Looming Motion Segmentation in Vehicle Tracking System using Wavelet Transforms

K. SUBRAMANIAM, S. SHUKLA, S.S. DLAY and F.C. RIND
Department of Electrical and Electronic Engineering
University of Newcastle-Upon-Tyne
Merz Court, Newcastle-Upon-Tyne NE1 7RU
UNITED KINGDOM

Abstract:-*:* This paper presents an algorithm to solve the problem of looming motion segmentation as a part of a vehicle tracking system [1]. The basic algorithm, uses the Gabor and Mallat Wavelet transforms to produce motion segmentation objects as an input to the present algorithm. When an object is looming in a scene, it is growing in all directions and treated as different objects by the motion segmentation algorithm[1]. The edges of these different parts of the object are connected locally in the scene. By applying the idea of edge connectivity, based on mathematical morphology[5], the segmented parts of the object can be connected together. This improved algorithm can detect and produce segmented results for both looming and translational motion. Applying Boolean algebra for connectivity decision, makes the implementation of the algorithm simple and increases the speed of computation. It also copes well with multiple objects looming in the scene. Reliability of the connected objects is increased by cross correlating the Mallat information (edge detection) and final result (connected segmented objects). The computational time is reduced further since the images are binary and are processed using morphological operators.

Key-Words: - Wavelet Transform, Looming motion segmentation , Morphology, Boolean algebra, Cross-correlation.

1. Introduction

The algorithm presented here aims to solve the looming motion segmentation problem associated with a larger system for vehicle detection and tracking. The basic system [1] was developed making use of Wavelet transforms. The two types of filters used from the wavelet family are the Gabor and the Mallat, for image flow field estimation and edge detection respectively.

Based on the Gabor wavelet transform[2], the image flow field between two consecutive frames is computed. Since the Gabor Wavelets are large, this method does not have a serious aperture or correspondence problem (which pixel in one frame corresponds to the moved pixel in next frame) problem, but the spatial resolution of the computed flow field is low.

Motion hypotheses from the image flow field can be extracted by detecting local maxima in the image flow histogram, which is formed over the flow field vectors. The advantage of the motion hypotheses is that it drastically reduce the correspondence problem for an image flow estimation algorithm based on more localised features, such as edges.

Edge detection is achieved using the Mallat Wavelet transform[3], which can be thought of as the grey-value gradient at different levels of resolution. Edge detection is achieved by convolution of the two-dimensional function that describes the local grey level of an image with a Mallat filter.

Integration over the sequence of frames is performed to determine which motion hypothesis in one pair of frames corresponds to that in next pair of frames. In order to avoid assumptions, about the type of movement objects may perform, we associate the motion hypothesis via the spatial overlap in the intermediate frame, i.e. the one that is common to both pairs.

At the end, segmentation is achieved, describing the task of dividing the data into different groups. These groups are analysed by processing the information separately. Each edge pixel is classified as belonging to that motion hypothesis for which it has the highest accordance value.

In motion segmentation algorithm [1], the assumption is made that objects move translationally. The algorithm behaves reasonably well even if the assumption of the motion algorithm is strongly violated.

The algorithm for the basic system fails to detect an object as a single object when looming or approaching directly in the scene. Instead, it shows the different parts of a single looming object as different objects in motion, which obviously belongs to one object. Mathematical morphology [5] used to implement the present algorithm is a tool for extracting image components that are useful in the representation and description of region shapes, such as image boundaries. This helps to reduce the information contained in an image and the connection of the segmented objects can be achieved by looking at the edges only. As these operators deal with the binary images, computational time is reduced. The process of cross-correlation further improves the accuracy of the object detection.

2. Investigation

The original algorithm successfully segments objects in translational motion. A problem arises when an object in the scene approaches directly, or looms. Regions of the object grows in different directions in space at the same time. The main algorithm detects the movement of the pixels in one direction and groups them as one object. Therefore, the segmented objects for looming motion appears as a different part of a single object as shown in figure 1. Frame1 and Frame2 is a computer generated looming square on a plain background. The images have a resolution of 128x128 pixels with 256 grey values. The segmentation produced by the original algorithm shows the looming object segmented into several part. The idea of edge connectivity arises from the fact that these different objects are locally connected in the scene.

3. Looming Motion Segmentation

The approach of the solution was made using the idea of edge linking based on mathematical morphology [5] and Boolean algebra [6] between the segmented results. This idea considers a line to be a chain of nodes which are connected with each other as long as the lines are locally connected. The cross-correlation is used to make sure that the final connected object is identical to the original.

3.1 Motion Computation

The input image sequence for the basic algorithm is subdivided into four sub regions as shown in figure 2. Each frame is sub divided into four parts to zoom in particular sectors of the scene and try to collect the information that might not be very clear on a larger scale. In figure 2, frame 1 and frame 2 are the looming square with the textured background which are subdivided into four sub-regions as A,B,C and D. The sub-region A1 from frame 1 and sub region A2 from frame 2 are used to compute the image flow field [4] using wavelet transforms [1] to produce segmented objects for sub region A shown in figure 3. The same can be done for the rest of the sub-regions using the basic system algorithm.

3.2 Edge Enhancement

The binary image processing is built upon three main operations. Their effect on the structural geometric image content makes them vital for morphological image processing [5]. The *erosion*, *dilation* and the *opening* are the morphological operators [6] used in the algorithm.

The *erosion* and *dilation* is performed on the image to enhance the edges of the motion objects. The *opening* filter is applied on the enhanced image to remove the noise.

The *erosion* of binary image *A* by *B* is denoted by $A \ \theta \ B$. *B* is called a *structuring element* (kernel). When the origin lies in the structuring element, erosion has a shrinking effect, $A \ominus B \subset A$.

Dilation of set *A* by structuring element *B* is defined by ;

$$\mathrm{A} \oplus \mathrm{B} = \left(\mathrm{A}^{c} \theta (-\mathrm{B})\right)^{c}, \qquad (5)$$

where $-B = \{-b : b \in B\}$ is the *reflection* of *B* through the origin and *c* denotes set complement. Geometrically, dilation can be obtained by translating the structuring element to each point of the image and then forming the union of all the translates:

$$A \oplus B = \bigcup_{x \in A} B_x \ . \qquad (6)$$

If the origin lies in the structuring element, then dilation expands the image, i.e. $A \oplus B \supset A$ which means that A is a subset of $A \oplus B$. Dilation provides a useful edge enhancement algorithm. The boundary of a binary image *A* is defined by;

$$\partial(A) = (A \oplus B) - A \ , \qquad (7)$$

where *B* is the 3×3 square centered at the origin. According to this equation (7), dilation adjoins pixels to the outside of the input image, and subtraction of the input leaves this outside edge. The output result is shown in the figure 4.

An often-used morphological filter is *opening*, which for set *A* and structuring element *B* is defined by;

$$A \circ B = \bigcup_{B_x \subset A} B_x \ . \qquad (8)$$

$A \circ B$ is the union of all translates of B that are subset of *A*: slide *B* around inside *A* and take as the filter output all pixels covered by the sliding structuring element. If the observed image is an ideal image unioned with noise, then the opening can serve as a restoration filter by choosing a structuring element that fits inside the ideal image but not inside the noise. An appropriate opening will remove (most of) the noise while keeping most of the information. The enhanced objects are labelled for the next stage processing.

3.3 Segmentation

The second stage of the algorithm, determines the connectivity between the labelled objects using Boolean Logic [6]. Boolean logic (also called binary logic) is a framework originally proposed to reason about propositions that are either true or false. This is the foundation for most artificial intelligence programs.

Connected regions of an object will have the same labels so if there are more than one object in motion in one sub region, it can be detected as well. This also solves the problem of connecting different objects in each sub-region as the motion object shown in figure 5.

The implementation of Boolean expression for the four sub-region of the frame i.e. A, B, C and D,

$$\begin{aligned}&(AB \bullet CD \bullet (BC + BD)) + \\ &(AD \bullet ((BC \bullet CD) + (BD \bullet AC))) = if.True\end{aligned} \qquad (9)$$

gives the output of the connected objects as single moving object.

At First, segmented objects are checked for connectivity between the objects of the sub-region A. If segmented object (edges) are connected to another region of the object, it then belong to the same object. Otherwise, it belongs to the other object of interest.

The other sub-regions are processed, in the similar manner which will group the motion object of each sub-region.

Next, using the similar method above, to find the connectivity between the sub-regions. The output of this stage, will group the actual motion of a single object as shown in figure 6.

3.4 Cross-Correlation

At the end of the algorithm, there is a need to ensure that the connected objects are the same as the original motion objects. The method of cross correlation between the Mallat wavelet information and the segmented result is computed for the final segmentation decision.

Cross correlation between the two unknown images can be computed to find the closest match by searching for the largest amplitude in the functions representing the images. The cross correlation $r_{12}(n)$ between the two image functions $x_1(n)$ and $x_2(n)$ representing the Mallat information and segmented result respectively where each containing N data sequence can be written as

$$r_{12} = \frac{1}{N}\sum_{n=0}^{N-1} x_1(n)x_2(n+x) \qquad (10)$$

4. Examples

At present, the most severe limitation of the basic tracking system is the restriction to translational motion. The looming motion segmentation system described in this paper, generates connected motion objects while preserving local segmentation decisions. The strength of the system is where it analyses the edges on the binary images for connectivity, which in return increases the computational accuracy and improves the speed of the algorithm.

Figure 1, shows the limitations of the basic algorithm. In the example shown, two consecutive frames of a computer generated square images are taken in account. The basic algorithm shows the final result as four different objects in motion. Some parts of the segmented regions are common to others as shown in Figures c) and d), and, e) and f). This gives rise to the basic idea of joining the objects using the edge linking procedure.

In figure 7, frame 14 is the car looming sequence on the road taken from a moving camera. The images have a resolution of 128x128 pixels with 256 grey values. The car shows the looming motion while it comes closer in the scene. The original frame is divided into four sub-regions. These sub-regions are used to compute the image flow field between the consecutive frames. Each sub-frame produces the segmented results, which are analysed for connectivity using the algorithm and the result is the connected objects as shown in figure7 b), c), d) and e). In the final process, these sub-figures are checked for connectivity for segmentation decision as shown in figure 7 f). From the final result, it is noticeable that the shadow of the car has been connected with the car because of the illumination effect in the scene. Although, the basic algorithm [1] coped with these image sequence initially, it fails when the car is looming and comes much closer in the scene.

The example shown above was tested using the real image sequence and the results obtained shows the success of the algorithm in detecting the looming objects.

5. Conclusion

The algorithm presented above tackles the problem of looming motion segmentation as related to the vehicle detection and tracking system. The algorithm makes use of an edge linking process between the sub-regions of consecutive single frames. The result could take the form of a graph, whose nodes represent the edges and whose vertices represent whether edges are connected or not. Thus, the separate parts of a single motion objects are connected as a square and locally segregated in the scene at the same time. Simple implementation of detecting the connection between edges using Boolean algebra increases the speed of the algorithm.
The original frame is divided into four parts and the motion is calculated in each subpart of the consecutive frames. It is possible to divide the frames in more subparts but this would slow the speed of computation. Also, as the object grows in all the directions (looming), there will be only four major quadrants that will show the motion in one direction with very little phase difference. Finally, zooming into the sub-regions further more will not add any more information to the result. Use of logical operators for the edge linking is performed by labelling the sub regions. This technique is easy to implement in software.

The algorithm was tested first using the basic computer generated images and then the real time image sequences. The result obtained shows that the algorithm works well with several objects and extends the potential of the main algorithm to track the approaching objects.

Reference

[1] Subramaniam K., Dlay S.S. and Rind F.C, July 1999, ‘Vehicle detection and tracking using wavelet transforms’, IEEE Proceeding of CSCC'99, 5011-5016.

[2] Daugman J.G, July 1988, ‘Complete discrete 2-D Gabor transform by neural networks for image analysis and compression’, IEEE Transactions on Acoustics, Speech and Signal Processing, 36-7, 1169-1179.

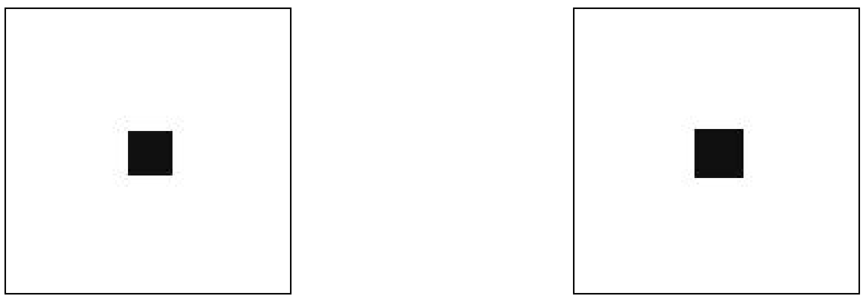
[3] Mallat S and Zhong S, 1992, ‘Characterisation of signal from multiscale edges’, IEEE Transactions on Pattern Analysis and Machine Intelligence, 14-7, 710-732.

[4] Horn B.K.P and Schunck B.G, 1981, ‘Determining optical flow’, Artificial Intelligence, 17, 185-203.

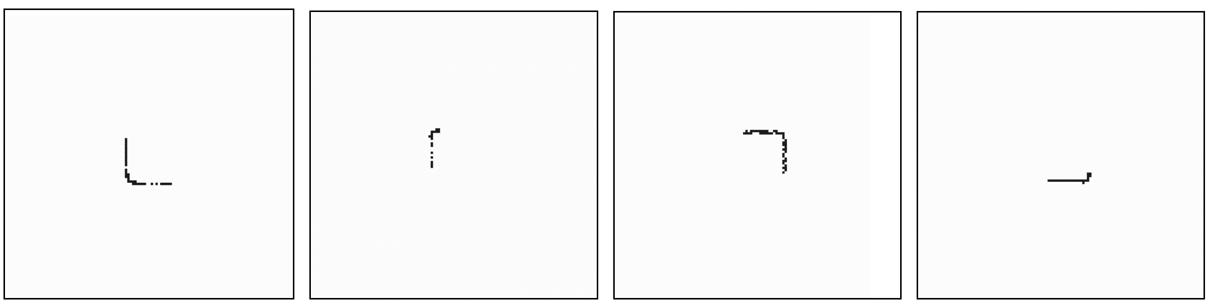
[5] Dougherty E.R., ‘An Introduction to Morphological Image Processing, SPIE Press, 1992.

[6] Sinha D. and Dougherty E.R., ‘Introduction to computer-based imaging systems’, SPIE Press, 1998.

Figure 1 a) Frame1 & b) Frame2 is the looming square which generated with computer simulation. The images have a resolution of 128x128 pixels with 256 grey values. The segmentation from the original algorithm shows different parts of the looming object (c,d,e,f).

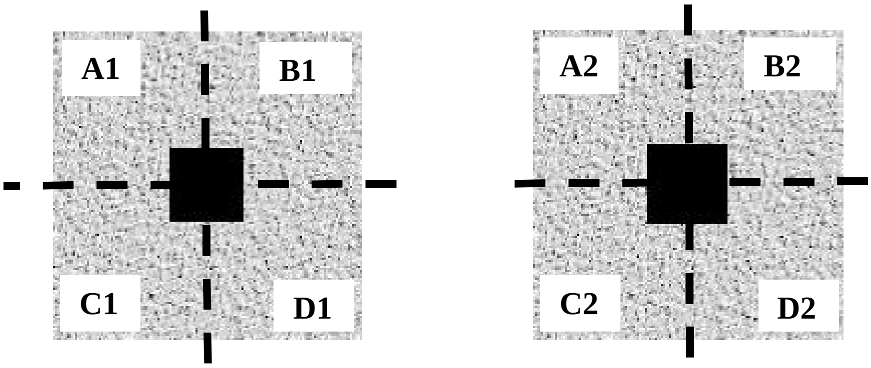


a) Frame 1 b) Frame 2



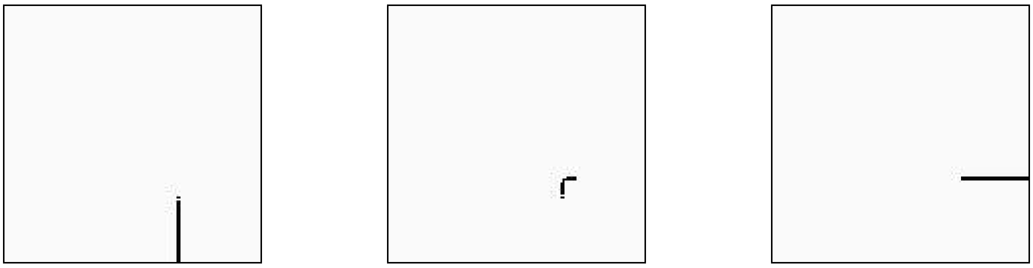
c) d) e) f)

Figure 2 : a) Frame 1 b) Frame 2, the looming square sequence with the textured background which are subdivided into four sub regions as A,B,C and D.



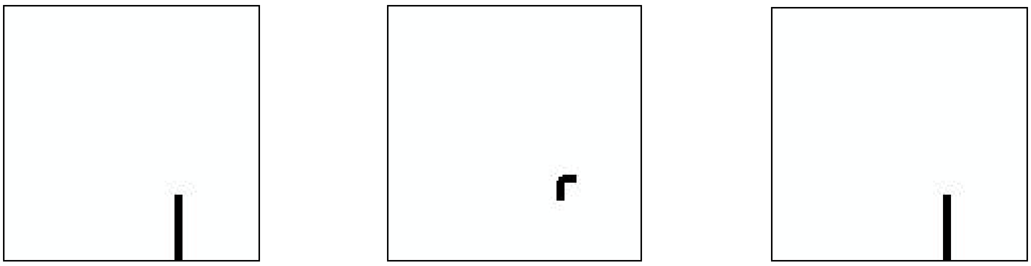
a) Frame 1 b) Frame 2

Figure 3: The segmented objects from looming square sequence of sub-region A, which consist of 3 output objects.



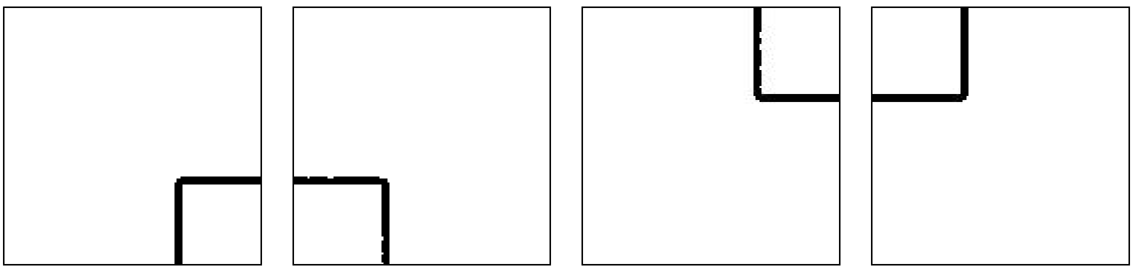
a) Object 1 b) Object 2 c) Object 3

Figure 4: The segmented objects from looming square sequence of sub-region A, which are processed in the enhancement stage.



a) Object 1 b) Object 2 c) Object 3

Figure 5: The segmented objects from Boolean algebra of sub-region A,B,C&D, which are checked for connectivity between them using Boolean function.



a) A b) B c) C d) D

Figure 6: The segmentation decision of the final stage.

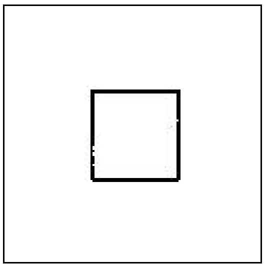
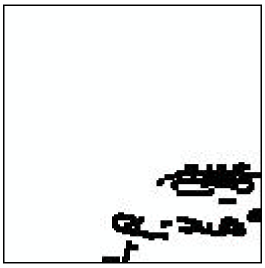
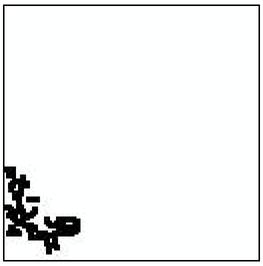


Figure 7: a) Frame 14 , is the car looming sequence on the road taken from moving camera. The images have a resolution of 128x128 pixels with 256 grey values. b),c),d) and e) are the segmented decision of the sub-region of A, B, C and D. f)The cross-correlation stage gives the final decision of the segmentation.

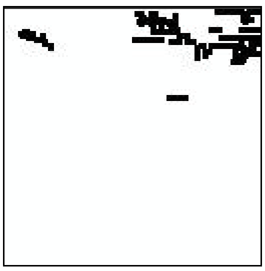
a) Frame 14



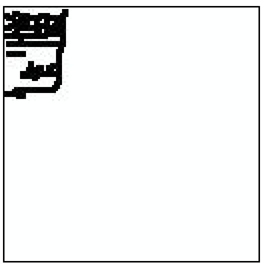
b)



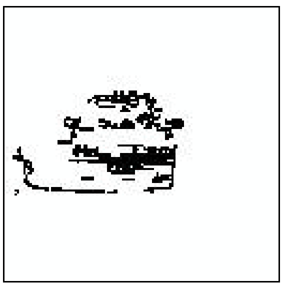
c)



d)



e)



f)