http://dx.doi.org/10.5370/JEET.2012.7.1.115

Context-aware Video Surveillance System

Tae-Ki AN* and Moon-Hyun Kim†

Abstract – A video analysis system used to detect events in video streams generally has several processes, including object detection, object trajectories analysis, and recognition of the trajectories by comparison with an a priori trained model. However, these processes do not work well in a complex environment that has many occlusions, mirror effects, and/or shadow effects. We propose a new approach to a context-aware video surveillance system to detect predefined contexts in video streams. The proposed system consists of two modules: a *feature extractor* and a *context recognizer*. The *feature extractor* calculates the *moving energy* that represents the amount of moving objects in a video stream and the *stationary energy* that represents the amount of still objects in a video stream. We represent situations and events as motion changes and stationary energy in video streams. The *context recognizer* determines whether predefined contexts are included in video streams using the extracted *moving* and *stationary energies* from a *feature extractor*. To train each context model and recognize predefined contexts in video streams, we propose and use a new ensemble classifier based on the AdaBoost algorithm, DAdaBoost, which is one of the most famous ensemble classifier algorithms. Our proposed approach is expected to be a robust method in more complex environments that have a mirror effect and/or a shadow effect.

Keywords: Ensemble classifier, Surveillance, Context aware, Video analysis, AdaBoost

1. Introduction

In semantic video analysis, researchers define context so that they can recognize activities and identify different behaviors based on contextual information from video streams. Nagel [1] defined the context of an action analysis process as a complex structure, composed of spatial structures, temporal changes, and action intent. Strat [2] defined the context as all information that may influence the way a scene is perceived. Medioni et al. [3] defined contextual information as the accessory information used during processing to help the process complete its task efficiently. They used two kinds of context: spatial context and mission context. Spatial context contains the spatial structures of the scene, symbolic names, and the static reference object. Mission context provides a priori expectations regarding scenarios for detection, specific recognition methods, and their parameters.

Several steps are used t recognize contexts in video streams, including detection objects, analysis trajectories of the objects, and comparison with a priori context models [18-20]. Medioni et al. [3] used two modular blocks. The first one detects and tracks moving regions, and generates trajectories of moving objects. The second module takes these trajectories as input, together with user-provided information. These blocks continue to have several problems related to the analysis of a video stream that has many moving objects with occlusions in a complex environment. The first module, moving-object detection, does not work well in the environment. The *context recognizer* also does not accomplish its mission to detect contexts.

We propose a noble context-aware system to detect the contexts in video streams. Our system also has two modules: a *feature extractor* and a *context recognizer*. The *feature extractor* creates stationary energy (SE) and moving energy (ME) features instead of object trajectories. Our features represent a more effective solution than object trajectories to reflect unexpected influences in a background image because of mirror effects or shadow effects. In describing the contexts, we consider these effects to be latent features.

However, we need a process to select features that represent a context adequately, as well as a process to combine the selected features because the features are too massive to for direct use in construct recognizers. We use an ensemble classifier algorithm to perform the selection procedure and the combination procedure simultaneously. We propose a new Diverse AdaBoost algorithm as the ensemble classifier. Our proposed Diverse AdaBoost outperforms Gentle AdaBoost proposed by Viola et al. [14] to detect the facial region.

We first give an overview of the context-aware system in Section 2, and then describe in detail the features from

† Corresponding Author: School of Information & Communication Engineering, Sungkyunkwan University, Korea (mhkim@ece.skku.ac.kr)
* Urban transit Research Center, Korea Railroad Research Institute, Korea / School of Information & Communication Engineering, Sungkyunkwan University, Korea (tkahn@krri.re.kr)
Received: October 21, 2010; Accepted: August 16, 2011

video streams in Section 3. The proposed Diverse AdaBoost algorithm is presented in Section 4 and the results of experiments are discussed in Section 5. We conclude the paper in Section 6.

2. Context-Aware System

Our context-aware video surveillance system has two modules: a *feature extractor* and a *context recognizer*. The overview of our proposed context-aware system is given in Fig. 1. The *feature extractor* generates ME or SE. ME represents the differences between a current video frame and the previous video frame, and SE represents the differences between a current video frame and the reference video frame.

The *context recognizer* detects the contexts in video streams using the features from the *feature extractor* with trained ensemble classifiers.

2.1 Feature extractor

The *feature extractor* acquires necessary information to represent a context. Medioni et al. [3] detected moving objects and analysis trajectories of the objects with this module. Application of this method is only possible to an environment with few objects and no occlusions. In practice, moving objects, reflections, and shadows could change the environment. In the environment, separating objects to find the trajectories of the objects correctly can be quite difficult.

We consider changes in video streams as variation of energy to apply a real-time video surveillance system in a complex environment. Bradski et al. [4] used motion history images, accumulated motion energy images, to determine the current pose of the object and to segment and measure the motions induced by the object in a video scene.

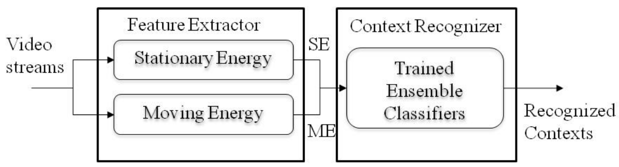


Fig. 1. Overview of the proposed context-aware system

We define and use two kinds of energy, SE and ME, to represent the contexts. Contexts about stable objects and moving objects are generally important in a video surveillance system. SE is used to express stable objects features and ME is used to express moving object features.

2.2 Context recognizer

The c*ontext recognizer* determines whether a predefined context is included by using the extracted features from the *feature extractor*. The recognizer has as many classifiers as predefined contexts. Each trained ensemble classifier decides that the context related the classifier is included in a video stream. We need to train each classifier related to each context. Medioni et al. [3] extracted basic contexts, such as mobile object properties and distance to an object of reference using extracted objects from the *feature extractor* and trajectories of objects. Subcontexts can be combined to recognize more complex contexts such as passing or avoiding a checkpoint.

We define the context model as a classifier trained by an ensemble algorithm, such as AdaBoost. Viola et al. [14] used Gentle AdaBoost algorithm to detect regions of the face and have high efficiency. We developed a new Diverse AdaBoost algorithm to improve the performance of the classifier by increasing diversity between classifiers. To confirm performance, we use UCI machine repository data, and show test results in Section 5.

3. Moving and Stationary Energy

We use moving and stationary energy to represent moving and static objects in video. The energies are features as inputs of train classifier models to recognize the contexts in video streams. We can use the energies to detect each object, and to detect more complex contexts. Each context model is trained with the energies. We consider all contexts in video streams as changes of energy instead of trajectories of each object. These features can be presented more efficiently in a complex environment.

3.1 Moving energy

Bobick et al. [5] defined a motion energy image as a vector threshold version computed for optic flow fields to prevent noise in motion computation from corrupting the process. Adelson et al. [6] proposed a temporal-spatial energy model that manages motion in x-y-t space to describe motion features, and a method to detect motion by direction in the temporal-spatial space.

We define ME as the difference between a current image and the previous image. Moving energy in region *i*, $ME(R_i)$, is calculated as

$$ME(R_i) = \frac{\sum_{(x,y)\in R_i} |I_c(x,y) - I_p(x,y)|}{\sum_{(x,y)\in R_i} I_c(x,y)} \tag{1}$$

where $I_c(x,y)$ is the intensity on pixel (x,y) in the current image, and $I_p(x,y)$ is the intensity on pixel (x,y) in the previous image. The image is divided into N regions, where R_i is the ith region of the image.

3.2 Stationary energy

We define the stationary energy as the difference between a current image and a reference image. Stationary energy in region i, $SE(R_i)$, is measured as

$$SE(R_i) = \frac{\sum_{(x,y)\in R_i} | I_c(x,y) - I_r(x,y) |}{\sum_{(x,y)\in R_i} I_c(x,y)} \tag{2}$$

where $I_c(x,y)$ is the intensity on pixel (x,y) in the current image, and $I_r(x,y)$ is the intensity on pixel (x,y) in the reference image which is also the background image. The image is divided in regions, where R_i is the i^{th} region of the image.

4. Diverse AdaBoost Model

The AdaBoost algorithm is one of the most famous algorithms to make ensemble classifiers by selecting weak member classifiers. The AdaBoost algorithm generates a strong classifier combined with several weak classifiers. AdaBoost finds a combination of weak classifiers by adjusting the weight through a repetitive process without changing the original training data set. One of the reasons the AdaBoost algorithm yields good results is the diversity among weak classifiers. However, no definite standard to measure diversity in the process exists [7, 14].

We propose a new Diverse AdaBoost algorithm called DAdaBoost, which adds the process considering diversity to generate an ensemble classifier. Whereas AdaBoost selects a member classifier to minimize the error in each cycle, DAdaBoost selects a member classifier not only to minimize error, but also to maximize the diversity among the member classifiers in each cycle. DAdaBoost selects candidate classifiers by measuring the difference between minimum error and the error of each weak classifier in each cycle. If the difference is smaller than a threshold, we select the weak classifier as a candidate classifier. We then calculate the diversity between the ensemble classifiers generated in the previous cycle and each candidate classifier. DAdaBoost can improve the generalization performance of the ensemble classifiers by considering diversity while sacrificing the accuracy of a weak classification in each cycle.

4.1 Diversity and accuracy

According to Krogh and Vedelsby [17], the relationship between error and diversity (ambiguity) is

$$E = \overline{E} - \overline{A} \tag{3}$$

where E is the ensemble generalization error, $\overline{E}$ is the weighted average of the individual classifier generalization errors, and $\overline{A}$ is the weighted average of the diversity, which they call the ensemble ambiguity. The diversity on input x of a single member of the ensemble is defined as

$$d_i(x) = \left(\overline{f}(x) - f_i(x)\right)^2 \tag{4}$$

$$\overline{f}(x) = \sum_i w_i f_i(x) \tag{5}$$

where $f_i(x)$ is the output of the classifier i, $\overline{f}(x)$ is a weighted ensemble average, and w_i is the weight of $f_i(x)$. From (3), we know that the ensemble generalization error is decreased as diversity increase and individual generalization errors decrease. They show the relationship between the diversity and accuracy of the ensemble. The measures cannot be straightforward because they can be calculated as parameters from final ensemble classifiers.

Many authors emphasize and evaluate the relationship between the diversity and accuracy of ensemble classifiers. However, no strict method exists for using the diversity while designing the regression ensemble classifiers. Li et al. [9] proposed a method to increase diversity of AdaBoost with SVM-based weak classifiers. In their method, in each cycle, the weak classifier with minimum error is selected. If D calculated by (6) and (7) is greater than a predefined threshold, DIV, the weak classifier is selected as a member of the ensemble classifier, or otherwise discard the weak classifier. The diversity measure proposed by Melville and Mooney [8] is used. The measures are based on disagreement/agreement measures. The diversity of the tth component classifier on input $\mathbf{x}_i$ is calculated as

$$d_t(\mathbf{x}_i) = \begin{cases} 1 & \text{if } h_t(\mathbf{x}_i) \neq f(\mathbf{x}_i) \\ 0 & \text{otherwise} \end{cases} \tag{6}$$

$$D = \frac{1}{TN} \sum_{t=1}^{T} \sum_{i=1}^{N} d_t(\mathbf{x}_i) \tag{7}$$

where T is the number of member classifier of the ensemble classifiers, N is the number of training samples, $h_t(\mathbf{x})$ is the prediction label of classifier t, and $f(\mathbf{x})$ is the combined prediction label of all the existing component classifiers.

However, they might not obtain a member classifier every cycle, and do not present the method to define DIV. They cannot always get ensemble classifiers that satisfy the condition with a predefined DIV. The amount of operations to calculate the diversity they use is increased according to the number of component classifiers.

Melville and Mooney [8] presented a method that directly constructs diverse member classifiers using additional artificially constructed training examples. The concept of the measures of diversity is similar; however, they should be redefined for the regression ensemble model, such as AdaBoost. This is a cumbersome procedure to make artificial data, and a very repetitive procedure to get the member classifiers for ensembles.

We have developed the disagreement measure as the degrees of diversity between classifiers on the given training data set. The degree of diversity between two classifiers on input $\mathbf{x}_k$ is calculated as

$$d_{i,j}(\mathbf{x}_k) = \begin{cases} 1 & \text{if } h_i(\mathbf{x}_k) \neq h_j(\mathbf{x}_k) \\ 0 & \text{otherwise} \end{cases} \quad (8)$$

$$D_{i,j} = \frac{1}{N}\sum\nolimits_{k=1}^{N} d_{i,j}(\mathbf{x}_k) \quad (9)$$

where $D_{i,j}$ is the generalization average degree of the diversity between classifier i, h_i, and classifier j, h_j, and N is the number of training samples. We calculate the diversity between not all the combined ensemble classifiers but only one ensemble classifier in j^{th} iteration and the candidate classifier. The amount of the number of calculation is not increased according to the number of the component classifiers because the diversity is independent of the number of the component classifiers.

4.2 AdaBoost

The AdaBoost algorithm generates a strong classifier from weak classifiers. It consists of several training steps. At each step, every weak classifier classifies all of the provided training samples. It then selects a weak classifier h_t that has the minimum average classification error. The weight for each training sample is modified as follows.

The weight for sample $\mathbf{x}_i$ is decreased if h_t classifies $\mathbf{x}_i$ correctly; otherwise, the weight is increased. After training T steps, the set of weak classifiers selected at each step is called an ensemble classifier. The weighted combination of weak classifiers belonging to the ensemble classifiers becomes the strong classifier. AdaBoost is regarded as a good algorithm to construct ensemble classifiers to minimize average generalization error. To improve the performance of the AdaBoost algorithm, we need to expand diversity among the weak classifiers of the final ensemble classifiers.

We propose a method to enhance diversity among ensemble classifiers, and to show a performance improvement compared to Gentle AdaBoost algorithm that aims at minimizing classification errors.

A typical AdaBoost algorithm is Gentle AdaBoost algorithm (see Table 1) [15]. This algorithm selects a small number of critical visual features from a very large set of potential features because of learning. AdaBoost provides an effective learning algorithm and strong bounds on generalization performance. The result shows a set of features with significant variety. The AdaBoost algorithm restricts the weak learner to a specific set of classification

Table 1. Gentle AdaBoost algorithm

1. Input : Training data set $\{(\mathbf{x}_1, y_1), (\mathbf{x}_2, y_2), \ldots, (\mathbf{x}_n, y_n)\}$, with labels $y_i \in \{+1, -1\}$

2. Initialize the weights of the training samples: $w_{1,i} = 1/n$, $\mathrm{i} = 1, \ldots, n$

3. Do this for $\mathrm{t} = 1, \ldots, T$
 - For each feature, train a classifier h_j that is restricted to using a single feature f_j
 - Calculate the error of the weak classifier
 $\varepsilon_j = \sum\nolimits_{i=1}^{n} w_{t,i} \mid h_j(\mathbf{x}_i) - y_i \mid$
 - Choose the classifier h_t with the lowest error ε_t
 - Update the weights of the training samples
 $w_{t+1,i} = \dfrac{w_{t,i}\beta_t^{1-e_i}}{C_t}$
 where $e_i = 0$ if example $\mathbf{x}_i$ is classified correctly, $e_i = 1$ otherwise, $\beta_t = \dfrac{\varepsilon_t}{1-\varepsilon_t}$, and C_t is a normalization constant.

4. Create a strong classifier
$H(x) = \begin{cases} 1 & \text{if } \sum\nolimits_{t=1}^{T} \alpha_t h_t(\mathbf{x}) \geq \frac{1}{2}\sum\nolimits_{t=1}^{T} \alpha_t \\ 0 & \text{otherwise} \end{cases}$ where $\alpha_t = \log\dfrac{1}{\beta_t}$

functions, each depending on a single feature. It uses each weak classifier as a threshold unit on a single feature, or single node decision tree.

4.3 Diverse AdaBoost

We present the proposed weak classifier selection method that considers not only classification errors but also diversity among classifiers. We compare the method with the typical classifier selection method that only tries to minimize errors.

Diversity is the opposite of accuracy, and is inversely proportional to accuracy. In other words, if the diversity rises, the accuracy falls, and vice versa.

By controlling the diversity and accuracy of a weak classifier in balance, we can generate ensemble classifiers with good performance. In Table 1, a weak classifier training method in Step 3 is used to minimize classification errors. In each iteration t, the weighted error of h_j is computed and used to update the distribution of weights on the training samples. The error is the sum of the weights of each incorrectly classified sample, and calculated as

$$\varepsilon_j = \sum\nolimits_{i=1}^{n} w_{t,i} \mid h_j(\mathbf{x}_i) - y_i \mid \quad (10)$$

where $w_{t,i}$ is the weights of $\mathbf{x}_i$ at t^{th} iteration step, h_j is a weak classifier j, and y_i is an expected output class on the input $\mathbf{x}_i$.

Minimizing classification errors cannot guarantee optimized ensemble classifiers since it does not reflect diversity at all. Thus, we need a plan that can add diversity additionally into the ensemble classifiers. We proposed an intelligent method to select a weak classifier at each iteration step, which not only satisfies error-minimization criteria, but also diversity criteria. After training, it results in diverse ensemble classifiers.

The Gentle AdaBoost algorithm chooses an ensemble classifier that has minimum error in each cycle. The DAdaBoost algorithm shown in Table 2 selects a member classifier of the ensemble classifiers through two steps. In the first step, it generates a candidate member set using the error of the weak classifiers and the lowest error in each cycle. We calculate the difference between the error of each weak classifier and the lowest error. If the difference of the weak classifier is under the threshold, the weak classifier is selected as the candidate classifier in the cycle.

We create the candidate classifier set in each iteration t, G_t, defined as

$$G_t = \left\{ h_j \mid |\varepsilon_j - \varepsilon_t| < \varepsilon_{thr},\ j = 1, \ldots, m \right\} \tag{11}$$

where h_j is candidate weak classifiers, ε_j is the error of the weak classifier j, ε_t is the minimum error in iteration t, ε_{thr} is a threshold value, and m is the number of features.

In the second step, DAdaBoost calculates diversity between the ensemble classifier generated in the previous cycle and each candidate classifier. We use the measure of the diversity as (8) and (9). We select the weak classifier that has the largest diversity as the member classifier of the ensemble classifier from the candidate classifier set in the iteration step.

Table 2. Diverse AdaBoost algorithm

1. Input : Training data set $\{(\mathbf{x}_1, y_1), (\mathbf{x}_2, y_2), \ldots, (\mathbf{x}_n, y_n)\}$, with labels $y_i \in \{+1, -1\}$

2. Initialize the weights of the training samples : $w_{1,i} = 1/n$, $i = 1, \ldots, n$

3. Do this for $t = 1, \ldots, T$
- For each feature, train a classifier h_j that is restricted to using a single feature
- Calculate the error of the weak classifier
$\varepsilon_j = \sum_{i=1}^{n} w_i \mid h_j(\mathbf{x}_i) - y_i \mid$
- Find the lowest error ε_t
- Create the candidate classifiers set, G_t
$G_t = \{h_j \mid |\varepsilon_j - \varepsilon_t| < \varepsilon_{thr},\ j = 1, \ldots, m\}$
- Calculate diversity between the ensemble classifier in the previous cycle, $h_p(\mathbf{x}_i)$, and each candidate classifier, $h_k(\mathbf{x}_i)$
$D_{p,k} = \frac{1}{n} \sum_{i=1}^{n} d_{p,k}(\mathbf{x}_i),\ k = 1, \ldots, L$
where $d_{p,k}(\mathbf{x}_i) = \begin{cases} 1 & \text{if } h_p(\mathbf{x}_i) \neq h_k(\mathbf{x}_i) \\ 0 & \text{otherwise} \end{cases}$
- Choose h_t with the largest diversity among the candidate classifiers, and update the lowest error ε_t
- Update the weights of the training samples :
$w_{t+1,i} = \frac{w_{t,i} \beta_t^{1-e_i}}{C_t}$
where $e_i = 0$ if example $\mathbf{x}_i$ is classified correctly, and $e_i = 1$ otherwise; and C_t is a normalization constant.

4. Create a strong classifier
$H(x) = \begin{cases} 1 & \text{if } \sum_{t=1}^{T} \alpha_t h_t(\mathbf{x}) \geq \frac{1}{2} \sum_{t=1}^{T} \alpha_t \\ 0 & \text{otherwise} \end{cases}$ where $\alpha_t = \log \frac{1}{\beta_t}$

5. Experiments

We performed experiments to verify the performance of our proposed DAdaBoost algorithm with UCI Machine Learning Repository data sets and the effectiveness of the proposed approach in detecting contexts from video streams.

5.1 DAdaBoost performance test

Our proposed Diverse AdaBoost, DAdaBoost, is compared with Gentle AdaBoost. Three benchmark data sets from the UCI Machine Learning Repository are used to evaluate the generalization performance of DAdaBoost.

Table 3 lists the specifications of data sets used in the experiment.

Table 3. Characteristics of the three data sets

Data set	Samples	Attributes	Classes
Sonar	208	60	2
Ionosphere	351	34	2
Diabetes	768	8	2

We use the decision stump consisting of a single-level decision tree as the weak classifier learning model. Table 4 shows the data set characteristics in the experiment. The experiments are performed a 1,000 times repetition process using a fivefold cross-validation method, and we test altering ε_{thr} from 0.01 to 1.0 with an increment step size of 0.01. The weak classification algorithm uses the decision stump method that has one node and is the simplest method in the decision tree.

Table 4. Performance results

Alg.	Generalization Average Error		
	Sonar	Ionosphere	Diabetes
GA	0.18350±0.01120	0.08674±0.00551	0.25060±0.00481
DA0.01	0.18094±0.01322	0.09178±0.00603	0.24832±0.00391
DA0.02	0.17338±0.01623	0.08788±0.00585	0.24822±0.00566
DA0.03	0.16912±0.01609	0.08122±0.00666	0.23953±0.00457
DA0.04	0.17269±0.01306	0.08099±0.00654	0.24951±0.00461
DA0.05	0.17942±0.01583	0.8339±0.00653	0.25997±0.00648
DA0.06	0.16881±0.01331	0.08412±0.00626	0.25600±0.00584
DA0.07	0.17650±0.01651	0.08007±0.00675	0.25232±0.00431
DA0.08	0.17381±0.01665	0.07533±0.00837	0.26362±0.00409
DA0.09	0.17526±0.01364	0.07663±0.01086	0.26793±0.00570
DA0.10	0.17904±0.01915	0.07811±0.00848	0.27329±0.00703

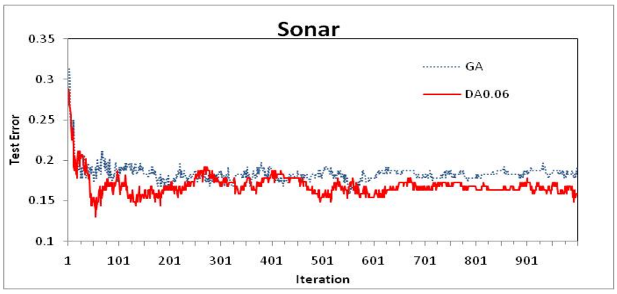
We also use Gentle AdaBoost algorithm as a comparison target algorithm well known by its excellent performance. We compare the performance test result of this algorithm and that of DAdaBoost, the proposed algorithm.

Table 4 shows the experimental results, including generalization errors with standard deviation of algorithms. DAdaBoost performs noticeably better than Gentle AdaBoost in some cases. The error threshold to be sacrificed is a very significant element of the average generalization error.

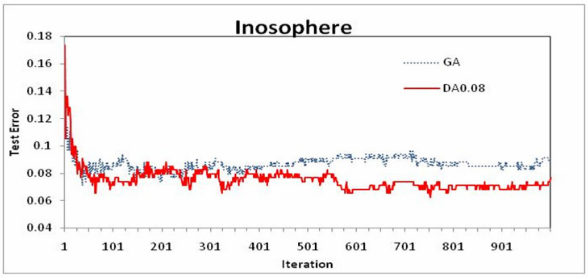
With regards to the Sonar data set, Diverse AdaBoost is superior to Gentle AdaBoost for all cases of ε_{thr}. When ε_{thr} is 0.01 and 0.02, the performance of the Diverse AdaBoost is low, but it is mostly superior to the other values in the ionosphere data set.

The proposed Diverse AdaBoost is much better than the Gentle AdaBoost in two data sets because the diversity obtained from the suggested algorithm is much bigger than that obtained from weak classification selected by Gentle AdaBoost with a minimum error in the repetition process. This experiment shows the performance of the weak classifiers. In the diabetes data set, the Diverse AdaBoost is superior only when ε_{thr} is from 0.1 to 0.4 and the Gentle AdaBoost threshold is much better in the other values because the weak classifiers generated by Gentle AdaBoost are sufficiently diverse in the diabetes data set.

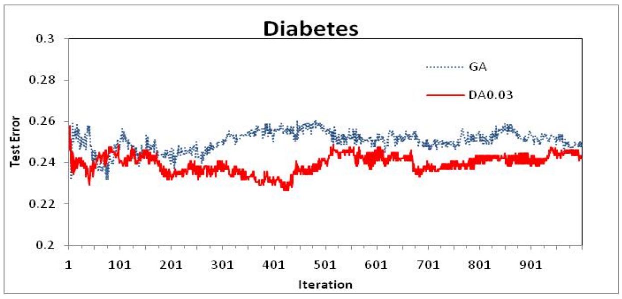
We know that DAdaBoost outperforms Gentle AdaBoost in all of the three data sets, if we determine a proper sacrifice range of error (see Fig. 2).



(a)



(b)



(c)

Fig. 2. Performance test results of Gentle AdaBoost and DadaBoost: (a) with ε_{thr} = 0.06 (DG0.06) on test subset Sonar; (b) with ε_{thr} = 0.08 (DG0.08) on test subset Ionosphere; (c) with ε_{thr} = 0.03 (DG0.03) on test subset Diabetes

5.2 Context-aware system performance test

We perform experiments to verify the performance of our context-aware video surveillance system using video streams of a subway platform of the Daegu Metro Line 1 in Korea. We use four different data sets: Train Arrival, Train Departure, Boarding Passengers, and Waiting Passengers (see Table 5). We show the sample images of each data set in Fig. 3. The size of the image used is 480 by 640 pixels, and the image is divided into 12 (3 by 4) regions. The size of each region is 160 by 160 pixels. The *feature extractor* calculates ME and SE of each region of the video streams.

Table 5. Characteristics of the subway data sets

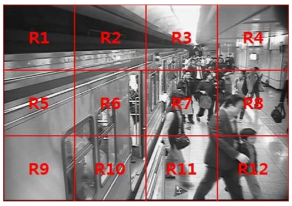
Data set	Samples	Attributes
Train Arrival	146	24
Boarding Passengers	248	24
Train Departure	219	24
Waiting Passengers	452	24

We generate the training models for each data set using the DAdaBoost algorithm on the 24 features, 12 MEs and 12 SEs from the *feature extractor*.

We show SE and ME from the *feature extractor* on the Passengers Boarding data set of the Daegu Metro Line 1 in Korea in Fig. 4. Although SE and ME are very simple attributes to represent the features of video streams, they are sufficient inputs to construct our ensemble classifiers.

The *context recognizer* determines the contexts on the

(a)

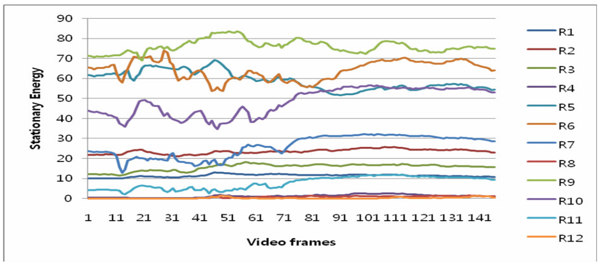


(b)

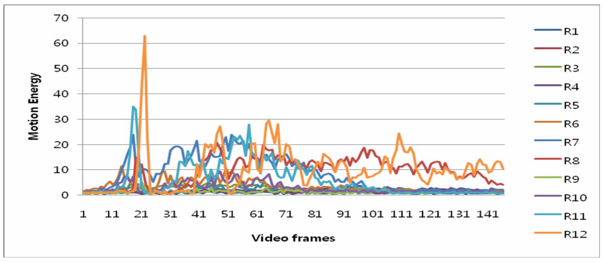
(c)

(d)

Fig. 3. Sample image of (a) Train arrival; (b) Boarding passengers; (c) Train departure; (d) Waiting passengers data set of Daegu Metro Line 1 in Korea



(a)



(b)

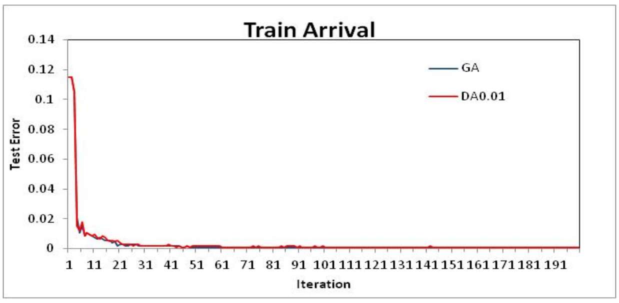
Fig. 4. The (a) SE; (b) ME from the *feature extractor* on Fig. 3(b)

input vectors, and each vector is composed of 12 MEs and 12 SEs from the *feature extractor*. Each experiment repeats the process 200 times using a fivefold cross-validation method. Then, ε_{thr} is changed from 0.01 to 1.0. We show the results of the performance experiments for the four contexts of subway platform as you can see in Table 6.

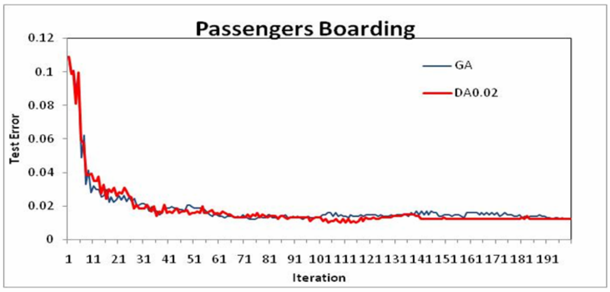
DAdaBoost outperforms Gentle AdaBoost in case of the Boarding Passengers and Train Departure data sets, but not in the case of the Train Arrival data set. Gentle AdaBoost is much better than DAdaBoost on the Train Arrival because the diversity obtained from DAdaBoost is not enough to

Table 6. Performance results of the proposed approaches

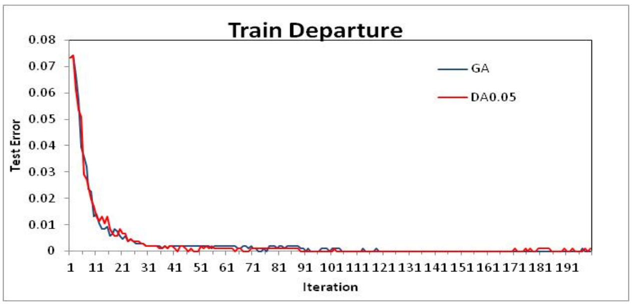
Alg.	Generalization Average Error		
	Train Arrival	Passengers Boarding	Train Departure
GA	0.00335±0.01361	0.01876±0.01390	0.00332±0.01067
DA0.01	0.00352±0.01359	0.01907±0.01469	0.00327±0.01059
DA0.02	0.00366±0.01373	0.01792±0.01475	0.00327±0.01084
DA0.03	0.00382±0.01452	0.02105±0.01500	0.00347±0.01061
DA0.04	0.00401±0.01408	0.01942±0.01446	0.00394±0.01030
DA0.05	0.00421±0.01492	0.01971±0.01439	0.00318±0.01057
DA0.06	0.00430±0.01492	0.02034±0.01481	0.00413±0.01027
DA0.07	0.00455±0.01491	0.02187±0.01404	0.00427±0.01063
DA0.08	0.00437±0.01491	0.02175±0.01451	0.00485±0.01085
DA0.09	0.00445±0.01491	0.02347±0.01527	0.00424±0.01107
DA0.10	0.00430±0.01513	0.02420±0.01530	0.00430±0.01058



(a)



(b)



(c)

Fig. 5. The performance test results of Gentle AdaBoost and DAdaBoost on (a) Train Arrival with ε_{thr} = 0.01 (DG0.01); (b) Boarding Passengers with ε_{thr} = 0.02 (DG0.02); (c) Train Departure with ε_{thr} = 0.05 (DG0.05).

compensate for the sacrifice of the accuracy in each cycle. We show the performance of Gentle AdaBoost and DAdaBoost on each data set of the Daegu Metro Line 1 in Korea in Fig. 5. The results of the experiment show that our proposed context-aware video surveillance system works well in the complex subway environment with many occlusions and mirrored surfaces.

6. Conclusion

We propose a new context-aware video surveillance system. The system consists of a *feature extractor* and a *context recognizer*.

We define the energies as a change of intensity, and consider the energies as features to detect contexts. ME and SE can represent changes, not only because of moving objects, but also because of the mirrored surface and the occlusion. We also introduce the ensemble classifier algorithm to train context models and to recognize contexts.

We propose DAdaBoost as an ensemble classifier that can add to the process of considering diversity when it selects a weak classifier in each cycle. It can improve the performance of the ensemble classifier by acquiring diversity while sacrificing the accuracy of a weak classifier in each cycle. We have verified the performance of the DAdaBoost using three data sets from the UCI machine Learning Repository database, and the performance of the context-aware video surveillance system using the video streams of the subway platform. We expect that our approaches can be used to construct an intelligent video surveillance system to detect contexts.

Acknowledgements

This research was supported by a grant (07 Urban Transit Standardization A01) from "The 2nd Phase of R&D on the Urban Transit Standardization" funded by the Ministry of Land, Transport and Maritime Affairs of the Korean government.

References

[1] H.H. Nagel, "From Image Sequences Towards Conceptual Descriptions," *in Image and Vision Computing*, vol. 6, no. 2, pp. 59-74, May 1988.

[2] Thomas M. Strat, "Employing Contextual Information in Computer Vision", *In Proceedings of ARPA Image Understanding Workshop*, pp. 217-229, 1993.

[3] Gerard Medioni, Isaac Cohen, Francois Bremond, Somboon Hongeng and Ramakant Nevatia, "Event Detection and Analysis from Video Streams", *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 23, vo. 8, pp. 873-889, Aug 2001.

[4] Gary R. Bradski and James W. Davis, "Motion segmentation and pose recognition with motion history gradients", *Int. Journal Machine Vision and Applications*, vol. 13, vo. 3, pp. 174-184, 2002.

[5] Aaron Bobick and James Davis, "an appearance-based representation of action", *in Proceedings of international conference on Pattern Recognition*, vol. 1, pp. 307-312, Vienna, Austria, Aug 1996.

[6] Edward H. Adelson and James R. Bergen, "Spatiotemporal energy models for the perception of motion", *Journal Optical Society of America*, vol.2, vo.2, pp. 284-299, Feb. 1985.

[7] Robert E. Schapire and Yoram Singer, "Improved Boosting Algorithms Using Confidence-rated Predictions" Machine Learning, vol. 37, No. 3, pp. 297-336, Dec. 1999.

[8] Prem Melville and Raymond J. Mooney, "Creating diversity in ensembles using artificial data", *Journal of Information Fusion*, vol. 6, No. 1, pp. 99-111, Mar. 2004.

[9] Xuchun Li, Lei Wang and Eric Sung, "AdaBoost with SVM-based component classifiers", *Engineering Applications of Artificial Intelligence 21*, pp. 785-795, 2008.

[10] Lumila I. Kuncheva and Christopher J. Whitaker, "Measures of Diversity in Classifier Ensembles and Their Relationship with the Ensemble Accuracy", *Machine Learning*, vol. 51, no. 2, pp. 181-207, May 2003.

[11] Louisa Lam, "Classifier Combinations: Implementations and Theoretical Issues", Multiple Classifier Systems, *Lecture Notes in Computer Science,* vol. 1857, pp. 78-86, Cagliari, Italy, 2000.

[12] Ron Kohavi, "A study of cross-validation and bootstrap for accuracy estimation and model selection", *in Proceedings of the 14th International Joint Conference on Artificial Intelligence*, pp. 1137-1143, 1995.

[13] Thomas G. Dietterich, "An Experimental Comparison of Three Methods for Constructing Ensembles of Decision Trees: Bagging, Boosting, and Randomization", *Machine Learning*, vol. 40, No. 2, pp. 139-157, Aug. 2000.

[14] Dianhong Wang and Liangxiao Jiang, "An improved attribute selection measure for decision tree induction", *in Proceedings of Fourth International Conference on Fuzzy Systems and Knowledge Discovery*, vol. 4, pp. 654-658, 2007.

[15] Paul Viola and Michael Jones, "Robust Real-time Object Detection", *International Journal of Computer Vision*, vol. 27, no. 2, pp. 137-154, 2004.

[16] Tin Kam Ho, "The Random Subspace Method for Constructing Decision Forests", *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 20, no. 8, pp. 832-844, Aug 1998.

[17] Anders Kroggh and Jesper Vedelsby, “Neural Network Ensembles, Cross Validation and Active Learning”, *Advances in Neural Information Processing Systems*, vol. 7, pp. 231-238, 1995.
[18] Ki-Yeol Eom, Tae-Ki AN, Gyu-Jin Kim, Gyu-Jin Jang, Moon-Hyun Kim, “Fast Object Tracking in Intelligent Surveillance System”, *LNCS 5593*, pp. 749-763, July, 2009.
[19] Ki-Yeol Eom, Tae-Ki AN, Gyu-Jin Kim, Gyu-Jin Jang, Jae-Young Jung, Moon-Hyun Kim, “Hierarchically Categorized Performance Evaluation Criteria for Intelligent Surveillance System”, in *Proceedings of the 2009 International Symposium on Web Information Systems and Applications*, pp. 223-226, May. 2009.
[20] Dong-Min Woo, Quoc-Dat Nguyen, “3D Building Detection and Reconstruction from Aerial Images Using Perceptual Organization and Fast Graph Search”, *Journal of Electrical Engineering & Technology,* Vol. 3, No. 3, pp. 436-443, Sep. 2008.

Tae-Ki AN received his MSc degree in Electronic Engineering from Kyungpook National University, Daegu, Korea, in 1996, and his PhD degree in Department of Electrical and Computer Engineering from Sungkyunkwan University, Seoul, Korea in 2011. He is a Chief Researcher in Korea Railroad Research Institute. His research interests are artificial intelligence, pattern recognition, and video analysis.

Moon-Hyun Kim received his MSc degree in Electrical Engineering from KAIST, Daejeon, Korea, in 1980, and the PhD degree in Computer engineering from University of Southern California in 1988. He is a full Professor in School of Information & Communication Engineering at Sungkyunkwan University. His research interests are artificial intelligence, digital image processing, and pattern recognition.