Automatic detection of adverse weather conditions in traffic scenes

Andrea Lagorio, Enrico Grosso
Computer Vision Laboratory
DEIR - University of Sassari, Italy
`{lagorio,grosso}@uniss.it`

Massimo Tistarelli
Computer Vision Laboratory
DAP - University of Sassari, Italy
`tista@uniss.it`

Abstract

Visual surveillance in outdoor environments requires the monitoring of both objects and events. The analysis is generally driven by the target application which, in turn, determines the set of relevant events and objects to be analyzed. In this paper we concentrate on the analysis of outdoor scenes, in particular for vehicle traffic control. In this scenario, the analysis of weather conditions is considered to signal particular and potentially dangerous situations like the presence of snow, fog, or heavy rain. The developed system uses a statistical framework based on the mixture of Gaussians to identify changes both in the spatial and temporal frequencies which characterize specific meteorological events. Several experiments performed on standard databases and real scenes demonstrate the applicability of the proposed approach.

1. Introduction

Visual surveillance systems are based on different methodologies and pertain a series of applications ranging from traffic control to human body tracking. One of the major challenges in visual surveillance is the ability to i) define the features of the object or event to be detected and ii) define a computational model to extract such features from an image sequence.

Dealing with dynamic scene analysis, several methods exist to analyze the motion in the images. The motion of 3D objects in space induces an apparent motion on the image plane. The apparent motion can be either computed on the basis of the displacement of few image features points or as a dense field of displacement vectors over the entire image plane. This vector field is often called optical flow $\vec{\mathbf{v}} = (u, v)$. A fundamental problem in the computation of the optical flow is the inherent ambiguity due to the 3D to 2D projection. Whenever a gray level pattern is moving on the image plane, the image motion can be due to a number of different phenomena in 3D space. Particularly in the decade spanning from 1985 to 1995, an abundance of algorithms were developed to compute the instantaneous velocity from image sequences, under different hypotheses [1-5].

Even though it is possible to extract several dynamic features from the optical flow field, the analysis of the optical flow for event detection is rather difficult because of the inherent ambiguities in the motion field. For example, it is quite difficult to distinguish the ego-motion and the eco-motion from the optical flow computed from an image sequence acquired from a moving camera [6,7]. Other techniques directly based on the image differencing for background subtraction [8] are very sensitive to high frequency motion and are ineffective for slow motion detection. Moreover these techniques can not cope with the motion of the camera.

For these reasons several techniques have been proposed which directly compute dynamic features from the image sequence. Among them the most commonly used method is based on the *mixture of Gaussian* (MOG) model [9,10]. In this framework the rationale is to statistically characterize each single pixel in the image, through a set of Gaussian probability distributions. Through this model it is possible to define the variability of each pixel over time. This methodology, coupled with a Fourier analysis of the frequency content of the input sequence, has been applied to determine a series of anomalous events. In particular, to identify potentially dangerous weather conditions, namely snow fall and fog. In the reminder of the paper the mixture of Gaussians model is further explained and the technique adopted to identify anomalous events is detailed. Several experiments from real scenes are presented.

2. Statistical analysis of image features

Each image pixel in an image sequence can be statistically characterized as a series of values changing over time:

$$\{X_1,..., X_t\} = \{I(x_0, y_0, t) : 1 \le t \le T\} \quad (1)$$

Where *I(x,y,t)* represents the intensity value of the pixel at position *(x,y)* and time *t*, in the image sequence.
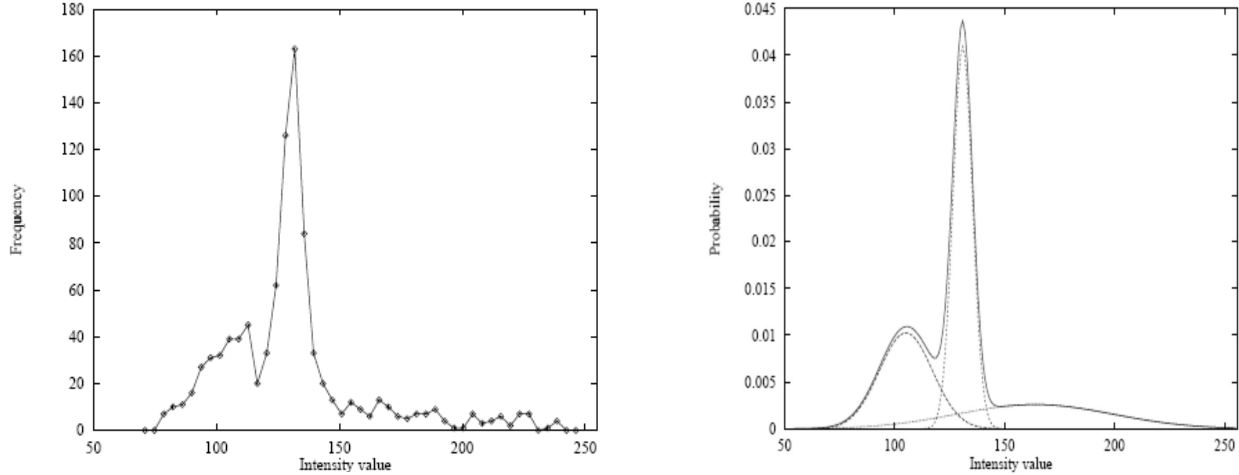
IEEE computer society

Figure 1: (left) Time histogram of an image pixel in a sequence and (right) the corresponding mixture of three Gaussian probability distributions. The ordinates report the frequency of each intensity value over time.

The recent evolution of the intensities can be modeled as the mixture of K Gaussian probability density distributions:

$$P(\mathbf{X_t}) = \sum_{i=1}^{K} \omega_{i,t} \cdot \eta(\mathbf{X_t}, \mu_{i,t}, \Sigma_{i,t}) \qquad (2)$$

Where K is the number of Gaussian distributions, $\omega_{i,t}$ is the individual weight of each Gaussian at time t (the sum being equal to 1), $\mu_{i,t}$ and $\Sigma_{i,t}$ are the mean the covariance matrix associated with the ith Gaussian at time t. The general expression of the Gaussian distribution is:

$$\eta(\mathbf{X_t}, \mu_t, \Sigma_t) = \frac{1}{(2\pi)^{\frac{n}{2}}|\Sigma_t|^{\frac{1}{2}}} \mathbf{e}^{-\frac{1}{2}(\mathbf{X_t}-\mu_t)^{\mathbf{T}}\Sigma_t^{-1}(\mathbf{X_t}-\mu_t)} \qquad (3)$$

In figure 1 an example of the time histogram of an image pixel over time and the related mixture of 3 Gaussian probabilities is shown.

At every frame, each pixel density is updated according to its intensity value. The update is performed by changing the Gaussian distribution from the pool with the mean value at a distance less than 2.5 times its standard deviation. This value has been experimentally found to be optimal to characterize the time variability of the pixel intensity [10]. The probability density distribution is modified by changing the weight associated to each Gaussian according to the following law:

$$\omega_{\mathbf{k,t}} = (1-\alpha)\omega_{\mathbf{k,t-1}} + \alpha(\mathbf{M_{k,t}}) \qquad (4)$$

Where $M_{k,t}$ is a binary value which determines if the considered pixel belongs to one of the Gaussians, α is a parameter associated with the learning capability of the distribution, and the sum of all K weights is equal to 1.

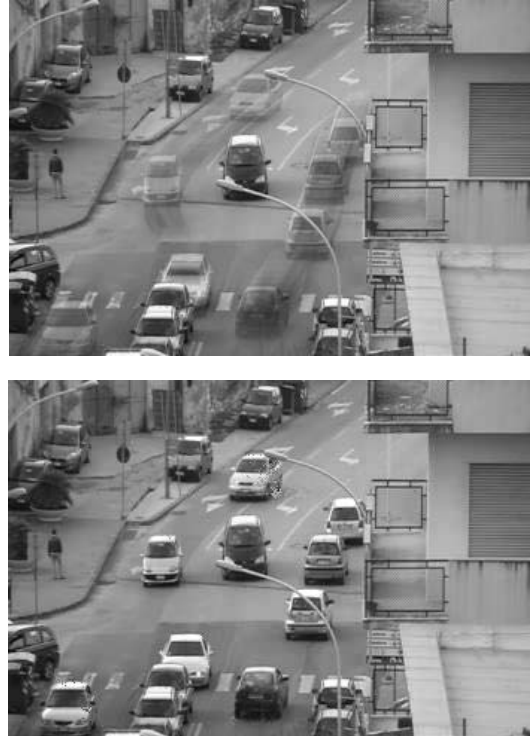In particular, the value of α determines the speed

Figure 2: Initialization of the Gaussian probabilities with the mean values (top) and the most recurrent values (bottom). In both cases the displayed intensities are the mean values of the Gaussian associated with the background.

274

Figure 3: Mean values of the background after 400 frames (top) and 1800 frames (bottom). It is worth noting how fast moving objects disappear as they are discarded by the model.

adaptation of the distribution to local changes. This is a key element to tune the model to the desired temporal frequency and thus detect either slow or fast objects from the background. The parameters of the selected Gaussian are updated according to:

$$\mu_t = (1 - \rho)\mu_{t-1} + \rho \mathbf{X}_t$$
$$\sigma_t^2 = (1 - \rho)\sigma_{t-1}^2 + \rho (\mathbf{X}_t - \mu_t)^T (\mathbf{X}_t - \mu_t) \quad (5)$$
$$\rho = \alpha\eta (\mathbf{X}_t \mid \mu_k, \sigma_k)$$

After the update, the Gaussians are ordered according to the ratio ω/σ. Having the lower weight and higher variance, the first Gaussian is always the one representing the (stationary) background in the scene. The advantage of this model over other methods is the capability to dynamically learn new "objects" or variations and incorporate them in the background model. The time to be elapsed for a new object to be "learned" depends on the tuning of the parameter α.

The model initialization is critical to correctly characterize the scene and the events. The initial parameters associated to each Gaussian are either determined by computing the mean values of the first Gaussian from a set of images or by selecting the Gaussian associated with the background as the one which is the most recurrent within the entire field of view. Assumed that the background covers a significant portion of the field of view, both methods perform almost equally well. An example of applying the two initialization methods is shown in figure 2 for a traffic scene, using 3 Gaussian distributions to model the sequence. The correct scene characterization is learned by the model after few frames. It is worth noting that slowly moving vehicles are merged with the steady objects in the scene, but gradually disappear as the model is updated. This effect is clearly evident in figure 3 as the ghosts of the moving vehicles is still visible in the foreground. The maximal velocity of the objects to be included in the steady background is determined by the learning speed of the Gaussian model.

In some of the experiments color image sequences have been used and the mixture model has been modified to handle color pixel values as well. In order to train the MOG with the color values, 9 Gaussian probability density functions are used which are merged, according to the relevance of each RGB color channel, into three main densities to code the variability of each pixel. The analysis of the color values proved to be more robust with respect to acquisition noise than the processing of 8 bit intensity values.

3. Detection of adverse weather conditions

Within the framework of traffic monitoring, it is important to detect anomalous or potentially dangerous events for the vehicles. Among all possible events, the detection of typical weather conditions is of great interest to deliver proper information to the drivers and also for timely alert of police patrols, thus reducing the probability of car accident or queues [11-15].

Differently from moving objects and other dynamic events, the change in the weather conditions affects the entire field of view. At the same time, it is important to still be able to identify moving objects, for example to classify cars from trucks and pedestrians. Toward this end it is necessary uniquely characterize the spatial and temporal features of each weather condition and how they affect the image sequence. A proper analysis of the spatio-temporal frequencies in the sequence allows to identify the different weather conditions.

In order to properly characterize the performances of the proposed system a mixed sequence has been produced including 62 frames for each different weather condition and mixed together in a single sequence 186 frames long. The resulting sequence reports all possible combinations in the transition between sunny to snow, and to fog.
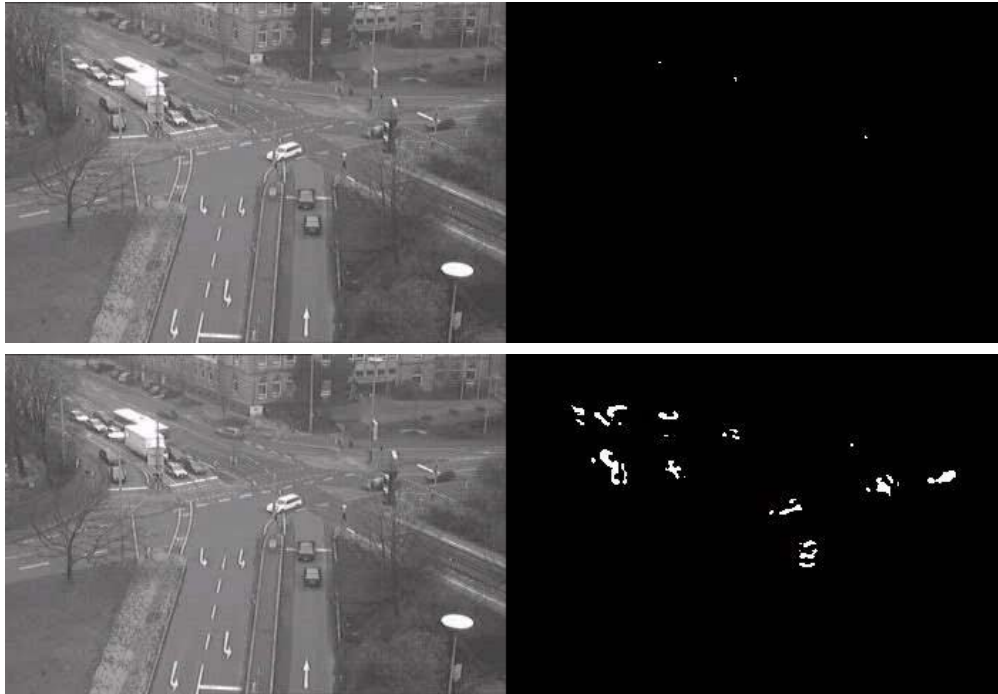
275

Figure 4: Traffic scene acquired during good weather and processed with a MOG model with α equal to 0.9 (top) and 0.01 (bottom). It is worth noting that no objects are detected with the high value of α, while the vehicles are correctly detected with the lower value.

Figure 5: Traffic scene in figure 4 but acquired during snow fall and processed with a MOG model with α equal to 0.9 (top) and 0.01 (bottom). The falling snow flocks are clearly detected with the high value of α, while the vehicles are still correctly detected with the lower value.

276

The error rate in the identification of the varying weather conditions is reported in table 1. Most of the reported errors are due to the unnatural abrupt transition between different weather conditions which can not be immediately accounted by the MOG model. In real weather changing conditions, the smooth change in the environment would produce a much lower error rate. As an example, processing a single weather condition, such as either snow fall or fog, the classification error is always equal to 0%

3.1. Snow fall detection

Snow falling is characterized by a number of small image blobs crossing the field of view in a single direction. This weather condition can be detected by selecting and counting the objects in the scene which are small in size and moving relatively fast across the image. The snow flocks will move, in general, faster than any other object in the scene, because they are very close to the camera lens.

The snow flocks can be detected simply by tuning the parameter α in equation (4) to adjust the learning speed of the mixture of Gaussians to the spatio-temporal frequency of the falling snow flocks. The discriminating value of the parameter can be determined by analyzing the same scene under different weather conditions. A thorough analysis has been carried out on several image sequences containing different weather conditions. In case of good weather conditions a very small value of α can be used, which is always below 0.1 (10% of the range). In case of snow the optimal value of α has been determined equal or greater than 0.9 (90% of the range). For this reason, a value of α equal to 0.9 has been applied to detect the presence of snow.

The results obtained from the analysis of the snow fall are presented in figure 4 and 5. The processed sequences has been obtained from the University of Karlsruhe public database of traffic scenes, which is a standard database for the evaluation of algorithms for traffic monitoring (http://i21www.ira.uka.de/image_sequences/).

The sequence contains at least 200 frames, with image resolution of 320x240 pixels with 24 bits per pixels to code the pixel color values. In order to train the MOG with the color values of the image pixels, 9 Gaussian probability density functions are used which are merged together in three main densities to code the variability of each pixel. In the processed sequences, the same scene view was captured under different weather conditions (snow fall and sunny) and analyzed with a MOG with α equal to 0.9 and 0.01. As it can be noticed, while the first value of α allows to detect the snow flocks, the second value clearly detects all moving objects.

Figure 6: Traffic scene reported in figure 4, under good weather conditions (top) and with fog (bottom).
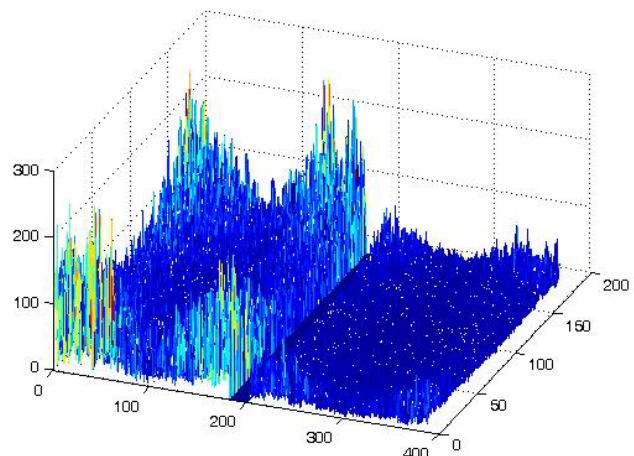

Figure 7: Amplitude of the Fourier spectrum for the two images in figure 6. The picture corresponding to the street scene with fog is represented by the section of the spectrum on the right hand side of the graph.

277

3.2. Detection of fog

Unlikely snow fall, the detection of fog requires the identification of a relevant change in the spatial frequencies of the entire image. In other terms, the effect of fog is a general blurring of the image which is almost equivalent to a low pass filtering. To detect this condition it is necessary to detect the change in the frequency content of the entire background image over time.

In order to analyze the distribution of the spatial frequencies in the image the Fourier Transform is computed over time and compared with the Gaussian distribution values associated with the stable background. The change in the amplitude of the Fourier spectrum over time determines if, and to what extent, the frequency content is changing within the entire field of view.

Figure 6 shows a sample traffic scene from the University of Karlsruhe public database, acquired from the same view point in good weather conditions and with fog. The Fourier spectrum for the two images is shown in figure 7. As expected, while many high frequency components are accounted in the image corresponding to the good weather, the frequencies computed from the fog image are much lower and uniformly distributed.

To detect the foggy weather condition it is required to determine if the Fourier spectrum is a simple plateau or it contains the series of peaks corresponding to a sharp image. The average of the values of the amplitude of the Forurier spectrum is generally sufficient. In the example presented in figure 7 the computed mean values of the Fourier transform amplitude are:

- 2927 for the image in the good weather condition,

- 1486 for the foggy weather condition.

The resulting difference between the two mean values is more than 50%, which is sufficient to set a robust threshold to discriminate the two conditions.

In order to further test the fog detection method, a sequence of 200 frames, with 320x240 RGB color images coded with 24 bit per pixels has been acquired from a camera placed at the window of a residential building, and overlooking a street. Two sample frames with and without fog are shown in figure 8.

To detect the foggy weather conditions, the reference mean frequency corresponding to the good weather were first computed. The first images with good weather were processed computing the Fourier transform of the pixel values corresponding to the Gaussian distribution associated with the stable background. The average of the Fourier spectrum was set as reference index V_{gw} for the good weather condition. The rest of the sequence was processed computing, every 10 frames, the mean value of the Fourier amplitude spectrum. This value constitute the index for the fog condition V_{fw}. Whenever the value of V_{fw} drops to 40% of V_{gw} the image is tagged as captured in fog

weather condition. Conducting this analysis on the locally captured sequence a classification error equal to 0% was reported.

The classification results of the snow and fog weather conditions are summarized in table 1.

Table 1: *Error rates in the classification of adverse weather conditions in the two experimental sequences.*

Weather condition	Classification Error
Fog sequence (200 frames)	
Fog / Sunny	0%
Mixed weather sequence (186 frames - from the U. Karlsruhe database)	
Fog	1.6%
Snow fall	19.8%
Sunny	1.6%

To detect the changing weather conditions in the second experiment reported in table 1, the two methods for the
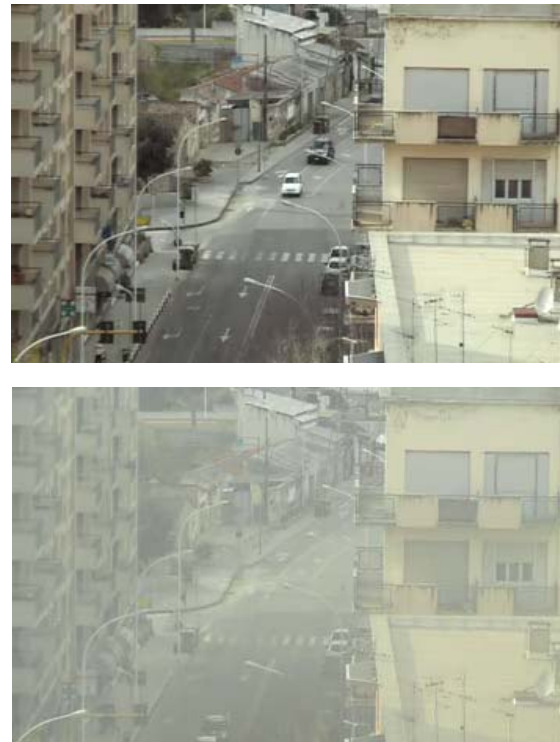

Figure 8: Traffic scene from a sequence of 200 frames, under good weather conditions (top) and with fog (bottom).

278

detection of fog and snow have been applied simultaneously to the entire sequence.

It is worth noting that the mixed weather sequence is composed of three different video streams recorded at different times. Therefore, the poor recognition rate accounted for the detection of snow fall is mainly due to the abrupt changes within the sequence.

4. Conclusions

The monitoring of activities in traffic scenes often requires being able to detect adverse weather conditions. This allows to improve the safety and the traffic flow control. In this paper two adverse weather conditions were analyzed: snow falling and fog.

The performed analysis is based on two different processes for the two considered weather conditions:

   i.   tuning the learning capabilities of the mixture of Gaussians model for the image sequence,

   ii.   and analyzing the frequency spectrum from the Fourier transform.

The method proved to be robust to image noise. Moreover, processing the sequence at different scale-space frequencies, in space and time, the system is capable of discriminating the weather conditions and to isolate the moving objects at the same time.

Future developments include the analysis of other adverse conditions such as rain and hail. We expect hail to produce the same effect as snow on the images, but we could not make any experimental tests. As for the rain, given that the water is almost transparent to light, a careful choice of both the camera optics and the viewing distance is required to capture enough moving texture associated with the falling water drops. The same process adopted to detect fog can be also extended to detect smoke in the scene, which may be produced by a fire. This can be accomplished by dividing the image into disjoint areas and analyzing each area independently.

Currently the system is implemented in Matlab code and requires between 7 and 15 sec to process each frame (including the computation of the Fast Fourier transform). This is due, not only to the inefficiency of the interpreter, but also to the non-optimal memory allocation to store the frame sequence. By re-coding the entire process with a compiled and optimized implementation a processing time reduction of 1/10 is expected. Moreover improvements in speed will be also obtained by processing down-sampled images.

Nonetheless, the processing time required does not limit the applicability of the approach. In fact, in real environments the weather conditions change rather slowly and within the range of several minutes. Therefore, the time constraints for this application are not the same of vehicle tracking. A timely response can be accounted even

processing the sequence (and providing a detection score) at the rate of 1 frame every 5 or 10 seconds.

5. References

[1] S.S. Beauchemin and J.L. Barron. The Computation of Optical Flow. ACM Computing Surveys (ACMCS1995) 27(3):433-467, 1995.

[2] A. Verri and T. Poggio. Motion field and optical flow: qualitative properties. IEEE Trans. Pattern Anal. Mach. Intell. 11(5):490-498, 1989.

[3] M. Tistarelli. Multiple Constraints to Compute Optical Flow. IEEE Trans. Pattern Anal. Mach. Intell. 18(12):1243-1250, 1996.

[4] M.J. Black and A. Jepson. Estimating Optical Flow in Segmented Images using Variable-order Parametric Models with Local Deformations. IEEE Trans. Pattern Anal. Mach. Intell. 18(10):972-986, 1996.

[5] J.R. Bergen, P. Burt, R. Hingorani, S. Peleg, "Three-frame algorithm for estimating two-component image motion", IEEE Trans. Pattern Anal. Mach. Intell. 14(9):886-896, 1992.

[6] E. De Micheli, V. Torre, S. Uras. The accuracy of the computation of optical flow and of the recovery of motion parameters. IEEE Trans. Pattern Anal. Mach. Intell. 15(5): 434-447, 1993.

[7] C. Fermuller, D. Shulman, and Y. Aloimonos. The Statistics of Optical Flow. Computer Vision and Image Understanding 82:1–32, 2001.

[8] S. McKenna, S. Jabri, Z. Duric, A. Rosenfeld, and H. Wechsler. Tracking groups of people. Comput. Vis. Image Understanding, 80(1): 42–56, 2000.

[9] W. Hu, T. Tan, L. Wang, and S. Maybank. A Survey on Visual Surveillance of Object Motion and Behaviors. IEEE Trans. Systems Man and Cybern. 34(3):334-352, 2004.

[10] C. Stauffer and W. Grimson. Learning Patterns of Activity Using Real-Time Tracking. IEEE Trans. Pattern Anal. Mach. Intell., 22(8):747-757, 2000.

[11] S. G. Narasimhan and S. K. Nayar. Contrast Restoration of Weather Degraded Images. IEEE Transactions on Pattern Analysis and Machine Intelligence, 25(6), June 2003.

[12] S. G. Narasimhan, S. K. Nayar. Shedding Light on the Weather. IEEE Computer Vision and Pattern Recognition (CVPR), Wisconsin, June 2003.

[13] S. K. Nayar, S. G. Narasimhan. Vision in Bad Weather. International Conference on Computer Vision (ICCV), Corfu, Greece, September 1999.

[14] N. Hautiére , J.-P. Tarel, J. Lavenant and D. Aubert. Automatic fog detection and estimation of visibility distance through use of an onboard camera. Machine Vision and Applications, 17(1):8–20, March 2006.

[15] H. Sakaino. Moving vehicle velocity estimation from obscure falling snow scenes based on brightness and contrast model. Proc. of International Conference on Image Processing, 3:905-908, June 2002.