Moving Object Classification in Infrared and Visible Spectra

Rania Rebai Boukhriss[a], Emna Fendri[b], Mohamed Hammami[b]
[a] Miracl-ISIMS, Sfax University, Sakiet Ezzit, Sfax, Tunisia
[b] Miracl-FS, Sfax University, Road Sokra, Sfax, Tunisia

ABSTRACT

This paper introduces a novel method of moving object classification in Infrared and Visible spectra. This method is based on a data-mining process by combining a set of best features based on shape, texture and motion. The proposed method relies either on visible spectrum or on infrared spectrum according to weather conditions (sunny days, rain, fog, snow, etc.) and timing of the video acquisition. Experimental studies are carried out to prove the efficiency of our predictive models to classify moving objects and the originality of our process with intelligent fusion of VIS-IR spectra.

Keywords: Moving object classification, machine learning, intelligent fusion, infrared spectrum, visible spectrum.

1. INTRODUCTION

The classification of moving objects in video sequences is a main step in smart computer vision systems. The goal is to improve the reliability of various automatic applications such as behaviors analysis, abnormal event detection and human re-identification, etc. Indeed, object classification in outdoors areas is still a challenging and attractive research topic due to the complexity and variety of its challenges, like variations of human postures, the color variation in human clothes and cars, the total or partial occlusions of objects, the large variation in weather conditions (such as snow, fog, rain, etc.) and lighting conditions. In the literature, object classification methods can be categorized primarily into four main approaches: Shape-based methods [1], Texture-based methods [2], Motion-based methods [3] and Hybrid methods [4]. The shape-based methods use objects' 2D spatial information such as the area, bounding box, silhouette and gradient of detected objects regions [1]. These methods are based either on machine learning techniques which exploit shape features extracted from the detected moving regions [5], or on the matching of models built onto candidate foreground regions [6]. The texture-based methods refer to the description of regions in an image by their texture in terms of the spatial variation in pixel intensities such as the Local Binary Pattern (LBP) [7], the histogram of oriented gradients (HOG) or the Gabor filter [2]. The motion-based methods use temporal features like periodicity, direction and speed of tracking objects for the classification task. They are used especially to distinguish non-rigid objects (humans) from rigid objects (vehicles) [8]. Finally, the idea of hybrid approach is to combine the shape, texture and/or motion analysis simultaneously to better exploit the advantages of each approach [4]. These methods use generally a machine learning technique for the moving object classification. In this paper, we introduce a new method to classify moving objects based on hybrid approach in Infrared (IR) and Visible (VIS) spectra using a data-mining process by combining the best features based on shape, texture and motion to ensure better classification of moving objects. The originality of our method lies in the intelligent fusion of VIS-IR spectra where the classification is performed either on visible spectrum or on infrared spectrum according to weather conditions (sunny days, rain, fog, snow, etc.) and timing of the video acquisition. Thus, the VIS spectrum is used in sunny days under normal weather conditions, while the IR spectrum will be used at night or in presence of fog, rain, snow, etc. The rest of the paper is arranged as follow. The outlines of our proposed approach are detailed in the next section. The experimental findings and results will be reported in section 3. Finally, the recapitulation of the presented approach and the future works are given in Section 4.

2. PROPOSED APPROACH

Our approach is composed of two main stages: off-line stage and on-line stage. In the first stage we adopt a data-mining process to construct two prediction models: one for classifying moving objects in VIS spectrum and the second for the IR spectrum. In the online stage, a classification of the detected moving objects is performed on one of the two spectra based on the decision of our method of intelligent fusion [9]. Fig. 1 shows the framework of the proposed approach.

Ninth International Conference on Machine Vision (ICMV 2016), edited by Antanas Verikas, Petia Radeva,
Dmitry P. Nikolaev, Wei Zhang, Jianhong Zhou, Proc. of SPIE Vol. 10341, 1034104
© 2017 SPIE · CCC code: 0277-786X/17/$18 · doi: 10.1117/12.2268414

Downloaded From: http://proceedings.spiedigitallibrary.org/ on 07/31/2017 Terms of Use: http://spiedigitallibrary.org/ss/termsofuse.aspx

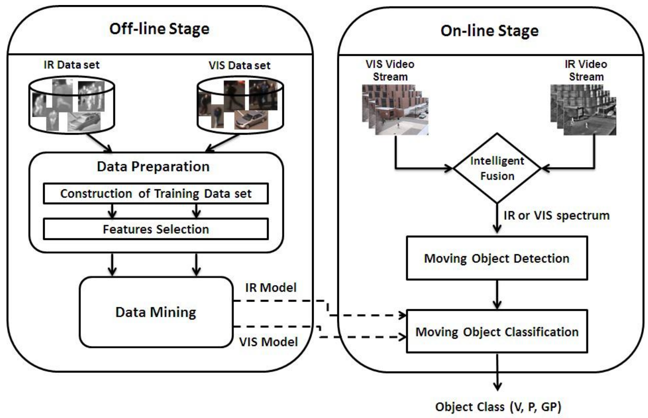


Figure 1. Proposed method for moving object detection and classification using an intelligent fusion of IR-VIS spectra.

2.1 Off-line Stage

2.2 Data Preparation

The study of the state of the art has allowed us to identify efficient features based on shape, texture and motion to ensure a better classification of moving objects and to build two tables of two-dimensional from our training corpus. In our work, the data preparation step is composed of two sub-steps. In fact, we start firstly with the construction of the training dataset. Secondly, a variable selection is performed to choose the efficient features which best discriminate the three classes of moving objects namely: Pedestrian (P), Vehicle (V) and Pedestrian Group (PG). The data-mining process for the classification of moving objects requires a large and representative training and test corpus. Our corpus consists of 13 famous IR and VIS sequences[1,2] recorded in different scenes and in various weather conditions that contain moving objects of three classes (P, V and PG). This database is divided to construct our training and testing data sets. The training data contains a total of 4583 of class P, 849 of class V and 1093 of class PG in each spectrum IR and VIS. Hence, the test data set is composed of 4682 of class P, 108 of class V and 996 of class PG in both spectra allowing the validation of the most appropriate prediction model for moving object classification. A thorough look on the shape, texture and motion features allowed the selection of ten features. In fact, we selected seven shape based features namely Height, Width [10], Aspect Ratio [11], Area [12], Perimeter [13], Compactness and Anthropomethry [11]; two famous texture based features that are the Histogram of Oriented Gradients (HOG) [11] and the Local Binary Pattern (LBP) [7] descriptors and one motion based features that is the velocity of the object [4].

A feature selection step is crucial to determine the best features which improve the classification performance. To achieve this selection, we used the Relief algorithm based filter approach, introduced by Kira and Rendell [14]. Due to its simplicity and effectiveness, the Relief algorithm is considered as one of the most successful algorithms for the evaluation of the quality of the features. Moreover, this algorithm has been proved as one of the fastest procedures for features selection [15].

2.3 Data Mining

The aim of the data mining step is to generate two predictive models: VIS predictive model and IR predictive model. These two models are constructed using supervised learning techniques and the features set defined above in order to classify moving objects in one of three classes P, V, or PG in IR and VIS spectra. In supervised learning, to generate a generic and efficient classifier, we must address three requirements: the number of relevant features, the size of the training set and the choice of the appropriate learning technique. Through a thorough study of the most used supervised learning techniques for classification task, we have selected 4 different learning algorithms: the Multilayer Perceptron Neural Network (MLP), SVM with Radial Basis Function Kernel (RBF Kernel), SVM with polynomial kernel (Poly Kernel) and Decision Trees. Each algorithm is considered as reference in its category. In fact, the MLP network is a popular learning algorithm since it defines the desired output and adjusts weight coefficients in such way that the calculated and desired outputs are as close as possible [16]. The most illustrative example of local neural system is the

[1] http://vcipl-okstate.org/pbvs/bench/
[2] http://www.ino.ca/Video-Analytics-Dataset

Downloaded From: http://proceedings.spiedigitallibrary.org/ on 07/31/2017 Terms of Use: http://spiedigitallibrary.org/ss/termsofuse.aspx

Support Vector Machine (SVM) of distinctive kernel functions. A critical step in support vector machine classification is choosing a suitable kernel. Indeed, choosing different kernel functions will produce different SVMs and may result in different performances [16]. In the literature, there are four types of kernel function used in SVM namely: linear kernel, polynomials kernel, sigmoid kernel and RBF kernel. It is well known that the two typical kernel functions are the radial basis function kernel and polynomial kernel [17]. For this, in our work, the SVM is experimented with polynomial kernel and RBF kernel. Concerning the decision tree algorithms, they have been very successful to solve large number of real world problems in data mining and machine learning. In 1986, Quinlan [18] has developed the first decision tree algorithm known as ID3. Later he proposed the C4.5 algorithm, a benchmark to which other learning algorithms have proved its performance. Therefore, we have selected the C4.5 learning technique based decision tree to compare their prediction models to the other prediction models constructed by the three learning techniques detailed above.

2.4 On-line Stage

In the online stage of our proposed approach, we start by classifying the VIS images in abnormal or normal weather condition using our intelligent fusion method of VIS-IR spectra. Then, according to the decision, we perform the detection and classification of moving objects relying on VIS spectrum or on IR spectrum.

2.5 Intelligent Fusion of IR-VIS spectra

In our previous work [9], we proposed an intelligent fusion method of VIS-IR video to benefit from the quality of both of them. In fact, the detection and classification of moving objects is carried out either on VIS spectrum or on IR spectrum according to weather conditions and timing of the video acquisition. Thus, the VIS spectrum is used in sunny days under normal weather conditions, while the IR spectrum will be used at night or in presence of fog, rain, snow, etc. In fact, a construction of the adequate prediction model for abnormal weather classification is performed. Then, the VIS images are classified into image in Normal conditions or in Abnormal conditions to detect Moving objects in IR spectrum or in VIS spectrum. Our experimentations on several sequences with different weather conditions, presented in [9], have proven the effectiveness of the generated prediction model and the originality of our intelligent fusion method.

2.6 Moving Object Detection

For the detection of a moving object, we adopted a method based on background modeling with a dynamic matrix and spatio-temporal analyses of scenes [19]. This method has proven its effectiveness in foreground segmentation and shown high performances facing a lot of important challenges such as sudden and gradual illumination changes, shaking camera, background component changes, ghost and foreground speed.

2.7 Moving Object Classification

The goal of this step is to classify each moving object detected either in the VIS spectrum or in the IR spectrum into one of the three classes: Pedestrian, Vehicle or Pedestrian Group. This classification is based on one of the two prediction models extracted from the offline step. In fact, if the detection of moving objects has been performed in the VIS spectrum, thus we use the VIS predictive model to classify moving objects. In the other case, we use the IR predictive model to classify moving objects detected from the IR spectrum.

3. EXPERIMENTAL STUDY

In order to evaluate the performance of the proposed approach, we have carried out three series of experiments. The first series presents the experiments concerning the selection of the most discriminative features in order to determine the best features with which the classification performance is the best. The second series deal with the experiments concerning the choice of the adequate prediction models in each spectrum VIS and IR. The third series seeks to evaluate the performance of the selected predictive models to classify of moving objects by proving the efficiency of the intelligent fusion of VIS-IR spectra.

3.1 First series of experiments

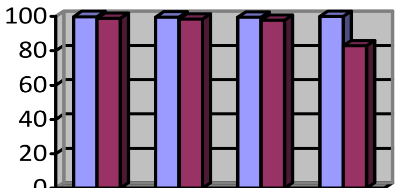
This experiment aims to evaluate the impact of features selection step on moving object classification results. To select the best features that improve the classification performance, we applied the Relief algorithm on our training dataset. By varying the sample data size, the Relief algorithm provides ordered weighted variables. We have distinguished from these results that both features of LBP and Perimeter have always low weights whatever the size of the sample data. For this, we have compared the moving object classification results on the test data set with two

Downloaded From: http://proceedings.spiedigitallibrary.org/ on 07/31/2017 Terms of Use: http://spiedigitallibrary.org/ss/termsofuse.aspx

different feature vectors. The first one contains our ten features detailed above in section 2.1.1. In the second feature vector we eliminated the LBP and the Perimeter features which is the result of our features selection algorithm. The results of this experiment prove the importance of our features selection step. In fact, the correct classification rate has improved from 97,41 to 99,01 in IR spectrum and from 96,39 to 99,14 in VIS spectrum by eliminating both features the LBP and the Perimeter.

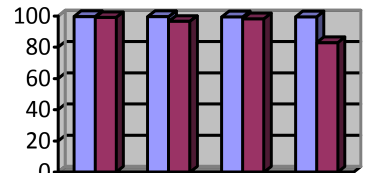
3.2 Second series of experiments

Regarding the construction of the best prediction models in each spectrum VIS and IR, we have proceeded to select the most adequate learning technique. In this study we have used different classifiers such as Multilayer Perceptrons (MLP), Support Vector Machines (SVM) with polynomial kernel and RBF kernel and C 4.5 based decision trees. These data mining algorithms were compared according to the rate of Total Correct Classification (TCC). The results of this experimentation are shown in figure 2 in which the best rates of TCC on the learning and test data sets are achieved by the SVM with polynomial kernel. In fact, the best classification rates obtained on the test data set are 99,14 for the IR spectrum and 99,01 for the VIS spectrum.



	SVM(Poly Kernel)	SVM(RBF Kernel)	MLP	C4.5
Training Set	99.92	99.68	99.54	99.97
Test Set	99.01	98.55	97.81	82.94

(a)



	SVM(Pol y Kernel)	SVM(RBF Kernel)	MLP	C4.5
Training Set	99,79	99,72	99,4	99,52
Test Set	99,14	96,68	98,34	82,94

(b)

Figure 2. Moving object classification results in VIS (a) and IR (b) spectra with different machine learning techniques.

3.3 Third series of experiments

The moving object classification is performed either on VIS spectrum or on IR spectrum according to weather conditions and timing of the video acquisition. Thus, the VIS spectrum is used in sunny days under normal weather conditions, while the IR spectrum will be used at night or in presence of fog, rain, snow, etc. In order to evaluate our proposed approach, we compared the classification results on the test data set using only the IR spectrum or the VIS spectrum against the intelligent fusion results. The use of our intelligent fusion method allowed to improve the TCC rate from 99,14 in IR spectrum and 99,01 in VIS spectrum to 99, 22. Moreover, some of qualitative results of our approach are shown in Figure 3. In the first three frames (a-c), we used the IR prediction model because these frames are classified with our intelligent fusion method as in abnormal weather conditions. In the last three images (d-f), the moving objects are classified with the VIS prediction model according to the decision of our method of intelligent fusion which has classified these frames as in normal weather conditions. These experimental results have proved the genericity of our prediction models and the effectiveness of our intelligent fusion method of IR-VIS spectra.

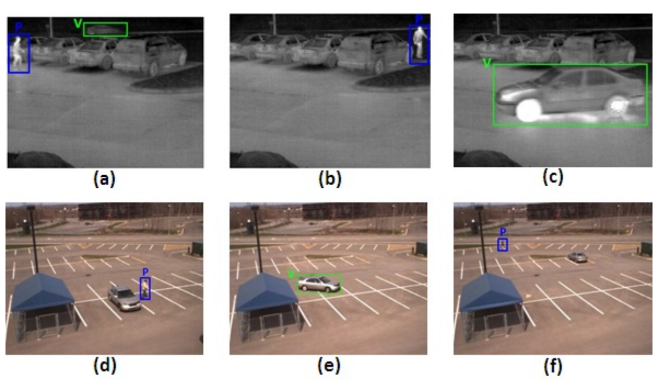


Figure 3. Qualitative Results of moving object classification with the intelligent fusion of VIS and IR spectra.

Downloaded From: http://proceedings.spiedigitallibrary.org/ on 07/31/2017 Terms of Use: http://spiedigitallibrary.org/ss/termsofuse.aspx

4. CONCLUSION

In this paper, we proposed a novel method for moving object classification using VIS spectrum or IR spectrum according to weather conditions (darkness, sunny days, fog, snow, etc.). For thus, in the offline stage we first extract a set of best features based on shape, texture and motion and then we generate a prediction model for each spectrum by supervised learning techniques. In the online stage, the classification of detected moving objects is performed on one of the two given spectrum at once according to the decision of our intelligent fusion method. Experimentations carried out on a large and representative training and test data sets has proven the efficiency of our prediction models for classifying moving objects based on an intelligent fusion of VIS-IR spectra. However, the development of a robust tracking algorithm seems necessary to address the occlusion problem.

REFERENCES

[1] Dedeoğlu, Y., Töreyin, B.U., Güdükbay, U. and Çetin, A.E., "Silhouette-Based Method for Object Classification and Human Action Recognition in Video," Int. Conf. Compt. Vision in Human-Computer Interaction, 64-77 (2006).

[2] Moctezuma, D., Conde, C., Diego, I.M. and Cabello, E., "Person detection in surveillance environment with HoGG: Gabor filters and histogram of oriented gradient," ICCV Workshops, 1793–1800 (2011) .

[3] Wang, L., Hu, W. and Tan, T., "Recent developments in human motion analysis," Pattern Recognition, 36(3), 585-601 (2003).

[4] Zhang, Z., Cai, Y., Huang, K. and Tan, T., "Real-time moving object classification with automatic scene division," Proc. IEEE ICIP, 149–152 (2007).

[5] Changalasetty, S. B., Badawy, A. S., Ghribi, W., Ashwi, H. I., Al-Shehri, A. M., Dhafer, A., Al-Shehri, A., Thota, L. S. and Medisetty, R., "Identification and Classification of Moving Vehicles on Road," Computer Engineering and Intelligent Systems, 4(8), 1-12 (2013).

[6] Raj Bharath, R. and Dhivya, G., "Moving Object Detection, Classification and its Parametric Evaluation," Int. Conf. on Information Communication and Embedded Systems, 1-6 (2014).

[7] Zhang, L., Li, S.Z., Yuan, X. and Xiang, S., "Real-time object classification in video surveillance based on appearance learning," IEEE Computer Society Conf. on Computer Vision and Pattern Recognition, 1–8 (2007).

[8] Dedeoğlu, Y., "Moving Object Detection, Tracking and Classiffication for smart video surveillance," Doctoral thesis (2004).

[9] Boukhriss, R.R., Fendri, E. and Hammami, M., "Intelligent Fusion of Infrared and Visible Spectrum for Video Surveillance Application," 14th Int. Conf. on Artificial Intelligence and Soft Computing, 768-777 (2015).

[10] Apatean, A., Rogozan, A. and Bensrhair, A., "Kernel and Feature Selection for Visible and Infrared based Obstacle Recognition," 11th IEEE Int. Conf. on Intelligent Transportation Systems, 1130-1135 (2008).

[11] Elhoseiny, M., Bakry, A. and Elgammal, A., "MultiClass Object Classification in Video Surveillance Systems-Experimental Study," IEEE Conf. on Computer Vision and Pattern Recognition Workshops, 788-793 (2013).

[12] Gurwicz, Y., Yehezkel, R. and Lachover, B., "Multiclass object classification for real-time video surveillance systems," Pattern Recognition Letters, 32(6), 805-815 (2011).

[13] Hota, R. N., Venkoparao, V. and Rajagopal, A., "Shape Based Object Classification for Automated Video Surveillance with Feature Selection," 10th Int. Conf. on Information Technology, 97-99 (2007).

[14] Kira, K. and Rendell, L.A., "A practical approach to feature selection," 9th Int. Conf. Machine Learning, 249–256 (1992).

[15] Dash, M. and Liu, H., "Feature selection for classification,"Intelligent Data Analysis,1(1), 131-156 (1997).

[16] Biswas, S. K. and Mia, M. M. A., "Image Reconstruction Using Multi Layer Perceptron (MLP) And Support Vector Machine (SVM) Classifier And Study Of Classification Accuracy," International Journal of Scientific & Technology R; esearch, 4(2), 226-231 (2015).

[17] Zanaty, E.A., "Support Vector Machines (SVMs) versus, Multilayer Perception (MLP) in data classification," Egyptian Informatics Journal , 13, 177–183 (2012).

[18] Quinlan, J.R., "Induction of Decision Trees," Machine Learning, 1, 81-106 (1986).

[19] Hammami, M., Jarraya, S. K. and Ben-Abdallah, H., "On line background modeling for moving object segmentation in dynamic scene," Multimedia tools and applications, 63(3), 899-926 (2013).

Downloaded From: http://proceedings.spiedigitallibrary.org/ on 07/31/2017 Terms of Use: http://spiedigitallibrary.org/ss/termsofuse.aspx