# A Context Space Model for Detecting Anomalous Behaviour in Video Surveillance

Arnold Wiliem*, Vamsi Madasu*, Wageeh Boles* and Prasad Yarlagadda*
*Queensland University of Technology, Brisbane, Australia.
Email: a.wiliem@student.qut.edu.au

***Abstract*—Having a good automatic anomalous human behaviour detection is one of the goals of smart surveillance systems' domain of research. The automatic detection addresses several human factor issues underlying the existing surveillance systems. To create such a detection system, contextual information needs to be considered. This is because context is required in order to correctly understand human behaviour. Unfortunately, the use of contextual information is still limited in the automatic anomalous human behaviour detection approaches. This paper proposes a context space model which has two benefits: (a) It provides guidelines for the system designers to select information which can be used to describe context; (b) It enables a system to distinguish between different contexts. A comparative analysis is conducted between a context-based system which employs the proposed context space model and a system which is implemented based on one of the existing approaches. The comparison is applied on a scenario constructed using video clips from CAVIAR dataset. The results show that the context-based system outperforms the other system. This is because the context space model allows the system to considering knowledge learned from the relevant context only.**

## I. Introduction

In the recent years, there has been a growing interest in developing automatic anomalous behaviour detection methods for video surveillance systems. These methods minimise human factors which can negatively affect the systems performance to detect security breaches. These factors are: fatigue and human limitations in effectively monitoring several video displays at the same time [1].

Anomaly detection can be defined as the problem of finding patterns in data that do not conform to expected behaviour [2]. From this definition, an anomalous pattern can be detected when it does not fall within the boundaries of normal patterns. Although it does not seem difficult to develop methods based on this straight forward strategy, there are a number of factors making anomaly detection challenging: [2] (a) Defining the boundary separating between normal patterns and anomalous patterns is difficult; (b) If the anomalies are the result of malicious actions then the malicious adversaries can adapt themselves so that the actions appear normal; (c) Normal behaviour may keep evolving so that the current notion of a normal pattern might not able to represent the normal patterns; (d) The exact notion of an anomaly may be different from one domain to another; (e) Availability of the training data; and (f) The presence of noise which behaves similarly to anomalies. So, the characteristics of the anomaly detection problem needs to be defined first before determining the appropriate method to use.

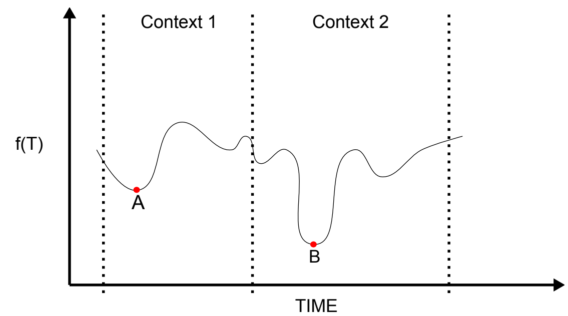


Fig. 1: A simple example of contextual anomalies on $f(T)$. Here anomalies are defined as the lowest values of $f(T)$. In general, point A is not the lowest value, however it becomes the lowest value when other values outside context 1 are excluded. So, in this case point A can be regarded as a contextual anomaly.

Some studies in the field of non-verbal behaviour suggest that it is nearly impossible to understand human behaviour without knowing the context in which the behaviour is observed [3]. In other words, a behaviour could only be regarded as anomalous in a particular context, but could be normal in the other contexts. For example, a person running on a train station platform when there is a train departing is normal, however a person running on a train station platform when the train schedule has finished for that day could be considered as anomalous. So, the notion of normal behaviour needs to be updated over time in order to sufficiently represent the normal behaviour model in the current context. This also suggests that it is almost impossible to have labeled training sets which can be used to generate such a normal behaviour model which sufficiently represents normal behaviour for all contexts. Not only that the notion of normal behaviour may change over different contexts, but it is also impossible to have a dataset enumerating all possible human behaviours [4].

With the characteristics described above, anomalous human behaviours can be categorised as complex anomalies, or specifically contextual anomalies. Contextual anomalies are a type of anomalies which only appears anomalous when the context in which they appear is considered [2]. Figure 1 shows a simple example of contextual anomaly.

There are two important aspects that need to be considered in order to devise a method which is able to detect contextual anomalies [2]: Contextual attributes and behavioural attributes. Behavioural attributes consist of any information describing behaviours and contextual attributes consist of any information describing contexts. These attributes are used when the system is attempting to detect anomalies.

Generally, there are two strategies for detecting contextual anomalies utilising both contextual and behaviour attributes[2]: (a) By reducing the detection problem into a simple anomaly detection by excluding information unrelated to the given context; (b) By building a normal behaviour model for each context. These strategies suggest that the system needs to have a capability to distinguish between different contexts. This can be done by initially modeling the contextual attributes.

While there are many features proposed to describe human behaviour in the video surveillance domain, it is still not clear how contextual information is used to describe a context. Despite its popular use in various domains [5], [6], the use of context for detecting anomalous behaviour in video surveillance is relatively new. Most approaches concentrate on how the normal behaviour model is updated [7], [8]. The update process only takes into account the normal behaviour which has not been seen previously. This implies that the approaches are unable to detect contextual anomalies which appear normal in the previous contexts. Actually, contextual information is implicitly used in rule-based approaches [9]. However, using a set of rules to describe contexts is only possible when the contexts are known a priori.

Recently, Tao et al [10] proposed a method which implicitly uses time as contextual information. They employ adaptive normal and abnormal behaviour models to detect anomalies. A type of behaviour can be classified either into normal or abnormal depending on its weight updated over a period of time. By using this approach, a normal behaviour which has not occurred for a long time will be reclassified into abnormal behaviour and vice versa. In spite of promising performance shown in the experiments, their work only suggests the possibility of using contextual information to detect contextual anomalies in video surveillance.

This work proposes a context space model which can be used for detecting anomalous behaviour in video surveillance. Specifically, both contextual and behavioural attributes will be modeled by the proposed context space model. The model is a generalisation of the context models proposed in computer vision and human computer interaction domains [11], [12]. The difference is that these models specify some elements to describe a context. For example, Jo et al [11] suggested that the information spaces should be constructed by a set of entities, roles and situations. Different contexts may have different entities, roles or/and situations. In our approach a context is not necessarily described by these elements. Having a generic model is very important, because different surveillance domains may have difference sets of parameters describing a context.

Besides the proposed model is more generic than the existing ones, implementing the proposed context space model also gives the following benefits:

- It provides a guideline for system designers to determine relevant information describing a context. Although the existing models define some elements to describe contexts, these elements may only be valid in the domain in which the models are proposed. It is not clear how to apply the models for describing contexts in other domains.
- It enables the system to distinguish between two different contexts. By using a representation of the proposed model, the system is able to distinguish between different contexts.

In the next section, the proposed context space model is discussed. Then, two existing context models proposed in other computer vision and human computer interaction domains are fitted into the proposed model. This is to show that the proposed model is a generalisation of the previous models. Finally, a comparative analysis between a system employing an existing approach and a system implementing the proposed model is discussed.

TABLE I: An example of context changes the meaning of a sentence. The subject that the word ‘he’ refers to in sentence S3, depends on what sentence precedes it. If it is S1 then ‘he’ refers to ‘A burglar’. On the other hand, if S2 precedes S3, it refers to ‘Paul’. In this case S1 and S2 set the context for the reader to understand the meaning of sentence S3.

| Statement ID | Statement |
|---|---|
| S1 | A burglar took a big television set from a house in the night. |
| S2 | Paul woke up in the morning and found his wife is still sleeping. |
| S3 | He walked out from the house quietly. |

## II. Context Space Model

According to the Oxford dictionary, the word ‘context’ is defined as the parts that immediately precede or follow any particular passage or ‘text’ and determine its meaning. For instance, a word ‘it’ in a passage could refer to different subject if we change the preceding parts. Hence, those parts become the context for that word (i.e. the word ‘it’). Another example can be found in Table I. In our case, we broaden this context definition into “any information that would help one to make inferences on the meaning of an object”. That information is regarded as independent variable and the subjects whose meaning are influenced by these independent variables are regarded as dependent variables. It is always assumed that both independent and dependent variables are discret variables.

As indicated in [5] contextual information is one of the important ingredients in the construction of a context. Most context-based approaches describe contextual information as any information that has influence in the understanding of a particular dependent variable [5], [6]. For example, the average number of people waiting on a train station platform late at night is less than that in the rush hour. Here, the time influences how one makes an inference on the average number of people. One may draw a conclusion that there is an abnormality by comparing between the current average number of people and the average number of people previously extracted from the same period of time. In this case, the current time becomes the contextual information.

When there are two kinds of information influencing each other, each of them can be considered as contextual information for the other. For instance, the information of a keyboard location and the information of a monitor location in a static image could be regarded as contextual information, as keyboards are usually found below a monitor and vice versa [13].

Based on our observations, one of the important properties of contextual information is that the information can be either classified as dependent or independent variables. In other words, the information must have a relationship. Let us suppose that a piece of information is classified as one of independent variables. This variable then must have a relationship in which it influences the meaning of the dependent variables.

The relationship could either be a one-way relationship, or a mutual relationship. Unlike in one-way relationship in which the dependent variables do not have influence on the independent variables, the dependent variables in the mutual relationship can have influence on the independent variables. In other words, in mutual relationship, the independent variables can be dependent variables and vice versa.

This relationship can be found automatically by using statistical tools such as regression and correlation analyses or manually by the experts. Automatic discovery methods are only available for the kind of information that can be materialised (e.g. as feature vectors for the systems processing information in the form of feature vectors). For example, a system may find contextual information by calculating the correlation over different pairs of variables. Any pair of variables which is correlated is regarded as contextual information. However, in some cases, it is easier for an expert to find it. For example, the distance between CCTV cameras and humans can influence the type of human behaviour features used by the system. An expert may know the kind of features the system should use in a situation where the people appearing in an image are so small that they can be represented as moving points. The kind of features used in this situation is different from the case where human limbs are observable in an image. Another example is that a human observer may be able to identify the relationship between a train schedule and the number of people standing on a platform. This piece of information could be fed into the system easily during the system design phase. So, the existence of a relationship between two kinds of information (i.e. either one way or mutual relationship) is what makes them as contextual information. Definition 3 gives the definition of contextual information used in this paper.

*Definition 1:* (Independent variables) Independent variables constitute information whose meaning is not influenced by other information.

*Definition 2:* (Dependent variables) Dependent variables constitute information whose meaning is influenced by independent variables.

*Definition 3:* (Contextual information) A set of information in which its members can be partitioned into two groups. These groups are independent variables and dependent variables groups.

Notice that the above definition considers dependent variables as contextual information. This is because of the possibility that an independent variable can become a dependent variable and vice versa (i.e. in the case of mutual relationships). In addition, although in the case of one way relationships some of the dependent variables may have little influence on independent variables, a group of dependent variables may have stronger influence on the independent variables. In this case, the relationship changes into a two way relationship.

If one selects a couple of different contextual information as independent variables and forms an information space in which these variables are regarded as the base, then such an information space is defined as a context space.

*Definition 4:* (Context space) Context space is defined as an n-dimensional information space formed by context parameters selected from the contextual information as its bases. The information defined over this information space is referred to as context-sensitive information.
A context space $\Theta$ is formally defined as follows. Given a set of contextual information $CI = \{ci_1 ... ci_n\}$, there exists subsets $CI^1$ and $CI^2$ where $CI = CI^1 \cup CI^2$ and $CI^1$ and $CI^2$ have a relationship. Either $CI^1$ or $CI^2$ is then chosen as the context space base $CSB$. The context space $\Theta$ defines the sets of context-sensitive information $CSI = \{ci_1, ci_2, ci_3, ..., ci_m\}$, where $CSI \in CI^y,\ y = \{1, 2\}$ and y depends on the selection of the base (i.e. if $CI^1$ is chosen as the base, then y is 2 and vice versa). Let us define a mapping function $\theta$ as $\theta : CSB \rightarrow CSI$. In this case, $CSB$ becomes the parameters (or context parameters) of $\theta$ and context is defined as the arguments of the function. Figure 2 presents an illustration of a 3 dimensional context space. A context is then defined as an argument of the mapping function $\theta$.

Here, Context Space Base and Context Sensitive Information can also be regarded as contextual attributes and behavioural attributes. As aforementioned, these attributes are the important ingredients to detect contextual anomalies.

*Definition 5:* (Context) Given a context space $\Theta$ and a function $\theta : CSB \rightarrow CSI$ (CSB: Context Space Base; CSI: Context Sensitive Information), a context C is defined as an argument of the function $\theta$.

To further clarify the context definition, let us consider the train station example which has been discussed previously. We know that the average number of people waiting on a platform has a one way relationship with the current time. In addition, the event of train arriving at the platform may also have a relationship with the average number of people waiting at the platform. For example, there would be more people waiting on the platform when the train is about to arrive than when the train has departed. In this case, the current time and train arrival event are selected as context parameters and a

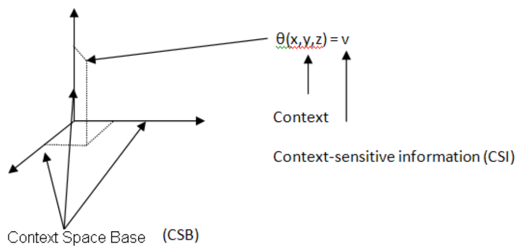


Fig. 2: A graphical illustration of a 3 dimensional context space.

TABLE II: Context space in the train station platform with current time and train arrival event as the context parameters. The number of people waiting on the platform becomes the context-sensitive information in the defined space. CT: current time; AE: arrival event

| | Average number of people on platform | |
|---|---|---|
| CT / AE | Train is arriving | Train is not arriving |
| 9.00 am | 40 | 25 |
| 11.00 am | 20 | 10 |
| 10.00 pm | 5 | 1 |
| 11.00 pm | 2 | 0 |

two dimensional context space is formed. The average number of people waiting on the platform becomes the context-sensitive information defined over this space. In other words, the meaning of the average number of people waiting on the platform depends on the given context. Table II shows an example of the context space.

As described in Definition 4, the interpretation of the context-sensitive information in the context can only be inferred when the values of context parameters are given. In our example, some values of context parameters are given in Table II. This makes it possible for the system to make an inference on whether or not there is an anomaly. For instance, if the system observes that there are 40 people waiting on the platform at 11.00 pm with no train schedule to arrive, it could flag or alert the human observers that there is an anomaly happening there. In other words, usually in the context of (11.00 pm, train not arriving), the average number of people observed is 2, so if the observed number of people is 40 then it is considered as anomalous.

### A. Context space model representations

Once the context space model is constructed, the next step is to determine its representation which will be used by the system. Depending on the model complexity, the representations can be as simple as a rule-based representation and increase in complexity to sets representation. For example, if the system is only required to handle a few identified contexts then a rule-based representation can be used. A matrix representation similar to the one in [12] can be used when the contexts can be enumerated and their number is large. Finally, when not all contexts are identified a priori, then a set representation can be used. By using the set representation the system monitors the values of the context parameters. When at least one of the parameters changes then a context change is detected. In this situation, the system creates a new instance of the context and put it into the set. So, set representation lets the system discover previously unidentified contexts by detecting context change.

### B. Source of contextual information

Since contextual information is important in the proposed model, the next logical question is to identify sources from which the information can be extracted. According to Pantic et al [14], contextual information is usually extracted from various sources. The source of contextual information could vary from one domain to another. One basic guideline is that one may concentrate on the existence of relationships between the context-sensitive information and the contextual information. When a piece of information has a relationship with context-sensitive information then the source of this information could be worth considering.

## III. Comparisons with existing context models

As stated previously, the proposed model is a generalisation of the existing models. In order to show this, two existing models in [11], [12] from the computer vision and human computer interaction domains are remodeled using the proposed context space model.

Jo et al [11] defined context as information spaces which are constructed by a set of entities, roles and situations. Different contexts lead to different types of services that a system would offer to the users. In this case, a set of entities, roles, situations and type of services become the contextual information. Equation 1 present a formal definition of Jo et al's model which is remodeled in terms of the proposed one.

$$\begin{aligned} CI &= \{entities,\ roles,\ situations,\ service\ types\} \\ CI^1 &= \{entities,\ roles\ and\ situations\} \\ CI^2 &= \{type\ of\ services\} \\ CSB &= CI^1\ and\ CSI = CI^2 \end{aligned} \tag{1}$$

Strat [12] defines context as a context set consisting of context elements. Context element is defined as a predicate involving any number of terms that refer to the physical, photogrammetric, or computational context of image analysis. According to Strat, different contexts may have different features and Computer Vision (CV) processes used by the system. In this case the context elements, feature sets and computer vision processes are defined as contextual information. So, the formal definition of the model described in terms of the

Fig. 3: Some images taken from the CAVIAR dataset

proposed model is presented as follows.

$$
\begin{aligned}
CI &= \{context\ elements,\ features\ sets,\ CV\ processes\} \\
CI^1 &= \{context\ elements\} \\
CI^2 &= \{features\ sets,\ computer\ vision\ processes\} \\
CSB &= CI^1\ and\ CSI = CI^2
\end{aligned}
\tag{2}
$$

From Equations 2 and 1 we can see that these models can be redefined in terms of the proposed model. In the proposed model, context may not necessarily be constructed by the parameters defined by the existing models. This point is very important since different surveillance domains may have different sets of contextual information.

## IV. Experiment and Results

This section presents a comparative experiment between a system implementing a context space model which we call "context-based system" and a system which does not use, or implicitly uses context space model which we call "existing system". Since the purpose of the experiment is to outline the advantages of using context space model in detecting anomalous human behaviour, this paper only presents the overall description of each system.

### *A. Scenario description*

We created a simple scenario from the video clips provided in CAVIAR dataset [1] for outlining the advantage of the context-based system over the existing system. The dataset consists of some situations that could be found in a typical office building lobby. Figure 3 presents an example of image feeds taken from the dataset.

In the scenario, it is assumed that people do not to walk into the hallway at point A (Figure 4) when after office hours. It is also assumed that this pattern is not identified during the system design. Or in other words, both systems have to discover this by themselves.

[1]http://homepages.inf.ed.ac.uk/rbf/CAVIAR/


Fig. 4: An illustration of the scenario in CAVIAR dataset.

In order to create such a scenario, the video clips are organised into two groups of ordered lists. Each group represents a different context. The list of videos are presented in Table III. Contexts 1 and 2 respectively represent situations during and after office hours. It was assumed that every video clip represents approximately 30 minutes of scenario time. This means that, the time stamp adds 30 minutes when a new video clip starts. The time stamp for the first video clips of context 1 and 2 is 8 am and 5.30 pm respectively. The office hours start at 8 am, and finish at 5.30 pm. Meet_Crowd video clip in context 2 contains some people walking into the hallway at point A. These are therefore deemed as anomalies.

All the video clips are concatenated into one large video stream. The video stream is then fed into the system being tested. By using this method, all systems would not be aware of the existence of these two different contexts. Apart from the video stream, each system is also given a stream of information about the current time.

### *B. Human behaviour feature description*

Both context-based and existing systems use the same human behaviour interest point based features proposed in [15] to represent human behaviour. Technically, interest point patches are extracted. Then, the tracking information provided by the dataset is used to associate the interest point patches to a person. These patches are then used to represent a person's behaviour.

Each person's behaviour is segmented into behaviour units which have lengths of one second. For example, if a person appears in the scene for 3 seconds, then his/her behaviour will be segmented into three different behaviour units. This segmentation is required in order to avoid making each behaviour too specific. Based on our observation, one second behaviour unit is sufficient for the dataset. Then, each behaviour unit is represented by the interest point patches extracted in the duration of the unit. This representation is able to distinguish between basic human actions and their direction of action (i.e. a person walking to the left is considered different from person walking to the right).

There are 292 behaviour units extracted from context 1 and 81 behaviour units extracted from context 2. These behaviour units were streamed into both systems.

TABLE III: The list of selected video clips for each context.

| Context | Video clip names |
|---|---|
| Context 1 | Rest_SlumpOnFloor, Rest_WiggleOnFloor, Meet_Split_3rdGuy, Browse_WhileWaiting1, Browse3, Browse4, Meet_WalkTogether2, Rest_WiggleOnFloor, Split, Walk3, Fight_OneManDown, Meet_WalkSplit, Browse_WhileWaiting2, Browse1, Meet_WalkTogether1, Rest_InChair and Browse2 |
| Context 2 | Fight_RunAway1, Fight_Chase, Rest_FallOnFloor and Meet_Crowd |

### C. Existing system description

The existing system implements the adaptive model approach proposed by Tao et al approach [10]. The system has a capability to adapt its normal behaviour model with the current context. A normal behaviour class can be reclassified into abnormal model, and vice versa. Technically, a weight is assigned to each behaviour class. This weight is increased whenever the incoming pattern (i.e. behaviour unit) is classified into the class. The weight also is decreased automatically due to the normalisation of the weights so that their sum must equal one.

We use the same parameter values as in their work [10]. Specifically, $Th_{w1}$, $Th_{w2}$ and $\alpha$ are set to 0.05, 0.25 and 0.1 respectively. $Th_{w1}$, $Th_{w2}$ and $\alpha$ are the minimum weight of a normal behaviour model to be still considered as normal, the maximum weight of an abnormal behaviour model, and the learning rate respectively. In order to construct Receiver Operating Characteristics (ROC) plot, we varied $Th_{\Lambda}$, the threshold deciding whether a behaviour pattern is abnormal.

### D. Context-based system description

Before describing the context-based system, a context space model needs to be constructed.

According to the scenario, the rate at which person walking into hallway at point A ($f_p$) and the office hour become the contextual information. The office hour will be represented in terms of time which is discreetised into hour units. So, the context space model can be presented as follows.

$$\begin{aligned} CI &= \{f_p,\ time\} \\ CI^1 &= \{time\} \\ CI^2 &= \{f_p\} \\ CSB &= CI^1 \ and \ CSI = CI^2 \end{aligned} \tag{3}$$

The context-based system takes an approach called reduction to point anomalies detection approach [2]. Technically, to detect contextual anomalies, the approach excludes information which is irrelevant to the current context. Then, a simple point anomalies detection is applied. The context-based system defines anomalies as any data which can be classified as an outlier. So, any outlier detection can be applied.

In order to exclude irrelevant information to the current context, the system utilises the context space model presented above. Since, it is assumed that the contexts are not known, then the system employs set representation. The representation will allow the system to discover new instances of context by monitoring the context parameters. A new instance of context is discovered when at least one of the context parameter values changes.

Finally, the context-based system employs a data stream clustering algorithm proposed in [16] to maintain frequency of occurrences of each behaviour class. A data stream clustering algorithm is chosen because it is able to update clustering results from the incoming data (i.e. behaviour unit). The algorithm determines whether the incoming data should be classified into one of the existing clusters or a new singleton cluster. Each cluster represents a behaviour class. The second reason is that because the cluster summary structure constructed by the algorithm enables the context-based system to retrieve information extracted from a given period of time. This can be achieved by periodically storing the clustering results into snapshots. Given two snapshots stored in two different time then to extract information between these two timestamps, the system employs a subtraction process which is described in [16]. By using the subtraction method, the system will be able to retrieve the behaviour classes and their members in a period of time between the two snapshots.

The system makes decision on every incoming behaviour unit. Initially, it classifies the behaviour unit into either one of the existing clusters or a new singleton cluster. Then, the commonality index value of the behaviour unit is calculated as follows.

$$CV(x, f_j) = \frac{1}{max_f} f_j \tag{4}$$

where x is the behaviour unit; $f_j$ is frequency of occurrence of the behaviour class in which x belongs; $max_f$ is the largest frequency of occurrence of the behaviour classes.

Once the commonality index value is calculated, the following simple thresholding rule is employed to decide whether the behaviour unit is abnormal.

$$CVLevel(CV_x) = \begin{cases} Normal & Th \geq CV_x \\ Abnormal & CV_x < Th \end{cases} \tag{5}$$

where $CV_x$ is the commonality index value of behaviour unit x and Th is predefined. In this experiment, Th was varied to generate the ROC plot.

### E. Results

In order to do comparative analysis between these two systems, we vary the $Th_{\Lambda}$ used in the existing system and Th which is used in the context-based system. The ROC plot is presented in figure 5. As we can see here, the context-based system clearly has a better performance than the existing system. This is because the capability of the context-based system to distinguish these two contexts and exclude the information extracted in context 1 when making decision in context 2.

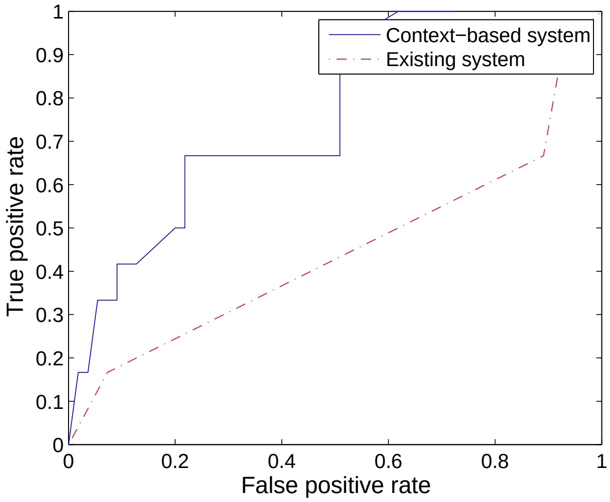


Fig. 5: Receiver Operating Characteristic (ROC) plot of context-based and existing systems.

In the second context, anomalous behaviours only appear in Meet_Crowd video clip. There are four people walking to point A. These people are detected as anomalous by context-based system as soon as they are walking toward point A. The existing system is unsuccessfull in detecting them as anomalies because the behaviour class belongs to these behaviour is classified into the normal behaviour model. This is because the existing system still utilises knowledge learned from context 1 which has a much larger number of occurrence of instances of this behaviour class when making decisions. Unlike the existing system, the context-based system only utilises knowledge learned from the current context.

Although the existing system employs model adaptation to reflect changes in visual context, knowledge learned in previous context is still considered when making decision. The knowledge will slowly be removed from the system over a period of time. Unlike the existing system, the context-based system automatically excludes the knowledge learned from other different contexts.

## V. Conclusions

One of the key aspect in evaluating the success of surveillance systems depends on their performance in detecting anomalous human behaviour which could lead to a security breach. Unfortunately, the current surveillane systems heavily rely on human observers. This limits the capability of these systems to become forefront crime fighting tools. The current automatic anomalous detection approaches which address these problems, employ various techniques starting from rule-based methods to statistics approaches. However, the use of contextual information in these approaches is still limited.

This paper proposed a context space model which provides guidelines for the system designers to determine which information could be used as contextual information. One of the primary requirements of a type of information could be used to describe a context is that there is a relationship between the information and the interpretation on human behaviour. Furthermore, any system implementing the proposed model will be able to distinguish between different contexts. Finally, it also is shown that the proposed model is a generalisation of the existing models in the computer vision and human computer interaction domains.

A comparative analysis was conducted in order to show the effectiveness of the proposed model for detecting anomalous human behaviour. To do this, a context-based system employing the context space model was implemented. The system was then compared to a system which employs an adaptive model proposed in [10]. Then the CAVIAR dataset was used to construct the experiment scenario containing two different contexts. From this experiment, it was shown that the context-based system performed better. This is due to the ability of the context-based system to distinguish different contexts and use only knowledge learned from the relevant context to detect anomalous behaviours.

## References


[1] A. Wiliem, V. Madasu, W. Boles, and P. Yarlagadda, “Detecting uncommon trajectories,” in *Computing: Techniques and Applications, 2008. DICTA ’08*, 2008, pp. 398–404.

[2] C. Varun, B. Arindam, and K. Vipin, “Anomaly detection: A survey,” *ACM Comput. Surv.*, vol. 41, no. 3, pp. 1–58, 2009.

[3] P. Bull, *Body Movement and Interpersonal Communication*. John Wiley and Sons Ltd, 1983.

[4] Z. Hua, S. Jianbo, and M. Visontai, “Detecting unusual activity in video,” in *Computer Vision and Pattern Recognition. CVPR. IEEE Conference on*, vol. 2, 2004, pp. 819–826.

[5] M. Marszalek, I. Laptev, and C. Schmid, “Actions in context,” in *Computer Vision and Pattern Recognition. CVPR. IEEE Conference on*, 2009, pp. 2929–2936.

[6] A. Torralba, “Contextual priming for object detection,” *Int. J. Comput. Vision*, vol. 53, no. 2, pp. 169–191, 2003.

[7] D. Duque, H. Santos, and P. Cortez, “Prediction of abnormal behaviors for intelligent video surveillance systems,” in *Computational Intelligence and Data Mining. CIDM. IEEE Symposium on*, 2007, pp. 362–367.

[8] C. Stauffer and W. E. L. Grimson, “Learning patterns of activity using real-time tracking,” *Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 8, pp. 747–757, 2000.

[9] Y. A. Ivanov and A. F. Bobick, “Recognition of visual activities and interactions by stochastic parsing,” *Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 8, pp. 852–872, 2000.

[10] X. Tao and G. Shaogang, “Incremental and adaptive abnormal behaviour detection,” *Computer Vision and Image Understanding*, vol. 111, no. 1, pp. 59–73, 2008.

[11] Jo, C. lle, L. C. James, D. Simon, and G. David, “Context is key,” *Communications of the ACM*, vol. 48, no. 3, pp. 49–53, 2005.

[12] T. M. Strat, “Employing contextual information in computer vision,” in *DARPA93*, 1993, pp. 217–229.

[13] I. Katz and H. Aghajan, “Exploring relationship between context and pose: A case study,” in *Cognitive Systems and Interactive Sensors*, 2007.

[14] M. Pantic, A. Pentland, A. Nijholt, and T. Huang, “Human computing and machine understanding of human behavior: A survey,” in *Artifical Intelligence for Human Computing*, 2007, pp. 47–71.

[15] P. Dollar, V. Rabaud, G. Cottrell, and S. Belongie, “Behavior recognition via sparse spatio-temporal features,” in *Visual Surveillance and Performance Evaluation of Tracking and Surveillance, 2005. IEEE International Workshop on*, 2005, pp. 65–72.

[16] C. C. Aggarwal, *Data Streams Models and Algorithms*. Springer, 2007.