c), A(a, c)\} \quad (2)$$

The entry $\mathfrak{R}(a, b)$ in the responsibility matrix reflects the accumulated evidence for how well-suited trajectory s_b to serve as an exemplar for trajectory s_a while taking into account other potential exemplar for trajectory s_a .

4. Update availability matrix \mathfrak{S} as

$$\mathfrak{S}(a, b) = \begin{cases} \min\{0, \mathfrak{R}(b, b) + \sum_{\forall c \text{ s.t. } c \neq a \wedge c \neq b} \{0, \mathfrak{R}(c, b)\}\} & \text{if } a \neq b \\ \sum_{\forall c \text{ s.t. } a \neq c} \max\{0, \mathfrak{R}(c, a)\} & \text{otherwise} \end{cases} \quad (3)$$

5. Identify the exemplar for each sample as

$$\check{\zeta}_a = \text{argmax}_b [\mathfrak{S}(a, b) + \mathfrak{R}(a, b)] \quad (4)$$

6. Iterate through steps 3-5 till the algorithm is converged or maximum number of learning iterations (t_{max}) is exceeded. The algorithm is considered to have converged if there is no change in exemplar identification for certain number of iterations ($t_{convergence}$).
7. If the number of exemplars identified are smaller than the desired number of medoids, set higher values of preference and vice versa. The algorithm is repeated till the desired number of exemplars are identified. An appropriate value of preference parameter, for identification of desired number of medoids, is searched using a bisection method.
8. Append exemplars $\check{\zeta}_a$ to the list of medoids $\mathbf{M}^{(i)}$ modeling the pattern i .

3.2 Step 2: Computation of possible normality ranges

After the identification of medoids $\mathbf{M}^{(i)}$ for pattern i , we intend to identify and pre-compute a set of possible normality ranges for a given pattern. Values of normality ranges for a given

pattern is determined by the inter-medoid distances within a given pattern. Hence, different patterns will have different set of possible normality ranges depending on the distribution of samples, and in turn medoids, within a pattern. In this step, a set of possible normality ranges $\mathbf{D}^{(c)}$ for the pattern c is computed as follows:

1. Identify the closest pair of medoids (i, j) (indexed by (p, q)) from $\mathbf{M}^{(c)}$ as follows:

$$(p, q) = \arg \min_{(i, j)} \text{dist}(M_i, M_j) \quad \forall i, j \wedge i \neq j \quad (5)$$

where $\text{dist}(\cdot, \cdot)$ is the distance function for a given feature space representation of trajectories.

2. Set $l = 1$.
3. Populate the distance array at index l using

$$\mathbf{D}_l^{(c)} = \text{dist}(M_p, M_q) \quad (6)$$

4. Remove the closest pair of medoids using

$$\mathbf{M}^{(c)} = \mathbf{M}^{(c)} - \{M_p, M_q\} \quad (7)$$

5. Set $l = l + 1$.
6. Iterate through steps 1-5 till there are no medoids left in $\mathbf{M}^{(c)}$.

3.3 Step 3: Selection of customized normality range for each medoid

After the identification of medoids and a set of possible normality ranges for a given pattern, we select different normality range for each medoid depending on the distribution of samples from the same and different patterns around a given medoid. The normality range is selected to minimize false positives (false identification of training samples from other patterns as a normal member of pattern that is being modeled) and false negatives (classification of normal samples of the pattern being modeled as anomalies). The algorithm for selection of customized normality range for each medoid, to enable multimodal m -Medoid based modeling of pattern, comprises of following steps:

1. Initialize significance parameter τ with the number of possible normality ranges for pattern c as computed in Step 2.
2. Sequentially input labeled training instances belonging to all classes and identify the closest medoid, indexed by r , using:

$$r = \arg \min_k \text{dist}(Q, M_k) \quad \forall k \quad (8)$$

where Q is the test sample.

3. Perform an anomaly test using the anomaly detection system, as proposed in section 4, assuming a one class classifier containing only pattern c represented by medoids set $\mathbf{M}^{(c)}$ using the current value of τ .
4. Increment false positive count $FP(r)$, corresponding to closest medoid M_r , each time when the sample is a normal member of pattern c but is identified as anomalous.

5. Increment false negative count $FN(r)$, corresponding to closest medoid M_r , each time when the sample is misclassified to pattern c .
6. Iterate through steps 2-5 for all the samples in DB .
7. Calculate Significance Parameter Validity Index ($SPVI$) to check the effectiveness of current value of τ for a particular medoid using:

$$SPVI(k, \tau) = \beta \times FP(k) + (1 - \beta) \times FN(k) \quad 0 \leq \beta \leq 1 \quad \forall k \quad (9)$$

where β is a scaling parameter to adjust the sensitivity of proposed classifier to false positives and false negatives according to specific requirements.

8. Set $\tau = \tau - 1$.
9. Iterate through step 2-8 till $\tau = 1$.
10. Identify the value of significance parameter for a given medoid as:

$$\widehat{\tau}_{(c,k)} = \arg \min_{\tau} SPVI(\tau, k) \quad \forall M_k \in \mathbf{M}^{(c)} \quad (10)$$

where $\widehat{\tau}_{(c,k)}$ is the dynamic significance parameter that have a different normality range for each medoid depending on the local density.

The space complexity of the proposed modeling algorithm in general feature space is $O(3 * n^2)$. The time complexity of our algorithm is the sum of time complexities of the three steps and is equivalent to $O(\omega * (n^2 + n^2 * \log(n))) + O((\#_{medoids} * \log(\#_{medoids}))) + O(|DB|^2 * \#_{medoids} * \log(\#_{medoids}))$ where

- $O(n^2)$ is the time complexity of affinity matrix computation
- $O(n^2 * \log(n))$ is the time complexity of message passing to compute availability and responsibility matrix
- ω is the number of times the modeling algorithm is repeated to identify m medoids. It has been observed that the value of ω normally lies in the range 3-10.
- $m * \log(m)$ is the time complexity of computing possible normality range
- $|DB| * m$ is the time complexity for selecting customized normality range for each medoid where $|DB|$ is the number of trajectories present in trajectory dataset DB .

4. Classification and anomaly detection

Once the m -Medoids based model for all the classes have been learnt, the classification of new trajectories is performed by checking the closeness of said trajectory to the models of different classes. The classification of unseen samples to known classes and anomaly detection is performed using following steps:

1. Identify k nearest medoids, from the entire set of medoids (\mathbf{M}) belonging to different classes, to unseen sample Q as:

$$k\text{-NM}(Q, \mathbf{M}, k) = \{ \mathbf{C} \in \mathbf{M} \mid \forall R \in \mathbf{C}, S \in \mathbf{M} - \mathbf{C}, \text{Dist}(Q, R) \leq \text{Dist}(Q, S) \wedge |\mathbf{C}| = k \} \quad (11)$$

where \mathbf{M} is the set of all medoids from different classes and \mathbf{C} is the ordered set of k closest medoids starting from the nearest medoid.

2. Initialize nearest medoid index ι to 1.
3. Set r to the id of ι^{th} nearest medoid and c to the index of its corresponding class.
4. Set $l = \widehat{\tau}_{(c,r)}$.
5. Identify the normality threshold d w.r.t. the medoid r using $\mathbf{D}^{(c)}$ as:

$$d = \mathbf{D}_l^{(c)} \quad (12)$$

6. Test sample Q is considered to be a valid member of class c if:

$$\text{Dist}(Q, M_r) \leq d \quad (13)$$

7. If the condition specified in eq. (13) is not satisfied, increment the index ι by 1.
8. Iterate steps 3-7 till ι gets equivalent to k . If the test trajectory Q has not been identified as a valid member of any class, it is considered to be an outlier and deemed anomalous.

The time complexity of MMC-GFS based classification and anomaly detection algorithm is $O(|\mathbf{M}|) + O(k)$ for anomalous samples where $|\mathbf{M}|$ is the total number of medoids belonging to all classes. However, for most of the normal samples the time complexity is $O(|\mathbf{M}|)$. The time complexity can be further reduced by using efficient indexing structure like kd-trees to index $|\mathbf{M}|$ medoids for efficient k -NM search.

5. Relative merits of m -Medoids based modeling and classification algorithms

In this section, we provide a comparative evaluation of the proposed multimodal m -Medoids (MMC-GFS) and localized m -Medoids (Khalid, 2010a) based frameworks (LMC-ES) for modeling, classification and anomaly detection. These frameworks can be characterized in terms of the following attributes:

- Ability to deal with multimodal distribution within a pattern
- Ability to deal with variety of feature space representation of trajectories
- Time complexity of generating m -Medoids based model of known patterns
- Time complexity of classification and anomaly detection using learned models of normality
- Scalability of modeling mechanism to cope with increasing number of training data

For the ease in understanding of the comparative analysis, simulation of the working of proposed modeling and classification algorithms for arbitrary shaped patterns having multimodal distributions is presented in Fig. 1. In the left image of Fig. 1, each point represents the training sample and instances belonging to the same class are represented with same color and marker. Squares superimposed on each group of instances represent the medoids used for modeling the pattern. Normality region generated using different frameworks for classification and anomaly detection is depicted in the right image of Fig. 1. Test sample is considered to be a normal member of the class if it lies within the normality

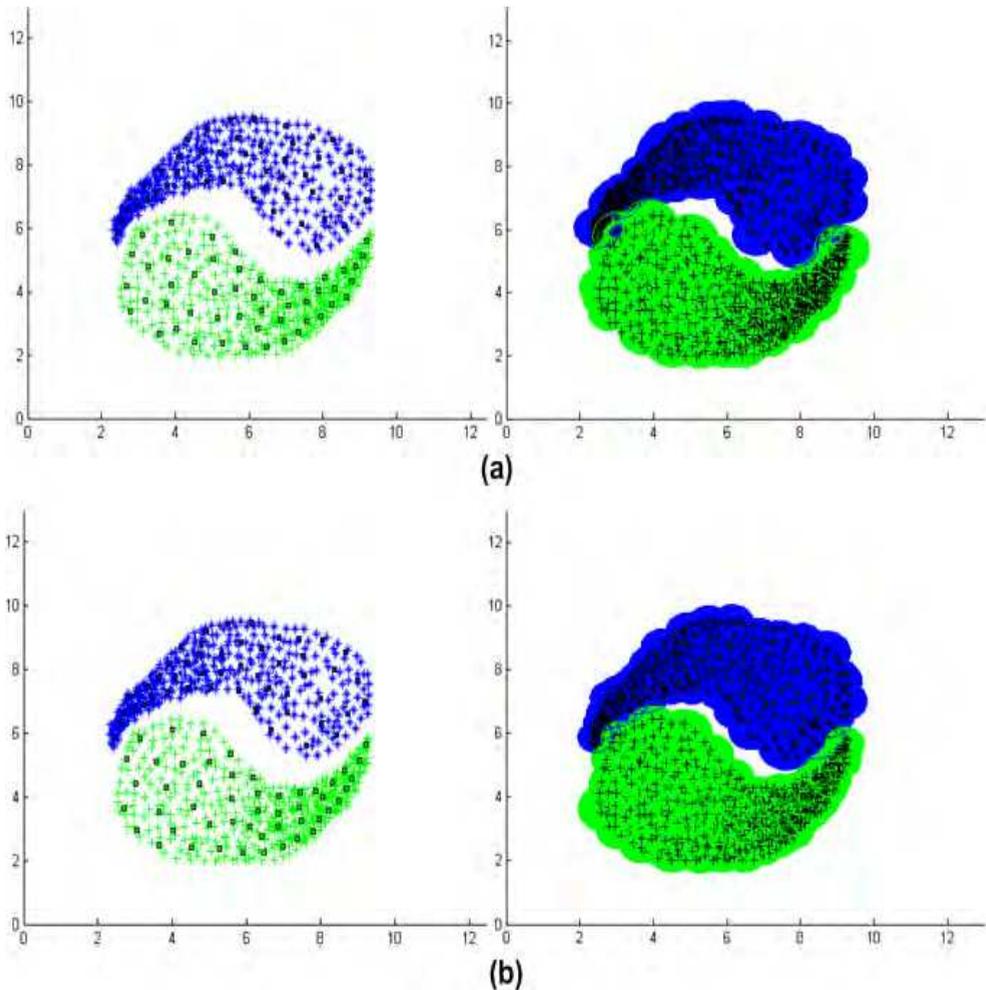


Fig. 1. m -Medoids based modeling of patterns using (a) LMC-ES framework (b) MMC-GFS framework.

region, else it is marked as anomalous. Visualization of LMC-ES and MMC-GFS based modeling is provided in Fig. 1(a) and Fig. 1(b) respectively.

Multimodal modeling using MMC-GFS frameworks caters for the multimodal distribution within a pattern. On the other hand, LMC-ES framework always assumes a unimodal distribution within a pattern and hence can not cater for the dynamic distribution of samples within a pattern. It is apparent from Fig. 1 that MMC-GFS frameworks have generated more accurate models that have accommodated the variation in sample density within a given pattern. LMC-ES framework performs a hard classification of unseen sample. A sample is classified to a pattern represented by the majority of medoids from a set of k nearest medoids. The sample may not lie in the normality region of a pattern to which it is classified and hence

deemed anomalous. However, it is likely that it may still fall in the normality region of the second closest but less dense pattern having larger normality range. The hardness of LMC-ES based classification algorithm will result in the misclassification of such samples. However, MMC-GFS based classification and anomaly detection algorithm does not give a hard decision and checks for the membership of test trajectories w.r.t. different patterns until it is identified as a valid member of some pattern or it has been identified as anomalous w.r.t. k nearest medoids. This relatively softer approach enables the MMC-GFS based classification algorithm to adapt to the multimodal distribution of samples within different patterns. This phenomena is highlighted in Fig. 2. The samples, represented by 'x' marker, will be classified to blue pattern but is marked as anomalous using LMC-ES classifier as it falls outside the normality range of dense medoids belonging to the closest pattern. On the other hand, soft classification technique as proposed in MMC-GFS frameworks will correctly classify the sample as normal members of green pattern. Another benefit of MMC-GFS framework is that it can be applied to any feature space representation of trajectories with a given distance function. On the other hand, LMC-ES can only operate in feature spaces with a computable mean.

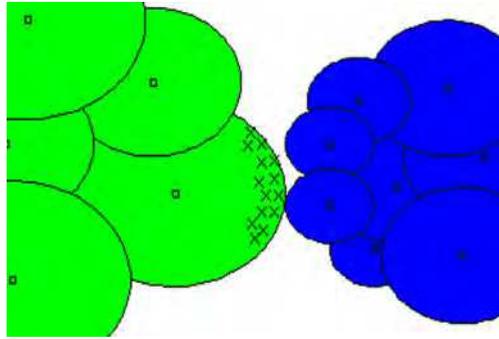


Fig. 2. Scenario for evaluating the adaptation of classification algorithms as proposed in different m -Medoids based frameworks.

Algorithms to generate m -Medoids model, as proposed in LMC-ES framework, is efficient and scalable to large datasets. On the other hand, the modeling algorithm of MMC-GFS is not scalable to very large datasets due to the requirement of affinity matrix computation. The space and time complexity is quadratic which is problematic for patterns with large number of training sample. However, this problem can be easily catered by splitting the training sample into subsets and selecting candidate medoids in each subset using algorithm as specified in section 3.1. The final selection of medoids can be done by applying the same algorithm again but now using the candidate medoids instead of all the training sample belonging to a given pattern. The classification algorithm of MMC-GFS framework is more efficient as compared to LMC-ES framework. This efficiency gain is due to the non-iterative unmerged anomaly detection with respect to a given medoid. The anomaly detection is done by applying a single threshold to the distance of the test sample from its i^{th} closest medoid as specified in eq. (13). On the other hand, LMC-ES implements iterative merged anomaly detection, which is more accurate but time consuming as compared to the modeling algorithm proposed in MMC-GFS framework. The time complexity of merged anomaly detection is $O(m * \log(m) - \tau * \log(\tau))$.

6. Experimental results

In this section, we present some results to analyze the performance of the proposed multimodal m -Medoids based modeling, classification and anomaly detection as compared to competitive techniques.

6.1 Experimental datasets

Experiments are conducted on synthetic SIM₂ and real life LAB (Khalid, 2010a;b; Khalid & Naftel, 2006), HIGHWAY (Khalid & Naftel, 2006) and ASL (Bashir et al., 2006; 2005a;b; 2007; Khalid, 2010b; Khalid & Naftel, 2006) datasets. Details of these datasets can be found in Table 1.

Dataset	Description	# of trajectories	Extraction method	Labelled (Y/N)
SIM ₂	Simulated dataset comprising of two dimensional coordinates.	arbitrary #	Simulation.	Y
LAB	Realistic dataset generated in the laboratory controlled environment for testing purposes. Trajectories can be categorised into 4 classes.	152	Tracking moving object and storing motion coordinates.	Y
HIGHWAY	Realistic vehicle trajectory dataset generated by tracking vehicles in a highway traffic surveillance sequence.	355	Tracking vehicles using PTMS(Melo et al., 2004) tracking algorithm.	Y
ASL	Trajectories of right hand of signers as different words are signed. Dataset consists of signs for 95 different word classes with 70 samples per word.	6650	Extracting (x, y) coordinates of the mass of right hand from files containing complete sign information.	Y

Table 1. Overview of datasets used for experimental evaluation

6.2 Experiment 1: Evaluation of m -Medoids based frameworks for classification and anomaly detection

The purpose of this experiment is to evaluate the performance of proposed MMC-GFS and LMC-ES based frameworks for classification of unseen data samples to one of the known patterns. The effectiveness of the proposed frameworks to perform anomaly detection is also demonstrated here. The experiment has been conducted on simulated SIM₂ dataset. Training

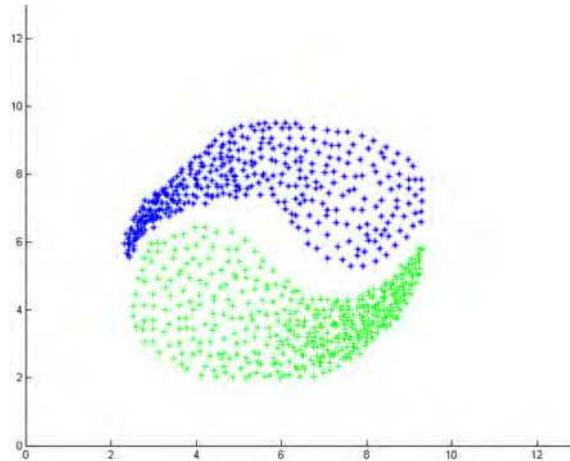


Fig. 3. Training data from SIM₂ dataset.

data from simulated datasets is shown in Fig. 3. Test data for SIM₂ dataset is obtained by generating 500 samples from a uniform distribution such that $(x, y) \in (U(1, 12), U(1, 12))$.

We have used 50 medoids to model a class using its member samples. The classification and anomaly detection results for SIM₂ dataset, using LMC-ES and MMC-GFS frameworks are presented in Fig. 4(a) and Fig. 4(b) respectively. Training samples are represented using '+' marker whereas classified normal samples are represented by small circles. Data points belonging to same class are represented with same colour and marker. Samples from test data which have been identified as anomalous are represented with a black 'x' marker. It is apparent from Fig. 4 that multimodal m -Medoids based classification system as proposed in MMC-GFS framework performs better classification and anomaly detection while catering for multimodal distribution within the modeled pattern. On the other hand, LMC-ES based framework performs unimodal modeling of patterns and therefore the classification system does not adjust well to the variation of density within a pattern.

After demonstrating the efficacy of proposed classification and anomaly detection approach on synthetic data, the experiment is then repeated on real life LAB and HIGHWAY datasets. LAB and HIGHWAY datasets are classified motion datasets and contain anomalous trajectories within the datasets themselves. Classified training data for these datasets is obtained by randomly selecting half of the trajectories from each of the normal patterns in the dataset. The remaining half of the trajectories from the normal patterns along with anomalous trajectories are extracted and used as test data. Training samples from the LAB and HIGHWAY datasets are shown in Fig. 5 and Fig. 6 respectively. For ease of visualization, samples from each class are plotted separately on the background scene. The starting point of each trajectory is marked in green.

Trajectories from LAB and HIGHWAY datasets are modelled using DFT-MOD based coefficient feature vectors. (Khalid, 2010b). Patterns are modeled using 20 medoids per pattern. Once the multimodal m -Medoids based model for all the classes have been learnt, classification of samples from the test data is done using the classifier as proposed in section

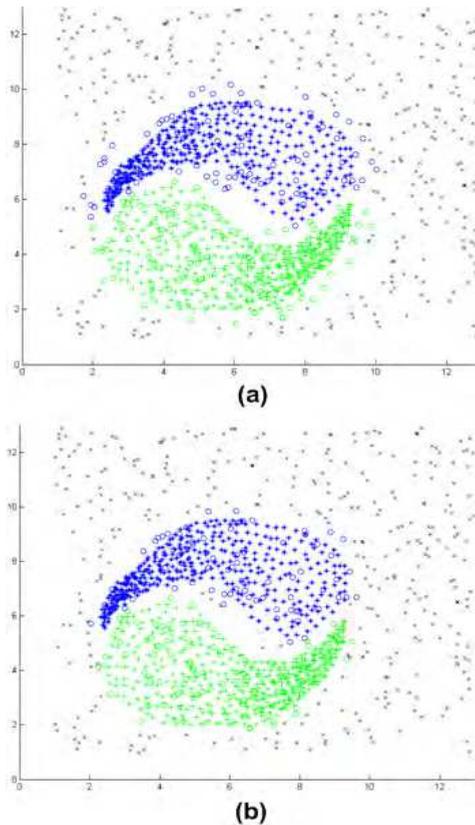


Fig. 4. Classification of test data, based on SIM_2 classes, using (a) LMC-ES framework (b) MMC-GFS framework.

4. Classification obtained by applying the MMC-GFS approach on LAB dataset is shown in Fig. 7. The matching of classification obtained for each trajectory with its ground truth shows that no trajectory is misclassified. Trajectories identified as anomalous are shown in Fig. 8. It is clear from Fig. 8 that anomalous trajectories are significantly different from the normal motion patterns as shown in Fig. 7. The classification experiment is also conducted on the HIGHWAY dataset and the results obtained are shown in Fig. 9. Trajectories filtered as anomalous are shown in Fig. 10. These experimental results give evidence to the claim that MMC-GFS based classification and anomaly detection system is an effective and robust approach that works well with real life motion datasets.

6.3 Experiment 2: Comparison of proposed classifiers with competitive techniques

The purpose of this experiment is to compare the performance of classifiers as proposed in LMC-ES and MMC-GFS frameworks. For comparison of our results with competitive techniques, we establish a base case by implementing three different systems for comparison including Mahalanobis and GMM classifier. Real life ASL dataset is used for the experiment.

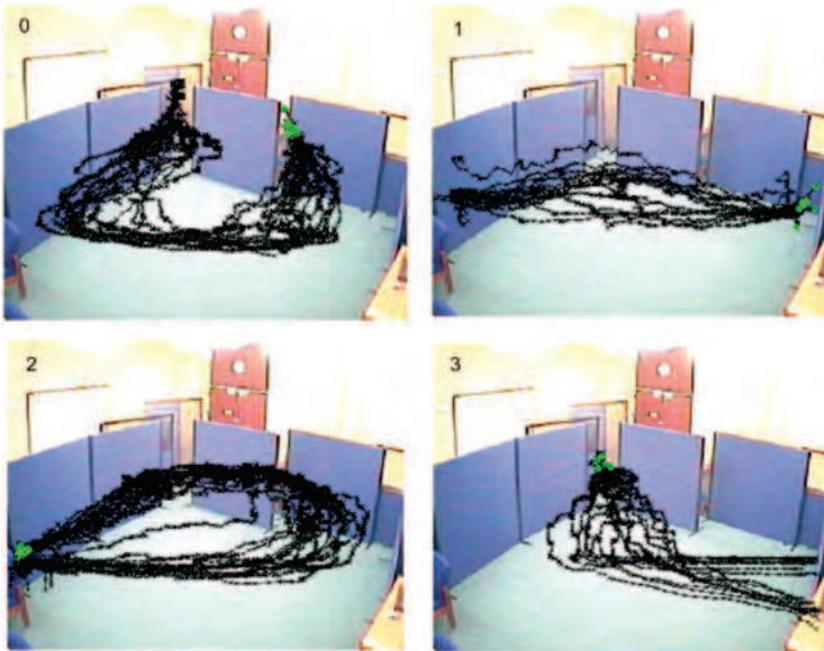


Fig. 5. Labelled training samples from LAB dataset. Trajectories belonging to different classes are plotted separately on background scene

Signs from different number of word classes are selected. Classified training data is obtained by randomly selecting half of the trajectories from each word class leaving the other half to be used as test data. Trajectories from ASL dataset are represented using DFT-MOD based coefficient feature vectors (Khalid, 2010b). Patterns are modeled using 20 medoids per pattern. We have computed single multimodal Gaussian for modeling of patterns for Mahalanobis classifier. Modeling of patterns and classification of unseen samples using GMM is based on the approach as described in (Bashir et al., 2005a). Each class is modeled using a separate GMM. The number of modes to be used for GMM-based modeling is automatically estimated using a string of pruning, merging and mode-splitting processes as specified in (Bashir et al., 2005a). Once the models for all the classes have been learnt, the test data is passed to different classifiers and the class labels obtained are compared with the ground truth. The experiment is repeated with different numbers and combinations of word classes. Each classification experiment is averaged over 50 runs to reduce any bias resulting from favorable word selection.

The accuracy of different classifiers for wide range of word classes from ASL dataset is presented in Table 2. Based on these results, we can see that the multimodal m -Medoids based classifier as proposed in MMC-GFS framework yield a superior classification accuracy as compared to other classifiers closely followed by unimodal LMC-ES framework. GMM yields good results for lower number of classes but its performance deteriorates for higher number of word classes. It can also be observed from Table 2 that the relative accuracy of proposed m -Medoids based MMC-GFS and LMC-ES classifiers increases with an increase in the number

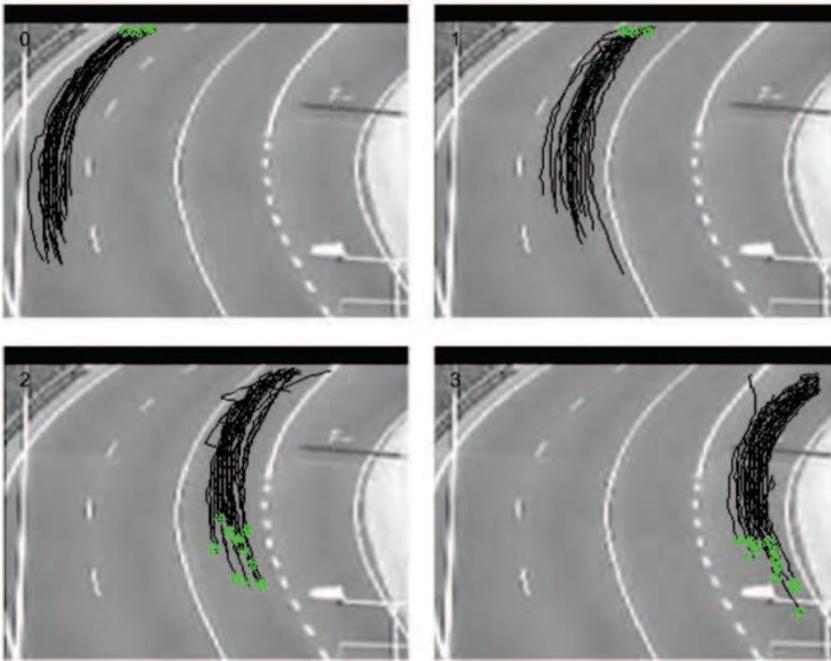


Fig. 6. Labelled training samples from HIGHWAY dataset. Trajectories belonging to different classes are plotted separately on background scene

	ASL (#classes : #samples)				
	2 : 70	4 : 140	8 : 280	16 : 560	24 : 840
MMC-GFS	0.99	0.94	0.91	0.86	0.83
LMC-ES	0.98	0.92	0.88	0.83	0.78
Mahalanobis	0.95	0.88	0.82	0.75	0.71
GMM	0.97	0.92	0.83	0.74	0.69

Table 2. Classification accuracies for different number of classes from ASL dataset

of classes as compared with competitive techniques; thus making them more scalable for larger number of classes. The superior performance of MMC-GFS, as compared to competitive techniques, can be explained by the fact that the proposed multimodal m -Medoids based frameworks do not impose any restriction on the probability distribution function of modeled patterns. The proposed frameworks can effectively model arbitrary shaped patterns and can effectively handle variation in sample distribution within a pattern as demonstrated in Fig. 1 and Fig. 4. On the other hand, the competitive approaches impose assumptions on the PDF of patterns (normally gaussian). These approaches do not have the capacity to handle multimodal distribution within a pattern. As a result, the model generated by these approaches will not give an accurate representation of complex patterns and hence result in poor classification performance as compared to the proposed multimodal m -Medoids based approaches.

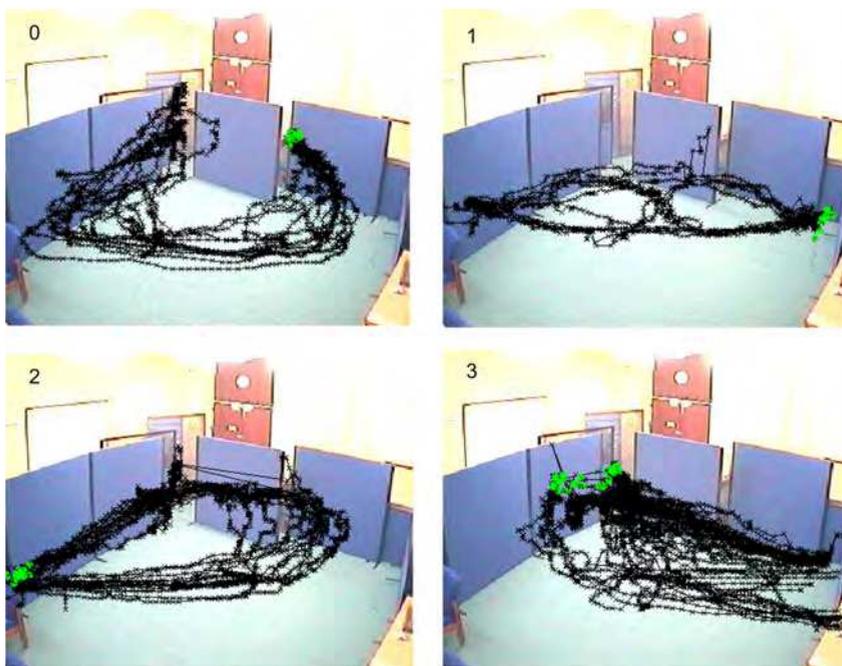


Fig. 7. Classification of test trajectories from LAB dataset

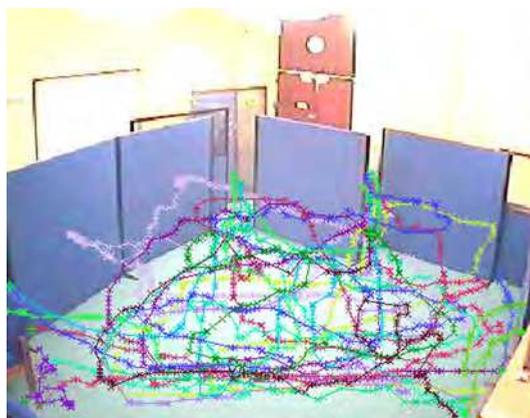


Fig. 8. Trajectories identified as anomalous from LAB dataset using proposed anomaly detection mechanism.

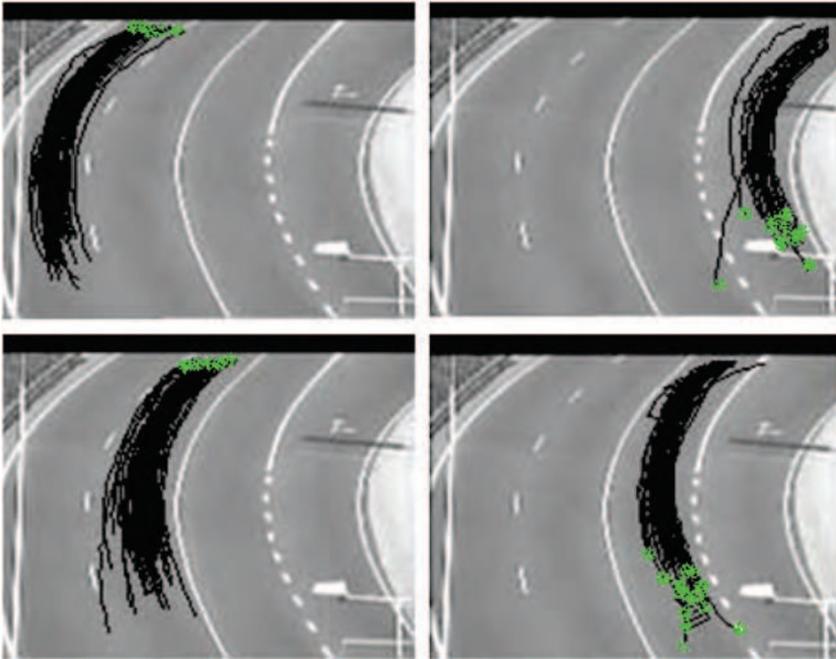


Fig. 9. Classification of test trajectories from HIGHWAY dataset

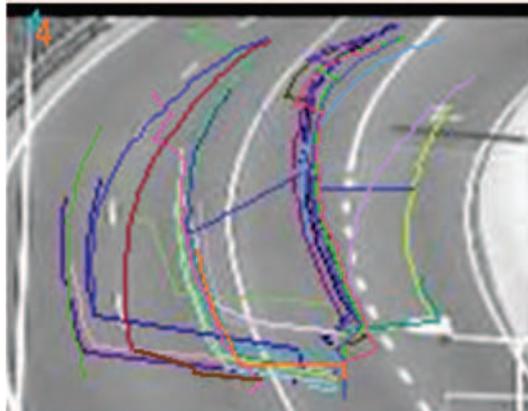


Fig. 10. Trajectories identified as anomalous from HIGHWAY dataset using proposed anomaly detection mechanism.

Similar experiment with ASL dataset (using similar experimental settings) has been conducted by (Bashir et al., 2007) using their proposed GMM and HMM-based classification system. They reported classification accuracies of 0.96, 0.92, 0.86 and 0.78 for 2, 4, 8 and 16 word classes respectively. A comparison of these classification accuracies with the results obtained using our approach reveals that classifiers from *m*-Medoids classifier family performs better than GMM and HMM-based recognition system (Bashir et al., 2007) despite the fact that our proposed classification approach is conceptually simpler and computationally less expensive.

6.4 Experiment 3: Quantitative evaluation of anomaly detection algorithms

Here we provide a quantitative evaluation and comparison of *m*-Medoids based anomaly detection algorithms, as proposed in MMC-GFS and LMC-ES frameworks, with competitors. We implemented three different anomaly detection techniques based on statistical test as proposed in (Khalid & Naftel, 2006), Grown When Required (GWR) novelty filter as proposed in (Marsland et al., 2002) and one-class classifier based anomaly detection as proposed in (Tax, 2001). (Khalid & Naftel, 2006) performs anomaly detection by using Mahalanobis classifier and conducting Hotelling's T^2 test. (Tax, 2001) perform anomaly detection by generating model of one class (referred to as target class) and distinguishing it from samples belonging to all other classes. There generation of model of the target class is done using SVM and GMM. For SVM-based one class classifier (OCC-SVM), we have used RBF kernel for the modeling of target class. For GMM-based one class classifier (OCC-GMM), we have used the approach as specified in Experiment 2 to generate the GMM-based model.

The experiment has been conducted using different number of word classes from ASL dataset. We have extracted half of the samples belonging to each word class for training purposes leaving the other half of the samples to be used as test data. DFT-MOD based coefficient feature vector representation of sign trajectories from training data is generated and used to generate models as required by the different classification approaches. MMC-GFS and LMC-ES framework based model of each class is generated using the algorithm as presented in section 3 and (Khalid, 2010a) respectively. Patterns are modeled using 20 medoids per pattern.

Once the model learning phase is over, anomaly detection using different techniques is carried out using test dataset. We would expect that few instances drawn from class X would be recorded as anomalous when tested against the same class, whereas nearly all instances would be detected as anomalous when tested against a different class Y. The experiment is repeated with different numbers and combinations of word classes. Each anomaly detection experiment is averaged over 50 runs to reduce any bias resulting from favorable word selection.

Fig. 11 reports the result in terms of percentage of correct anomaly detection using various number of word classes from ASL dataset. The results demonstrate the superiority of anomaly detection using *m*-Medoids based MMC-GFS and LMC-ES frameworks. The anomaly detection accuracies obtained using MMC-GFS algorithm are higher than unimodal LMC-ES based anomaly detection algorithm. MMC-GFS and LMC-ES performs better than OCC-SVM, OCC-GMM, GWR and Mahalanobis framework-based Naftel's method. The superior performance of proposed approach as compared to state-of-the-art techniques is due to the fact that our approach gives importance to correct classification of normal sample

and to the filtration of abnormal samples during the model generation phase. On the other hand, OCC-SVM generates good model of normal classes but classifies many of the abnormal samples as member of normal classes whereas GWR gives extra importance to filtering abnormal samples and in the process, identifies many normal samples as abnormal.

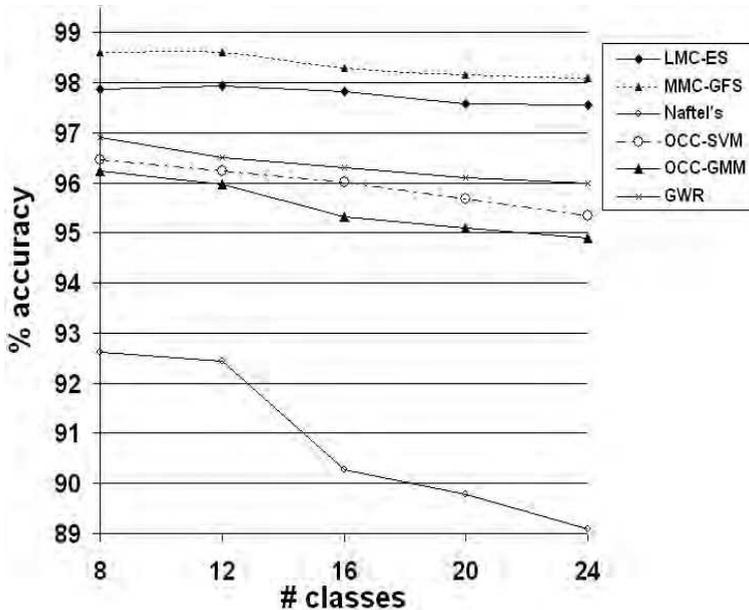


Fig. 11. Percentage anomaly detection accuracies for different number of classes from ASL dataset

7. Discussion and conclusions

In this chapter, we have presented an extended m -Medoids based framework, referred to as MMC-GFS, for modeling of trajectory-based motion patterns. The strength of the proposed approach is its ability to model complex patterns without imposing any restriction on the distribution of samples within a given pattern. Once the multimodal m -Medoids model for all the classes have been learnt, the classification of new trajectories and anomaly detection is then performed using a proposed soft classification and anomaly detection algorithm which is adaptive to multimodal distributions of samples within a pattern. The strength of this technique is its ability to model complex patterns without imposing any restriction on the shape of patterns. MMC-GFS can be used for modeling, classification and anomaly detection in any feature space with a computable similarity function.

Experimental results are presented to show the effectiveness of proposed MMC-GFS classifier. Modeling of pattern and classification using proposed frameworks is unaffected by variation of sample distribution within a pattern as demonstrated in Fig. 4. Quantitative comparison of MMC-GFS based classifiers with competitive techniques demonstrates the superiority of our multimodal approach as it performs consistently better than commonly used Mahalanobis, GMM and HMM-based classifiers.

Experiments are also conducted to show the effectiveness of anomaly detection capabilities of proposed frameworks. Anomaly detection results for different classes of ASL datasets, using different variants of proposed anomaly detection algorithm, are presented. It has been shown that anomaly detection using multimodal MMC-GFS frameworks gives better anomaly detection accuracies as compared to the unimodal LMC-ES approach. Although LMC-ES enables the anomaly detection system to adapt to the normality distribution of individual classes, it is insensitive to the variation of distributions within a pattern which results in degradation of its performance as compared to MMC-GFS frameworks. Comparison of proposed anomaly detection algorithms with an existing approach demonstrates the superiority of our approach as they consistently perform better for different number of classes.

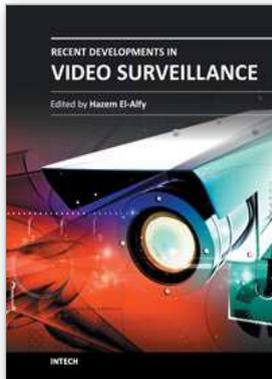
The application of proposed MMC-GFS based modeling and recognition system is not only limited to trajectory-based behavior recognition, but can also be applied to other recognition tasks that are critical in video surveillance application. Some of the applications where MMC-GFS based modeling and classification system can be applied include but is not limited to object recognition in surveillance videos, gait recognition, scene recognition etc.

8. References

- Aghbari, Z., Kaneko, K. & Makinouchi, A. (2003). Content-trajectory approach for searching video databases, *IEEE Transaction on Multimedia*, Vol. 5 of 4, pp. 516–531.
- Bashir, F. I., Khokhar, A. A. & Schonfeld, D. (2006). View-invariant motion trajectory based activity classification and recognition, *ACM Multimedia Systems, special issue on Machine Learning Approaches to Multimedia Information retrieval*, pp. 45–54.
- Bashir, F. I., Khokhar, A. A. & Schonfield, D. (2003). Segmented trajectory based indexing and retrieval of video data, *IEEE International Conference on Image Processing*, Vol. 3, Barcelona, Spain, pp. II– 623–6.
- Bashir, F. I., Khokhar, A. A. & Schonfield, D. (2004). A hybrid system for affine-invariant trajectory retrieval, *Proc. MIR'04*, pp. 235–242.
- Bashir, F. I., Khokhar, A. A. & Schonfield, D. (2005a). Automatic object trajectory based motion recognition using gaussian mixture models, *IEEE International Conference on Multimedia and Expo*, Netherland, pp. 1532– 1535.
- Bashir, F. I., Khokhar, A. A. & Schonfield, D. (2005b). Hmm based motion recognition system using segmented pca, *IEEE International Conference on Image Processing*, Genova, Italy, pp. 1288–1291.
- Bashir, F. I., Khokhar, A. A. & Schonfield, D. (2007). Object trajectory-based activity classification and recognition using hidden markov models, *IEEE Transactions on Image Processing*, Vol. 16 of 7, pp. 1912–1919.
- Brotherton, T., Johnson, T. & Chadderdon, G. (1998). Classification and novelty detection using linear models and a class dependent elliptical basis function neural network, *Proc. IJCNN Conference*, Vol. 2, Anchorage, pp. 876–879.
- Cai, Y. & Ng, R. (2004). Indexing spatio-temporal trajectories with chebyshev polynomials, *ACM SIGMOD/PODS Conference*, France, pp. 599–610.
- Chan, K. & Fu, A. (1999). Efficient time series matching by wavelets, *Proc. of International Conference on Data Engineering*, Sydney, pp. 126–133.

- Chang, S. F., Chen, W., Horace, J. M., Sundaram, H. & Zhong, D. (1998). A fully automated content based video search engine supporting spatiotemporal queries, *IEEE Transactions on Circuits and System for Video Technology*, Vol. 8 of 5, pp. 602–615.
- Dagtas, S., Ali-Khatib, W., Ghafor, A. & Kashyap, R. (2000). Models for motion-based video indexing and retrieval, *IEEE Transactions on Image Processing*, Vol. 9 of 1, pp. 88–101.
- Faloutsos, C., Ranganathan, M. & Manolopoulos, Y. (1994). Fast sub-sequence matching in time-series databases, *Proceedings of the 1994 ACM SIGMOD International Conference on Management of Data*, pp. 419–429.
- Frey, B. & Dueck, D. (2007). *Science*, Vol. 315, pp. 972–976.
- Hamid, R., Johnson, A., Batta, S., Bobick, A., Isbell, C. & Coleman, G. (2005). Detection and explanation of anomalous activities: Representing activities as bags of events n-grams., *CVPR*, pp. 1031–1038.
- Hsu, C.-T. & Teng, S.-J. (2002). Motion trajectory based video indexing and retrieval, *IEEE International Conference on Image Processing*, Vol. 1, pp. 605–608.
- Hu, W., Tan, T., Wang, L. & Maybank, S. (2004). A survey on visual surveillance of object motion and behaviors, *IEEE Transactions on Systems, Man & Cybernetic*, Vol. 34 of 3, pp. 334–352.
- Hu, W., Xiao, X., Fu, Z., Xie, D., Tan, T. & Maybank, S. (2006). A system for learning statistical motion patterns, *IEEE Transactions on Pattern Analysis and Machine Learning*, Vol. 28 of 9, pp. 1450–1464.
- Hu, W., Xiao, X., Xie, D., Tan, T. & Maybank, S. (2004). Traffic accident prediction using 3-d model based vehicle tracking, *IEEE Transactions on Vehicular Tech*, Vol. 53 of 3, pp. 677–694.
- Hu, W., Xie, D., Fu, Z., Zeng, W. & Maybank, S. (2007). Semantic based surveillance video retrieval, *IEEE Transactions on ImageProcessing*, pp. 1168–1181.
- Hu, W., Xie, D., Tan, T. & Maybank, S. (2004). Learning activity patterns using fuzzy self-organizing neural networks, *IEEE Transactions on Systems, Man & Cybernetic*, Vol. 34 of 3, pp. 1618–1626.
- Jeanin, S. & Divakaran, A. (2001). Mpeg-7 visual motion descriptors, *IEEE Trans. Circuits Syst. Video Technol.*, Vol. 11 of 6, pp. 720–724.
- Jin, Y. & Mokhtarian, F. (2004). Efficient video retrieval by motion trajectory, *Proceedings of British Machine Vision Conference*, Kingston, pp. 667–676.
- Johnson, N. & Hogg, D. (1995). Learning the distribution of object trajectories for event recognition, *Proceedings of British Conference on Machine Vision*, pp. 582–592.
- Keogh, E., Chakrabarti, K., Pazzani, M. & Mehrota, S. (2001). Locally adaptive dimensionality reduction for indexing large time series databases, *Proc. ACM SIGMOD Conference*, pp. 151–162.
- Khalid, S. (2010a). Activity classification and anomaly detection using m-mediods based modeling of motion patterns, *Pattern Recognition*, Vol. 43, pp. 3636–3647.
- Khalid, S. (2010b). Motion based behaviour learning, profiling and classification in the presence on anomalies, *Pattern Recognition*, Vol. 43, pp. 173–186.
- Khalid, S. & Naftel, A. (2005). Evaluation of matching metrics for trajectory based indexing and retrieval of video clips, *Proceedings of IEEE WACV*, Colorado, USA, pp. 242–249.
- Khalid, S. & Naftel, A. (2006). Classifying spatiotemporal object trajectories using unsupervised learning in the coefficient feature space, *Multimedia Systems*, Vol. 12, pp. 227–238.

- Marsland, S., Shapiro, J. & Nehmzow, U. (2002). A self-organising network that grows when required, *Neural Networks*, Vol. 15, pp. 1041–1058.
- Melo, J., Naftel, A., Bernardino, A. & Santos-Victor, J. (2004). Viewpoint independent detection of vehicle trajectories and lane geometry from uncalibrated traffic surveillance cameras, *Proc. of International Conference of Image Analysis and Recognition*, Porto, Portugal, pp. 454–462.
- Owens, J. & Hunter, A. (2000). Application of the self-organising map for trajectory classification, *Proceedings of Third IEEE International Workshop on Visual Surveillance*, Dublin, Ireland, p. 77.
- Rea, N., Dahyot, R. & Kokaram, A. (2004). Semantic event detection in sports through motion understanding, *Proceedings of Conference on Image and Video Retrieval*, Dublin, Ireland, pp. 88–97.
- Roberts, S. & Tarassenko, L. (1994). A probabilistic resource allocating network for novelty detection, *Neural Computation*, Vol. 6, pp. 270–284.
- Shim, C. & Chang, J. (2001). Content based retrieval using trajectories of moving objects in video databases, *Proceedings of IEEE 7th International Conference on Database Systems for Advanced Applications*, pp. 169–170.
- Shim, C. & Chang, J. (2004). Trajectory based video retrieval for multimedia information systems, *Proceedings of ADVIS*, pp. 372–382.
- Tax, D. (2001). One-class classification, *PhD thesis, Delft University of Technology*.
- Vlachos, M., Kollios, G. & Gunopulos, D. (2002). Discovering similar multidimensional trajectories, *Proceedings of the International Conference on Data Engineering*, San Jose, CA, pp. 673–684.
- Xiang, T. & Gong, S. (2005). Video behaviour profiling and abnormality detection without manual labelling, *Proc. IEEE International Conference on Computer Vision*, Vol. 2, London, UK, pp. 1238–1245.
- Xiang, T. & Gong, S. (2006). Incremental visual behaviour modelling, *The 6th IEEE Workshop on Visual Surveillance*, Graz, pp. 65–72.
- Yacoob, Y. & Black, M. (1999). Parameterized modeling and recognition of activities, *Computer Vision and Image Understanding*, Vol. 73, pp. 232–247.
- Yeung, D. & Chow, C. (2002). Parzen window network intrusion detectors, *Proc. International Conference on Pattern Recognition, IEEE*, Canada, pp. 385–388.
- Zhang, D., Gatica-Perez, Bengio, S. & McCowan, I. (2005). Semi-supervised adapted hmms for unusual event detection, *Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition*, pp. 611–618.
- Zhong, H., Shi, J. & Visontai, M. (2004). Detecting unusual activity in videos, *Proc. Of CVPR*, Washington D.C., pp. 819–826.



Recent Developments in Video Surveillance

Edited by Dr. Hazem El-Alfy

ISBN 978-953-51-0468-1

Hard cover, 122 pages

Publisher InTech

Published online 04, April, 2012

Published in print edition April, 2012

With surveillance cameras installed everywhere and continuously streaming thousands of hours of video, how can that huge amount of data be analyzed or even be useful? Is it possible to search those countless hours of videos for subjects or events of interest? Shouldn't the presence of a car stopped at a railroad crossing trigger an alarm system to prevent a potential accident? In the chapters selected for this book, experts in video surveillance provide answers to these questions and other interesting problems, skillfully blending research experience with practical real life applications. Academic researchers will find a reliable compilation of relevant literature in addition to pointers to current advances in the field. Industry practitioners will find useful hints about state-of-the-art applications. The book also provides directions for open problems where further advances can be pursued.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Shehzad Khalid (2012). Behavior Recognition Using any Feature Space Representation of Motion Trajectories, Recent Developments in Video Surveillance, Dr. Hazem El-Alfy (Ed.), ISBN: 978-953-51-0468-1, InTech, Available from: <http://www.intechopen.com/books/recent-developments-in-video-surveillance/multivariate-m-mediods-based-modeling-and-classification-in-any-feature-space-representation-of-traj>

INTECH

open science | open minds

InTech Europe

University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China

Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821

© 2012 The Author(s). Licensee IntechOpen. This is an open access article distributed under the terms of the [Creative Commons Attribution 3.0 License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.