A COMPARITIVE STUDY OF MOG AND KNN FOR FOREGROUND DETECTION

[1]Kritika Singh, [2] Dr. Vishal Gupta,
[1]P G Student, [2]Associate Professor,
[1]Department of Computer Science & Engineering,
[1]Ambedkar Institute of Advanced Communication Technologies and Research, New Delhi, India

Abstract: ***Background Subtraction is a technique that detects foreground from images or videos streams. In Background Subtraction, the foreground elements are categorized after comparing the current frame with the background reference frame. Therefore, the background reference frame should be updated constantly so that it becomes adaptable to the changes in the background of the image or a video. This paper discusses and compares two such Background Subtraction techniques: Mixture of Gaussians (MOG) and K-Nearest Neighbor (KNN). The experimental analysis is done based on some important properties of both the approaches.***

Keywords - Background Subtraction, Mixture of Gaussian (MOG), K-Nearest Neighbor (KNN).

I. INTRODUCTION

There are two approaches of Foreground detection: Derivative algorithms and Background subtraction algorithm, as shown in Figure 1. Background Subtraction technique has been used in the surveillance systems since a long time. This paper mainly focuses on two of the Background Subtraction techniques: Mixture of Gaussian (MOG) and K-Nearest Neighbor (KNN). Background Subtraction algorithms can also be classified on the basis of parameters, as shown in Figure 2.

The simplest case of the parametric Background Subtraction represents pixels in the background model with only one probability density function. However, more than one Gaussian Density Function is used to detect the moving objects present in the background such as tree leaves waving, snowing, different lightening conditions and so on. For that purpose, a Mixture of Gaussians is used. MOG can model multiple background objects for each pixel [1]. Whereas non parametric Background Subtraction algorithms are used to handle arbitrary density functions when a density function is complex and it is difficult to model parametrically. One such approach is K-Nearest Neighbor (KNN). This technique gives good results for the videos that have rapidly changing background.

There are various cases that arise while analyzing videos streams. Some of them are as follows:

Case A: Video captured by still camera, i.e. background is still or changes very gradually.
Case B: Video captured by a moving camera, i.e. both background and foreground object are moving.
Case C: Video in which foreground is still but the background keeps changing.

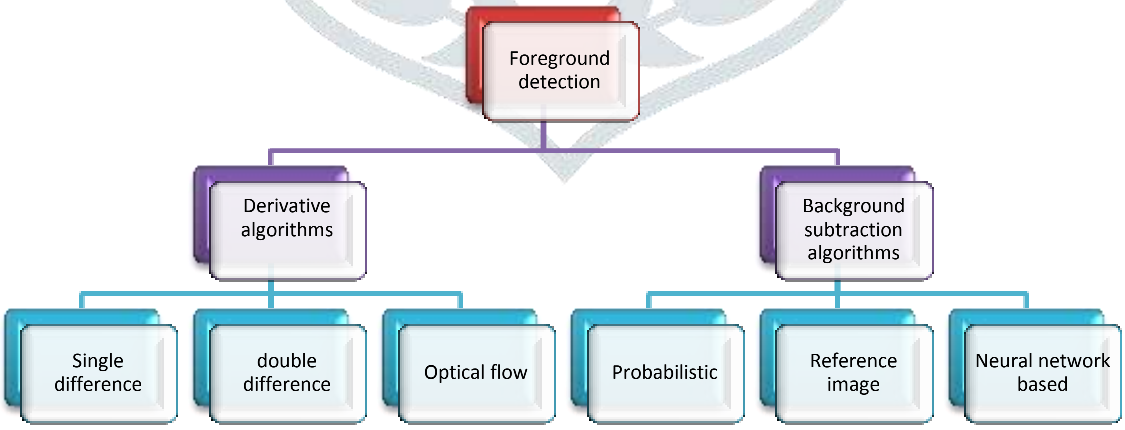

Figure 1 Foreground detection classification [2]

In this paper case A and case B are used to analyze the results of MOG and KNN. For the implementation Open CV 3.0 tool with Python 2.7.14 was used. Also, to remove the noise from the output, four morphological operations are used: Dilation, Erosion, Opening and Closing.

The rest of the paper is organized as follows: section II presents the literature survey. Section III briefly discusses the classification of Background Subtraction algorithm. Section IV summarizes the concept of morphological operations used in this paper and section V presents the results and analysis of both the approaches. Lastly, all the important aspects of the study are highlighted in the conclusion.

II. LITERATURE SURVEY

Many researchers have presented various Background Subtraction methods over time. The concept remains similar for almost every approach, the first or previous frame is used to generate a background reference model. The foreground object is detected by comparing the current frames with the generated background model and then the model is updated. Many background modeling techniques are classified into different categories such: pixel based, region based and hybrid methods. They can also be grouped as parametric and non- parametric approaches. One of the most popular pixel based method, Gaussian distribution model was first introduced by Wren et al. [3] for modeling the background at each pixel. Since it used only single Gaussian function, it was not adaptive to the sudden changes in the background [4]. The Gaussian mixture model was proposed by Stauffer and Grimson [5], [6] for removing the unnecessary disturbance caused by the background changes like tree leaves waving and water waves. In this method every pixel is modeled with the K Gaussian functions rather than a single Gaussian function [7]. An adaptive Gaussian mixture model was introduced by Zivkovic [8], [9] for updating the GMM parameters. In [10] Lee showed that the convergence rate can be improved without making any changes to the stability of GMM. Maddalena et al. [11] proposed a non parametric approach known as self-organisation through the artificial neural networks. Harris et al. [12] compared the two algorithms MOG and KNN on the basis of speed of the output videos. They showed that KNN output gives slower speed than the original speed of the video. In 2000, 9 different algorithms were surveyed by Mc Ivor [13]. This survey shows the description and the comparison between the models. In 2004, Piccardi [14] reviewed 7 methods and categorized them on the basis of the speed, memory and the accuracy. This review gives a very good insight to the readers about complexity of the methods. The most appropriate model can be chosen as per the application requirement. In 2005, Cheung and Kamath [15] show that the various methods can be classified into recursive or non-recursive techniques. Followed by this classification, Elhabian et al. [16] presented a survey in the background modelling based on the recursive and the non-recursive techniques. In 2010, Other surveys and comparisons of different algorithms for background subtraction can be found in the literature [17].

III. BACKGROUND SUBTRACTION CLASSIFICATION

The classification of background subtraction can also be done on the basis of the parameters. They are categorized as:

1. Parametric Background Subtraction
2. Non-Parametric Background Subtraction

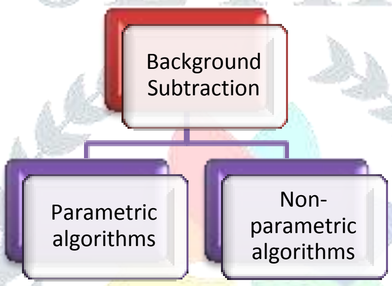

Figure 2 Background Subtraction classification

1. Parametric Background Subtraction

Most common way of modeling the background is by Normal (Gaussian) distribution, based on its pixel's intensity values. In a static scene, the intensity value of a pixel could be monitored and modeled with a Normal distribution. However, this model is challenging because the pixel intensities for moving objects tend to change. The movement from the wind causes this frequently. For instance, for some time a pixel may correspond to sky appearing blue and maybe another time it could correspond to a leaf or a cloud. Hence the pixel intensities keep changing. Mixture of Gaussians is a very popular parametric approach.

1.1 Mixture of Gaussians (MOG)

This technique models the background pixels with a mixture of Gaussians. The parameters and weights of the Gaussians are trained in the training stage and then used in the background subtraction. Then the mixture of Gaussians generates the probability of each pixel [5], [8]. On the basis of its probability, the pixel is categorized as foreground or background.

The background values are stored in the templates. For the detection, these values are compared with the pixels of the current frame. This goes on for all the other templates one by one. Before the process starts, all the pixels are counted as foreground; when the pixel matches any template based on some pre defined threshold it is categorized as foreground or background.

2. Non-parametric Background Subtraction

The Normal distribution model of pixel intensities with no parameters is used in this technique. The pixel intensity probability is measured independently for each frame. This model accurately generates the background model non parametrically. This model is very sensitive to the changes and is more robust so it adapts to the background very easily. The idea is to capture the most recent information of the current frame and update it continuously to make it adaptable to the changes. An example of non-parametric algorithm is KNN.

2.1 K-Nearest Nighbors (KNN)

KNN algorithm stores all available cases and then it classifies the new samples based on the similarity measure. It is a widely used non parametric technique. In this algorithm, all the training samples are needed to be stored before the classification process. This could be a drawback if a very large data set is used.

The training samples are described by n- dimensional numeric attributes. Each sample represents a point in an n- dimensional space. When any unknown sample needs to be classified, k nearest neighbour classifier searches the pattern space for the k training samples that are closest to the unknown sample. Euclidean distance is used to measure the closeness. Euclidean distance between two points, A = (a1, a2,....an) and B = (b1, b2,....bn) is given by the following equation:

$$d(A, B) = \sqrt{\sum_{n=1}^{\infty}(a_i - b_i)^2} \qquad (2.1)$$

IV. MORPHOLOGY

Morphology is a set of operations widely used in the image processing. A structuring element is used as an input image and it creates a same size output image. In a morphological operation, the value of each pixel in the output image is generated by comparing the corresponding pixel in the input image with its neighbors. The morphological operations used in this paper are Dilation, Erosion, Opening and Closing [2].

Dilation- Dilation adds more pixels to the boundaries of objects in an image. The size and shape of the structuring element decides the number of pixels to be added to the boundary. The working of dilation is explained in Figure 3. The foreground object is larger in size as compared to the original foreground object because extra pixels are added.

Erosion- In contrast to Dilation, in Erosion pixels from the object boundaries are removed by using a structuring element. The working of erosion is explained in Figure 4. The foreground object is smaller in size as compared to the original foreground object because the pixels are removed.

Opening- It is Erosion followed by Dilation using the same structuring element. Usually the smaller objects are removed from the foreground. The size of the foreground object is same as the original.

Closing- It is Dilation followed by Erosion. The same structuring element is used for both the operations. This operation can remove small holes present in the foreground object. The size of the foreground object is same as the original.

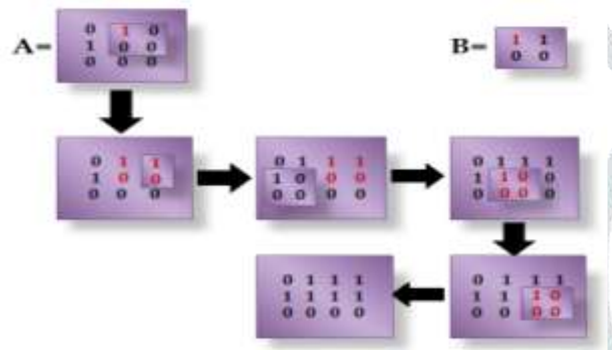

Figure 3 Working of Dilation

As it can be seen from Figure 3, the original 3*3 matrix is expanded to 3*4. Therefore, the result received after applying Dilation is a bigger than the original object.

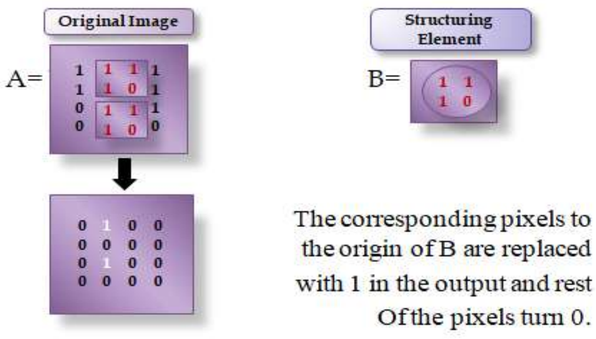

Figure 4 Working of Erosion

As it can be seen from Figure 4, most of the elements in the output matrix are turned to zero. Therefore, the result received after applying Erosion is a smaller than the original object.

V. EXPERIMENTAL RESULTS AND ANALYSIS

In this paper, MOG and KNN are applied on two cases:
Case A: Video captured by still camera, i.e. background is still or changes very gradually.
Case B: Video captured by moving camera, i.e. foreground and background both are moving.

The final analysis of both the algorithms is done based on various factors. Video used for this algorithm was of .mp4 format and coding was done in Python. After the successful retrieval of the foreground, the focus was to remove the noise from the output by using morphological operations like Dilation, Erosion, Opening and Closing.

Case A: Video captured by still camera, i.e. background is still or changes very gradually.

Figure 5 Reference image

Figure 6 MOG result

Figure 7 KNN result

As it can been seen in Figure 6 and Figure 7 that the human beings are detected as the foreground however, some noise is also introduced as the waving of tree leaves. In order to remove that noise, the morphological operations are used. The results after applying four morphological operations are as follows:

Figure 8 Dilation result of MOG

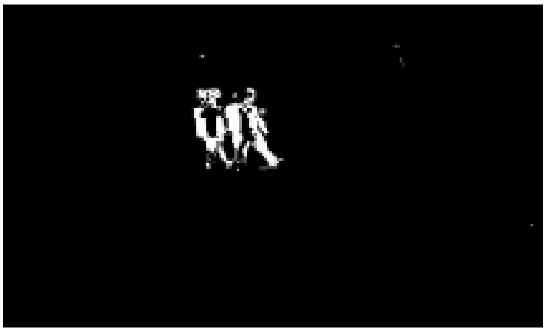

Figure 9 Dilation result of KNN

The resulting foreground after going through Dilation is slightly bigger than the original one. The results of MOG and KNN after applying Dilation are shown in Figure 8 and Figure 9 respectively.

Figure 10 Erosion result of MOG

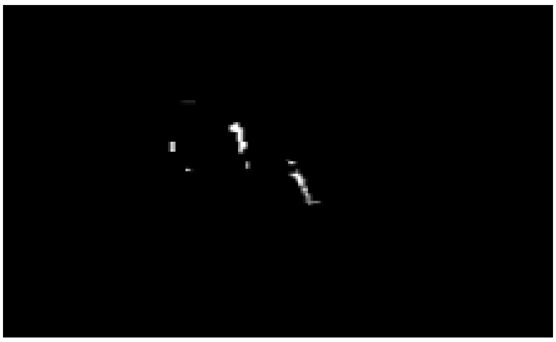

Figure 11 Erosion result of KNN

The resulting foreground after going through Erosion is slightly smaller than the original one. The results of MOG and KNN after applying Erosion are shown in Figure 10 and Figure 11 respectively.

Figure 12 Opening result of MOG

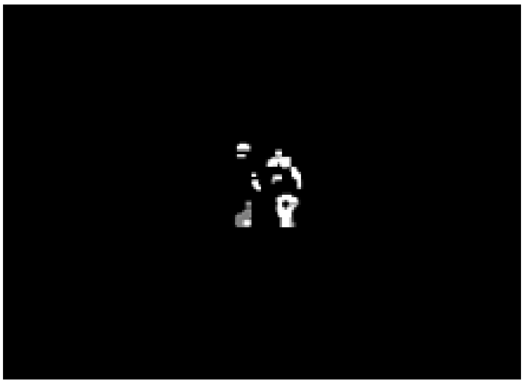

Figure 13 Opening result of KNN

In Opening, first Erosion is performed and then after that with the same structuring element Dilation is performed. The size of the foreground object after applying Opening remains the same. The results of MOG and KNN after applying Opening are shown in Figure 12 and Figure 13 respectively.

Figure 14 Closing result of MOG

Figure 15 Closing result of KNN

Closing is Dilation technique followed by Erosion. In this technique also size of the detected foreground remains the same. The results of MOG and KNN after applying Closing are shown in Figure 14 and Figure 15 respectively.

Case B: Video captured by moving camera, i.e. foreground and background both are moving.

Figure16 Reference image

Figure 17 MOG result

Figure 18 KNN result

The results after applying the morphological operations are shown.

Figure 19 Dilation result of MOG

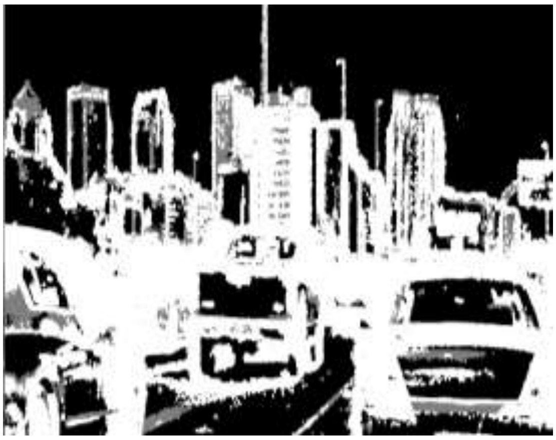

Figure 20 Dilation result of KNN

The results of MOG and KNN after applying Dilation are shown in Figure 19 and Figure 20 respectively.

Figure 21 Erosion result of MOG

Figure 22 Erosion result of KNN

The results of MOG and KNN after applying Erosion are shown in Figure 21 and Figure 22 respectively.

Figure 23 Opening result of MOG

Figure 24 Opening result of KNN

The results of MOG and KNN after applying Opening are shown in Figure 23 and Figure 24 respectively.

Figure 25 Closing result of MOG

Figure 26 Closing result of KNN

The results of MOG and KNN after applying Closing are shown in Figure 25 and Figure 26 respectively.

Observations and Results:
The results of both the algorithms show that:

- The MOG algorithm does not show any visible time lag between the reference mask and the foreground output.
- In MOG algorithm, the details (in the foreground object) are better visible.
- In MOG algorithm, for case A (video taken by still camera), the Erosion operation gives a slightly smaller foreground object than the original object, whereas the Dilation operation gives a slightly bigger foreground object.
- In MOG algorithm, for case A, the Closing operation was able to remove only little noise, also the details in the foreground object were not clearly visible, whereas the Opening operation was successfully able to remove the noise for both the cases. Therefore it gives the best output as compared to other operations.
- Although, the KNN algorithm was successful in removing the noise from the foreground, the output shows a visible time lag. The output video stream was much slower than the original video stream. This is due to the fact that KNN classification uses the instance based nearest neighbors classifiers which are faster training but slower at the classification.
- For case B (video taken by a moving camera), clearly the results of KNN algorithm were better than the MOG algorithm. Hence, it is proved by the output that the MOG is very sensitive to the rapidly changing background due to which a lot of noise is seen in the foreground object.

The comparison of MOG and KNN is shown in Table 1.

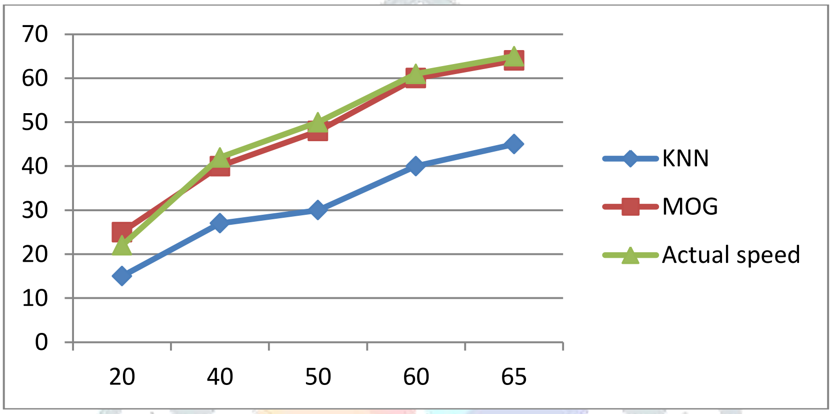

Figure 27 Output speed of the videos after applying MOG and KNN

Figure 27 the graph shows different speed of the videos after applying MOG and KNN. It is clear from the graph that the KNN gives a much slower speed as compared to the MOG.

Table 1 Comparison of Mixture of Gaussian (MOG) and K-Nearest Neighbors (KNN)

	MOG	KNN
Time lag	It does not have any time lag in the output.	The output video is slower than the original. This is due to the fact that KNN classification uses the instance based nearest neighbors classifiers which are faster training but slower at the classification.
Details in the foreground object before applying morphological operations	The details in the foreground object like hands and legs of human beings are easily visible.	The details are slightly blurred as compared to MOG but they are still distinguishable.
Erosion	Noise is removed	Noise is removed
Dilation	For case A Noise is reduced but some of it is still visible.	Gives better results.

Size of the foreground object	• Smaller after applying Erosion. • Bigger after applying Dilation. • Same after applying opening and closing.	• Smaller after applying Erosion. • Bigger after applying Dilation. • Same after applying opening and closing.
Foreground object	The details are hardly visible after applying dilation because it adds extra pixels to fill in any holes in the foreground object.	The details were much clearly visible as compared to MOG- dilation results.
Video taken by a moving camera	MOG does not work best when background keeps rapidly changing.	KNN gives good results even when the background was changing continuously.

VI. Conclusion

This paper briefly discusses two Background Subtraction algorithms: MOG and KNN. Both the algorithms were applied on two different cases and the results were analyzed on the basis of time lag, foreground object size and noise. To remove the noise, four morphological operations were applied on all the outputs. It can be seen from the output that although, MOG gives good results for case A (video captured by a still camera), it is very sensitive to the background changes. Hence a lot of noise can be seen in the output of case B, where the background keeps changing rapidly. For case A, after applying MOG, the details in the foreground object such as hands and legs were more clearly visible than after applying KNN. The output shows that KNN algorithm works better for case B and MOG algorithm works better for case A. The drawback of KNN is that it provides the slower output than the original videos streams.

VII. Acknowledgment

I would like to show my gratitude to Dr. Vishal Gupta (Associate Professor, Ambedkar Institute of Advanced Communication Technologies and Research, New Delhi, India) who provided insight and expertise that greatly assisted this paper. We would like to thank everyone, who has supported in creating and reviewing this paper.

References

[1] W. Power and J. A. Schoonees, “Understanding background mixture models for foreground segmentation,” in Proc. IVCNZ, Nov. 2002, pp. 267–271.

[2] Nawaz, Muhammad, et al. "Precise foreground detection algorithm using motion estimation, minima and maxima inside the foreground object." IEEE Transactions on broadcasting59.4 (2013): 725-731.

[3] C. Wren, A. Azarhayejani, T. Darrell, A.P. Pentland IEEE Trans. Pattern Anal. Mach Intell., 19 (7) (1997), pp. 780-785.

[4] M. Piccardi IEEE Int. Conf. Syst. Man Cybern, 4 (Oct. 2004), pp. 3099-3104 The Hague, The Netherlands.

[5] C. Stauffer, E. Grimson, “Adaptive background mixture models for realtime tracking”, in: Proc. IEEE Int. Conf. Comput. Vis. Pattern Recognit., vol. 2, 1999, pp. 246e252.

[6] C. Stauffer, E. Grimson, IEEE Trans. Pattern Anal. Mach. Intell. 22 (8)(Aug. 2000) 747e757.

[7] Benezeth, Yannick, et al. "Review and evaluation of commonly-implemented Background Subtraction Algorithms." Pattern Recognition, 2008. ICPR 2008. 19th International Conference on. IEEE, 2008.

[8] Z. Zivkovic, “Improved adaptive Gaussian mixture model for Background Subtraction”, in: Proc. 17th Int. Conf. Pattern Recognit., vol. 2, Aug. 2004, pp. 28e31.

[9] Z. Zivkovic, F. van der Heijden, Pattern Recognit. Lett. 27 (May 2006) 773e780.

[10] D.S. Lee, Pattern Anal Mach. Intell. IEEE Trans. 27 (5) (May 2005) 827e832.

[11] L. Maddalena, A. Petrosino, IEEE Trans. Image Process. 17 (7) (2008) 1168e1177.

[12] Simaremare, Harris, Hadi Melwanda, and Abdillah Abdillah. "The Comparison Of Vehicle Speed Accuracy Using Video Based Mixture Of Gaussian 2 Method and K-Nearest Neighbor Method." Seminar Nasional Teknologi Informasi Komunikasi dan Industri. 2017.

[13] Bouwmans, Thierry. "Traditional and recent approaches in background modeling for foreground detection: An overview." Computer Science Review 11 (2014): 31-66.

[14] Piccardi, Massimo. "Background Subtraction Techniques: a review." Systems, man and cybernetics, 2004 IEEE international conference on. Vol. 4. IEEE, 2004.

[15] Sen-Ching, S. Cheung, and Chandrika Kamath. "Robust techniques for Background Subtraction in urban traffic video." Visual Communications and Image Processing 2004. Vol. 5308. International Society for Optics and Photonics, 2004.

[16] Elhabian, Shireen Y., Khaled M. El-Sayed, and Sumaya H. Ahmed. "Moving object detection in spatial domain using background removal techniques-state-of-art." Recent patents on computer science 1.1 (2008): 32-54.

[17] Chalidabhongse, Thanarat Horprasert, et al. "A perturbation method for evaluating Background Subtraction algorithms." Joint IEEE International Workshop on Visual Surveillance and Performance Evaluation of Tracking and Surveillance. 2003.