

Integration of Spatio-Temporal Information for Motion Detection by Means of Fuzzy Reasoning

M. Barni, F. Bartolini, V. Cappellini, F. Lambardi

Dipartimento di Ingegneria Elettronica, Università di Firenze,
Via S. Marta 3, 50139, Firenze
e-mail barto@cosimo.die.unifi.it

Abstract. In this paper a motion detection system based on fuzzy reasoning is presented. Each pixel of a frame of the sequence is attempted to be classified as belonging to one of four classes (moving, still, uncovered background, covered background). The classes are treated as fuzzy sets, and as such, they are characterized by membership functions. After an initialization step, the degree of membership of each pixel to each class is refined by the application of a reasoning module driven by a set of fuzzy rules. Such fuzzy rules are designed so that the spatio-temporal correlation of image sequences is exploited by integrating information extracted from a small spatio-temporal neighborhood. The proposed system results to be flexible, thanks to the use of the reasoning approach (rules can be easily changed), and robust, thanks to the use of fuzzy logic.

1 Introduction

Motion detection techniques play an important role in the field of image sequence processing. Sensing the objects of the scene that are moving is useful, for example, to reduce the area of the images to be compressed in video coding systems, to produce alarms when an intrusion is occurring in restricted areas, or to count vehicles for traffic control applications.

Motion detection techniques are aimed to detect if (and possibly where) something is moving in the imaged scene.

When the goal is simply the surveillance of restricted areas it is not usually needed to detect where moving objects are localized but only whether they are present or not, this allows to deal with image noise by integrating the information obtained over the whole image [2,5].

For more complex tasks, as for example traffic monitoring, it is, on the contrary, required to segment the image in moving objects and static areas. For this goal two classes of algorithms can be identified. The algorithms of the first class analyze the differences between the current image and a reference static background. In order to deal with global illumination changes, the static background has to be adaptively updated [6] or edge images (less affected by the variation of the ambient lighting) have to be used [7]; nevertheless, the sensitivity of the algorithms of this class to sudden illumination changes remains a major drawback.

The algorithms of the second class, on the other side, analyze the differences between couples of successive frames of the sequence. For changing pixels detection, likelihood ratio tests are usually used [8]. Furthermore to enhance robustness it is important to exploit the a priori knowledge about motion spatio-temporal coherence, in that spatial and temporal neighboring pixels usually move in a similar way. Spatio-temporal integration of motion information can be, for example, achieved by modeling the image sequence as a 3D gaussian markov random field and by estimating the segmentation with a maximum a posteriori criterion [9]. This method is quite robust, but it is difficult to be driven by a priori heuristic knowledge about the image sequence. Furthermore, it does not allow to distinguish covered and uncovered areas from moving objects. Another approach for integrating spatio-temporal information is based on a set of heuristic rules that allow to assign each image pixel to a class identifying the behavior of the corresponding point in the 3D real world [10].

In the system described in this paper integration is achieved by imposing some heuristic rules to hold in a spatio-temporal neighborhood of each pixel. The adopted rules allow to label each pixel as belonging to a moving object (class M), to an area of covered (class C) or uncovered (class U) background, or to the still background (class S). The use of 4 classes instead of 2 only (still/moving) permits to better shape the mask of moving objects given that only class M and C pixels are assigned to it. We also propose to cast the heuristic rules in the framework of fuzzy reasoning. Fuzzy reasoning is, in fact, very suitable for modeling human knowledge about real world and also for producing very robust algorithms [4].

2 Fuzzy reasoning

Fuzzy or approximate reasoning is the core of any fuzzy system, since it is in charge of deriving actions or taking decisions by starting from imprecise or vague data. According to the original approach proposed by Zadeh [11], and further developed in [12] and [13], the fuzzy reasoning module is primarily characterized by a set of fuzzy rules, by means of which raw data are processed and decisions are taken. In the case of a multi-input-single-output (MISO) system, fuzzy rules have the form

R_1 : if x_1 is A_{11} and ... x_n is A_{1n} then y is B_1

...

R_m : if x_1 is A_{m1} and ... x_n is A_{mn} then y is B_m

where x_i are the input linguistic variables, y is the output linguistic variable and A_j , B_j are fuzzy sets. The problem of finding the value of the output variable y (i.e. the fuzzy set associated to it) by starting from the values assumed by the variables x_i is called the *fuzzy inference* or *fuzzy reasoning process*. Such a process involves three main steps:

1. determination of the matching between the variables x_i and the sets A_{ji} ;

2. computation of the output of each single rule R_j ;
3. aggregation of the outputs of the single rules.

A comprehensive survey of the possible approaches that can be followed to perform each of the above steps is outside the scope of this note, in the following only the strategy used throughout the motion detection system will be described. Interested readers may find more details in [12,13,4].

Let λ_j be the degree of activation of the j -th rule. To evaluate λ_j 's, let us consider first the quantities

$$\rho_{ji} = \max_{u \in U} \{\min\{\mu_{x_i}(u), A_j(u)\}\}, \quad (1)$$

where $\mu_{x_i}(u)$ is the membership function of the fuzzy set assumed by the linguistic variable x_i , and ρ_j represents the degree of matching between the actual value of the variable x_i and the fuzzy set A_j . Once ρ_j have been calculated the activation degrees λ_j are computed as

$$\lambda_j = \min_{i=1,n} \{\rho_{ji}\}. \quad (2)$$

Steps 2 and 3 of the fuzzy inference process, i.e. computation of the output of each rule and combination of such outputs, is accomplished by means of the formula

$$\mu_y(u) = \max_{j=1,m} \{\min\{\lambda_j, \mu_{B_j}(u)\}\}. \quad (3)$$

It should be noted that being y a linguistic variable, the output of the first three steps of the inference process is a fuzzy set, whose memberships function is defined by means of equations 1 through 3. The defuzzification of y can be carried out in many ways [13]. In this work the following simple rule is applied to produce a crisp version Y of the linguistic variable y

$$Y = \max_{u \in U} \{\mu_y(u)\}. \quad (4)$$

A generalization of the scheme of fuzzy system given above comprises the possibility of including an else statement [11,13].

3 Overview of the System

As outlined in Sections 1 and 2, the system aims at classifying pixels as being in one of 4 possible states: moving (M), still (S), covering (C) and uncovering (U). In particular, for each frame, pixels are assigned a degree of membership to one of the above 4 classes.

As a first step pixels classification is initialized by relying on frame differences. Then a former set of fuzzy rules is applied in order to refine the rough classification produced by the initializer. The status of a pixel in position (i, j) is determined by reasoning upon the states of the pixels lying in a 3×3 window

centered in (i, j) and the states of the pixels in position (i, j) in the previous and the subsequent frame.

In order to understand how the fuzzy reasoning module operates, some definitions must be given first. Six linguistic variables are taken into account by the system: 2 of them, x_{-1} and x_{+1} , refer to the status of pixels (i, j) in the previous and subsequent frame respectively, whereas the other 4 are related to the global status of the pixels inside the 3×3 working window in the current frame. Let us call the variables belonging to this second group x_m , x_s , x_c and x_u .

Fuzzy sets assigned to x_{-1} and x_{+1} are defined in a space U' consisting of 4 elements: let us call these elements m (moving), s (still), c (covering) and u (uncovering). Such fuzzy sets are built by specifying for each element in U' the possibility that the previous/subsequent pixel is a moving, still, covering or uncovering pixel (in Table 1 the fuzzy sets used to define the rules the reasoning module consists of are depicted).

Fuzzy set	Elements of U'			
	m	s	c	u
moving	1	0	0.1	0.1
covering	0.2	0	1	0.1
uncovering	0	0.2	0.1	1
still	0	1	0.1	0.1

Table 1. Values of the membership functions of the four fuzzy sets.

With regard to x_m , x_s , x_c and x_u , they account for the global status of the pixels in the current frame. To build the space U_M fuzzy sets referring to x_m belong to, the degrees of membership to the class M of pixels inside the working window are summed. Let us call such a sum M -sum; the possible values that the M -sum can assume form the U_M space. A fuzzy set defined in U_M gives the possibility, for each value in U_M , that the M -sum takes that value. Let us consider for example the statement *most of the pixels in the working window are moving (belong to class M)*; such a statement can be cast in the framework of approximate reasoning by introducing a fuzzy set whose shape is illustrated in Figure 1, the fuzzy set meaning that for a window in which *most* of the pixels are known to belong to the M -class the possibility that the M -sum is, for example, 2.5 equals 0.4. Of course, similar considerations hold for the linguistic variables x_s , x_c and x_u .

By means of fuzzy inference, the inputs to the reasoning modules are processed and pixel classification significantly improved. Following the first reasoning module a second step is carried out to further refine pixel classification. For this second step a different set of rules is applied since different configurations are likely to occur with regard to the output of the initializer.

The fuzzy reasoning module produces four memberships with the possibility that the pixels belong to the M , S , C and U classes. At this point, defuzzification

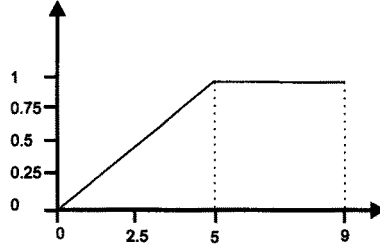


Fig. 1. The function describing the *most of the pixels* fuzzy set.

is performed to definitely classify pixels: more specifically, each pixel is assigned to the class it exhibits the maximum membership to.

4 Initialization and set of rules

With regard to the initialization of memberships, a procedure derived from [9] is adopted. A likelihood-ratio-like parameter L is evaluated inside a 5×5 window W centered at the current pixel position:

$$L = \frac{1}{n^2} \sum_{(i,j) \in W} \delta(i,j)^2 + \frac{1}{\sum_{(i,j) \in W} \Delta x^2(i,j)} \left\{ \left[\sum_{(i,j) \in W} \Delta x(i,j) \delta(i,j)^2 \right]^2 + \left[\sum_{(i,j) \in W} \Delta y(i,j) \delta(i,j)^2 \right]^2 \right\} \quad (5)$$

where $\delta(i,j)$ is the frame difference computed for pixels in position (i,j) and $\Delta x(i,j)$ and $\Delta y(i,j)$ are the horizontal and vertical distances of pixel in position (i,j) from the center of the window. The value of this parameter is used as input to the 4 memberships initialization functions (see Figure 2).

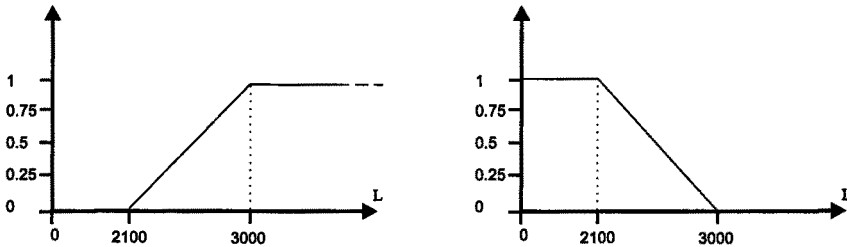


Fig. 2. Initialization functions used for the Moving, Covering and Uncovering fuzzy sets (*Left*) and for the Still fuzzy set (*right*).

In this section the complete list of fuzzy rules the system consists of is given. Let us begin with the first reasoning step, the following rules are applied to the output of the initializer:

- R_1 : if x_p was *moving* and
 x_s is *still* and
at least some pixels in the current frame are *uncovering*,
 then the current pixel is *certainly uncovering*;
- R_2 : if x_p was *moving* or *covering* and
 x_s is *moving* or *uncovering* and
at least some pixels in the current frame are *moving*,
 then the current pixel is *certainly moving*;
- R_3 : if x_p was *still* and
 x_s is *moving* and
at least some pixels in the current frame are *covering*,
 then the current pixel is *certainly covering*;
- R_4 : if x_p was *still* or *uncovering* and
 x_s is *covering* or *still* and
at least some pixels in the current frame are *still*,
 then the current pixel is *certainly still*;
- else
- G_1 : if *most* of the pixels in the current frame are *still*
 then the current pixel is *certainly still*;
- G_2 : if *most* of the pixels in the current frame are *moving*
 then the current pixel is *certainly moving*;
- else
- H_1 : if *most* of the pixels in the current frame are *covering*
 then the current pixel is *certainly covering*;
- H_2 : if *most* of the pixels in the current frame are *uncovering*
 then the current pixel is *certainly uncovering*;

The above rules have been designed by taking into account the situations that are most likely to occur in the analysis of image sequences; when an unusual pixel configuration is encountered which is not encompassed by these rules, the fuzzy engine produces a vague output to be refined during subsequent steps. The second set of rules is very similar to the first one, the only difference being that rules H_1 and H_2 are at the same level as rules G_1 and G_2 . The rationale for such a minor modification is that, due to the particular initialization adopted, the first module is likely to reason upon images in which the degree of membership to the M -class is equal to those of the C and U -classes.

5 Results and Conclusions

The proposed system has been tested on synthetic and real world images. The results obtained by processing a sequence representing vehicles on a highway lane are here described.

In Figure 3 a frame of the processed sequence is shown: in the field of view both the lanes are included. In Figure 4 the values computed by the algorithm for the four membership functions at each pixel location are visualized (the images labeled M, S, C and U display, respectively, the Moving, the Still, the Covering and the Uncovering membership functions). In Figure 5 the results of



Fig. 3. A frame of the sequence representing vehicles on a highway.

the defuzzification process are shown. In each map the pixels belonging to the respective class are highlighted. From the analysis of the maps it is evident that the algorithm allows the estimation of the direction of motion, even if it does not estimate speed. For vehicles approaching the camera, covered areas are, in fact, behind and uncovered are on the back. The converse it is true for vehicles moving in the opposite direction. Furthermore the classification into four classes instead of only two (i.e. moving/still) helps in managing ambiguous situation and increases, then, the robustness of the algorithm.

The presented fuzzy reasoning approach to motion detection allows to combine the flexibility of the systems based on reasoning algorithms and the robustness which is characteristic of fuzzy methodologies.

6 Acknowledgment

This work was partially funded by the *Progetto Finalizzato Trasporti II* of CNR (Italian National Research Council) under grant nr. 96.00027.PF74 .

References

1. S.B. Chae, J.S. Kim and R.H. Park "Video coding by segmenting motion vectors and frame differences," vol. 32, no. 4, Apr. 1993, pp. 870-876.
2. I. Distein, "A new technique for visual motion alarm," *Pattern Recognition Letter*, vol. 8, no. 5, Dec. 1988, pp. 347-351.
3. P.G. Michalopoulos, "Vehicle detection video through image processing: the Autoscope system," *IEEE Transactions on Vehicular Technology*, vol. 40, no. 1, Feb. 1991, pp. 21-28.

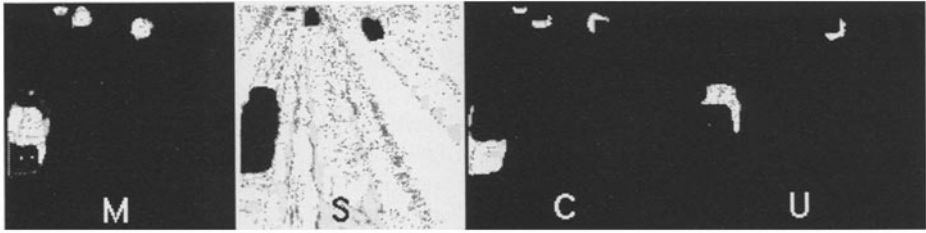


Fig. 4. Representation of the values computed by the system for the four membership functions at each pixel position (values are normalized in the range 0-255).



Fig. 5. Maps of the pixels belonging to the four classes (Moving, Still, Covering, Uncovering).

4. T. Terano, K. Asai and M. Sugeno, *Fuzzy Systems Theory and its Applications*, Academic Press Inc. 1992.
5. P. Iliev and L. Tsekov, "Motion detection using image histogram sequence analysis," *Signal Processing*, vol. 30, no. 3, Feb. 1993, pp. 373-384.
6. K.P. Karmann, A. Von Brandt and R. Gerl, "Moving object segmentation based on adaptive reference images," *Proc. EUSIPCO-90.*, Barcelon, Sept. 18-21, 1990.
7. M.A. Vicencio, B. Qiu and M.G. Hartley, "Algorithms and architectures for the analysis of road-traffic movements," *Proc. 3rd Int. Conf. Image Proc. and Appl.*, Warwick, UK, July 1989.
8. R. Jain and H.H. Nagel, "On the analysis of accumulative difference pictures from image sequences of real world scenes," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 1, no. 2, Apr. 1979, pp. 206-214.
9. P. Bouthemy and P. Lalande, "Recovery of moving object masks in an image sequence using local spatio-temporal contextual information," *Optical Engineering*, vol. 32, no. 6, June 1993, pp. 1205-1212.
10. R. Jain, "Extraction of motion information from peripheral processes," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 3, no. 5, Sept. 1981, pp. 489-503.
11. L. A. Zadeh, "Outline of a new approach to the analysis of complex systems and decision processes", *IEEE Trans. Syst. Man Cybern.*, vol.SMC-3, pp. 28-44, 1973.
12. D. Dubois and H. Prade, *Possibility Theory: an Approach to Computerized Processing of Uncertainty*, New York, Plenum Press, 1988.
13. C. C. Lee, "Fuzzy logic in control systems: fuzzy logic controller - Part I and II", *IEEE Trans. Syst. Man Cybern.*, vol.SMC-20, no.2, pp. 404-435, Mar./Apr. 1990.