IJISET - International Journal of Innovative Science, Engineering & Technology, Vol. 1 Issue 2, April 2014.

www.ijiset.com

ISSN 2348 - 7968

Video Segmentation in Dynamic Texture Using Weber's Law for Monitoring Different Environments

Seetharaman.K [#1], Palanivel.N [#2], Sowmya.D[#3]
Associate Professor, Department Of Computer Science and Engineering[#1]
Annamalai University, Chidambaram,
Assistant Professor, Department Of Computer Science and Engineering[#2]
Manakula Vinayagar Institute of Technology, Puducherry.
UG Student, Department Of Computer Science and Engineering[#3]
Manakula Vinayagar Institute of Technology, Puducherry.

Abstract- **Dynamic texture is the texture which is in motion. Dynamic texture segmentation is a challenging task, as the texture can change in shape and direction over time. In this paper segmentation of Dynamic textures into distinct regions has been done. For the segmentation of dynamic texture, three different techniques are combined together to obtain better segmentation. Two local texture descriptors called Local binary pattern and Weber local descriptor are used. These descriptors, when used in spatial domain helps to segment a frame of video into distinct regions based on the histogram of the region. Also using the same texture descriptors in temporal domain, it is possible to obtain the dynamic texture in a given video. In addition to these texture descriptors, optical flow of pixels is used for detecting dynamic texture in a video, as optical flow is the natural method for detection of motion. After computing the three different features from multiple split sections of a group of video frames, individual histograms are obtained for each of the split sections of the video. These histograms each are converted to a single value called the Cumulative. The cumulative so obtained is compared with the threshold and filtered to obtain the dynamic texture. Since the histograms are converted to a single value, the computation of threshold is very easy as the whole set of values of Cumulative falls in two different set of values. The magnitude of motion detected depends on the threshold selected.**

***Keywords*— Dynamic texture— Dynamic texture, segmentation, Cumulative, Texture descriptor, optical flow.**

1. INTRODUCTION

DYNAMIC textures or temporal textures are textures with motion .Dynamic textures could be loosely described as visual processes, which consist of a group of particles with random motion .The particles can be macroscopic (e.g., foliage flying in the wind), microscopic (e.g., fire plume and smoke), or even moving objects (e.g., a flock of birds). Potential applications of DT analysis include remote monitoring and various types of surveillance in challenging environments, such as monitoring forest fires to prevent natural disasters, traffic monitoring, homeland security applications, and animal behavior for scientific studies, video synthesis, motion segmentation, and video classification. Segmentation is one of the basic problems in computer vision. Meanwhile, DT segmentation is very challenging compared with the static case because of their unknown spatiotemporal extension, the different moving particles, and stochastic nature of the motion fields. DT segmentation is to separate the different groups of particles showing different random motion. In this paper, we propose a new method based on both appearance and motion information for the segmentation of dynamic textures. For the appearance of DT, we use local spatial texture descriptors to describe the spatial mode of DT; for the motion of DT, we use the optical flow and local temporal texture descriptors to represent the movement of objects, and employ the Histogram of Oriented Optical Flow (HOOF) approach to organize the optical flow of a region. To compute the distance between two HOOFs, we develop a new distance measure based on Weber's Law, which is simple and efficient. The motivation to employ both the appearance and motion modes for the DT segmentation is that DTs might be different from their spatial mode (i.e., appearance) and/or temporal mode (i.e., motion field). Combining the spatial and temporal modes, we exploit the fuse of discriminant features of both the appearance and motion for the robust segmentation of cluttered DTs. For the spatial mode of DT, we use the simple but effective local texture descriptor, i.e., local binary pattern (LBP) and Weber local descriptor (WLD) .LBP is robust to monotonic gray-scale changes caused, e.g., by illumination variations. It is widely used, e.g., in texture analysis. WLD is also simple and effective for texture classification. We just use one of its components, i.e., differential excitation. However, we still call it

WLD for consistency. Here, we use both LBP and WLD because they are complementary to each other

2. RELATED WORK

There are several methods for dynamic texture segmentation and the methods for dynamic texture segmentation mainly fall under two models: generative models and discriminative models... A generative model is a model for randomly generating observable data, typically given some hidden parameters. After that, they proposed the layered dynamic texture (LDT) to represent a video as a collection of stochastic layers of different appearance and dynamics and then proposed a variational approximation for the LDT that enables efficient learning of the model.

In this paper a new method for segmentation of dynamic texture has been proposed. The general idea of any dynamic texture segmentation using texture as the key involves identifying areas of similar texture using the histogram and keeping them as a single set. The same technique is followed in this paper with a slightly different approach. The proposed new method involves using the texture descriptors used just as in [1, 8] and also using the optical flow proposed by Berthold K.P. Horn and Brian G. Schunck and further improved as in [10] and combining all three features followed by the new method of converting the entire histogram to a single value called the Cumulative for better visual identification of the dynamic texture.

3. PROPOSED SYSTEM

A new method is proposed based on both appearance and motion information for the segmentation of dynamic textures. For appearance – we use local spatial texture descriptors to describe the spatial mode of DT. For motion - we use the optical flow and local temporal texture descriptors to represent the movement of objects, and employ the Histogram of Oriented Optical Flow (HOOF) approach to organize the optical flow of a region. *:* The framework followed in this paper is shown in figure 1. The framework shows that a video is given as input. Now the whole video is split into multiple sections. For each section the features are computed which includes LBP_{TOP}, WLD_{TOP} and optical flow. The computation of LBP_{TOP} and WLD_{TOP} results in output which is in the form of histogram with each split section having a unique histogram. Similarly the output after computation of optical flow is in terms of vector quantities. Unlike Hierarchical splitting followed in [1, 8], the splitting is done equally into parts and the size of the split section depends on the video so considered for segmentation. Unlike [1, 8] the histogram comparison is not done bin by bin. Instead the whole histogram is converted to a single value called the Cumulative which is a value that is calculated from the histogram of each split section of video. This value represents the uniqueness of each of the histogram of the split section and this value is compared to a threshold for segmenting the dynamic texture. The calculation of optical flow helps in identification of the motion of pixels in vector quantities. Thus calculation of optical flow not only determines the motion but also the direction of motion of pixels. In this paper only the magnitude of motion detected is taken into account.

4. CLASSIFICATION AND TRAFFIC RELATED VIDEO

In recent years, the use of video systems for traffic monitoring has shown promise over that of traditional loop detectors. The analysis of traffic video can provide global information, such as overall traffic speed, lane occupancy, and individual lane speed, along with the capability to track individual cars. Because video systems are less disruptive and less costly to install than loop detectors, interest has grown in building and using large camera networks to monitor different aspects of traffic, such as traffic congestion.

Most of the existing work in monitoring traffic uses a vehicle segmentation and tracking framework. First, a potential vehicle is segmented from the scene using motion cues. Once segmentation is performed, the objects are tracked between frames using various methods, such as rule-based systems or Kalman filters. The vehicle tracking framework has the disadvantage that its accuracy is dependent on the quality of the segmentation. The segmentation task becomes more difficult with the presence of adverse environmental conditions, such as lighting (e.g. overcast, glare, night), shadows, occlusion, and blurring. Furthermore, segmentation cannot be performed reliably on low resolution images where the vehicles only span a few pixels. Tracking algorithms also have problems when there are many objects in the scene, which is typically the case for highways scenes with congestion.

In this project, we model the motion field as a whole using a generative probabilistic model, the dynamic texture. The motion from the video sequence is abstracted and represented as a generative model, thus avoiding the difficulties of segmentation and tracking. Furthermore, using a probabilistic model allows for a rich variety of distance functions and classifiers that are based on probability and information theory. In particular, classification can be performed using the probabilistic SVM framework, which combines the generalization guarantees of the large-margin SVM method, with the robustness of the underlying probabilistic models. For retrieval, the probabilistic model allows for advanced systems such as Bayesian retrieval.

ISSN 2348 - 7968

Experimental results show that this framework is useful in tasks such as traffic video retrieval and classification of congestion in traffic sequences, and that the model is robust to lighting conditions, occlusion, and blurring.

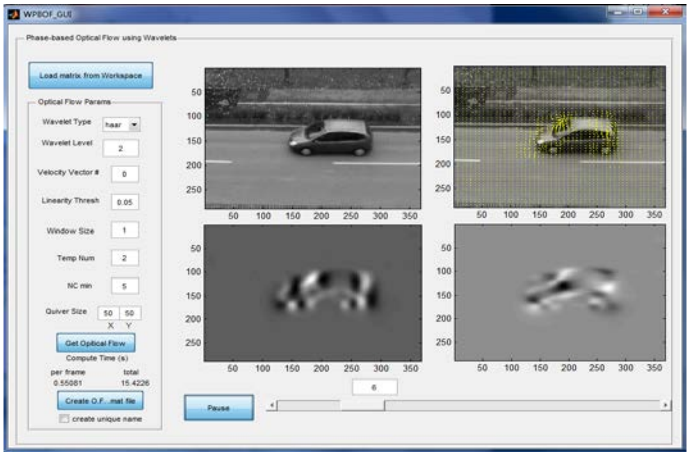


Fig.1 Flow chart of Framework

5. MOTION SEGMENTATION USING MIXTURE OF DYNAMIC TEXTURES

To segment dynamic texture from a video two texture descriptors are computed from the input video. These texture descriptors are used as spatial–texture descriptor when used in XY plane of a video. These are called temporal-texture descriptors when they are used in XT and YT plane. Since these texture descriptors are used in both spatial and temporal domain, they can be called as spatiotemporal descriptors. One family of visual processes that has relevance for various applications of computer vision is that of, what could be loosely described as, visual processes composed of ensembles of particles subject to stochastic motion. The particles can be microscopic, plumes of smoke, macroscopic, e.g. leaves and vegetation blowing in the wind, or even objects, e.g. a human crowd, a flock of birds, a traffic jam .The applications range from remote monitoring for the prevention of natural disasters, e.g. forest fires, to background subtraction in challenging environments, e.g. outdoors scenes with vegetation, and various type of surveillance, e.g. traffic monitoring, homeland security applications, or scientific studies of animal behavior.

Despite their practical significance, and the ease with which they are perceived by biological vision systems, the visual processes in this family still pose tremendous challenges for computer vision. In particular, the stochastic nature of the associated motion fields tends to be highly challenging for traditional motion representations such as optical flow, which requires some degree of motion smoothness, parametric motion models, which assume a piece-wise planar world, or object tracking, which tends to be impractical when the number of subjects to track is large and these objects interact in a complex manner.

The main limitation of all these representations is that they are inherently local, aiming to achieve understanding of the whole by modeling the motion of the individual particles. This is contrary to how these visual processes are perceived by biological vision: smoke is usually perceived as a whole, a tree is normally perceived as a single object, and the detection of traffic jams rarely requires tracking individual vehicles. Recently, there has been an effort to advance towards this type of holistic modeling, by viewing video sequences derived from these processes as dynamic textures or, more precisely, samples from stochastic processes defined over space and time

In this, we extend the simple dynamic texture model to a mixture of dynamic textures, where the observed video is an instance of one of several possible dynamic texture models. With this framework we are able perform motion segmentation of a video sequence, along with clustering on a set of video sequences. Experimental results show that the model is capable of achieving segmentation and clustering that are perceptually plausible.

Fig.2 Motion Segmentation

6. FEATURES USED FOR SEGMENTATION

A. Texture Descriptors

To segment dynamic texture from a video two texture descriptors are computed from the input video. These texture descriptors are used as spatial–texture descriptor when used in XY plane of a video. These are called temporal-texture descriptors when they are used in XT and YT plane. Since these texture descriptors are used in both spatial and temporal domain, they can be called as spatiotemporal descriptors.

XT plane denotes the change in pixels row-wise over temporal domain. YT plane denotes the change in pixels column-wise over temporal domain. LBP and WLD are computed in all three planes and hence it is called as LBP_{TOP} and WLD_{TOP} respectively.

LBP histogram for edge detection is computed in the very same way for all the three planes considered (XY, XT, YT) and it is computed as,

$$F_{edgeness} = \frac{|\{p|Mag(p) > T\}|}{N}$$

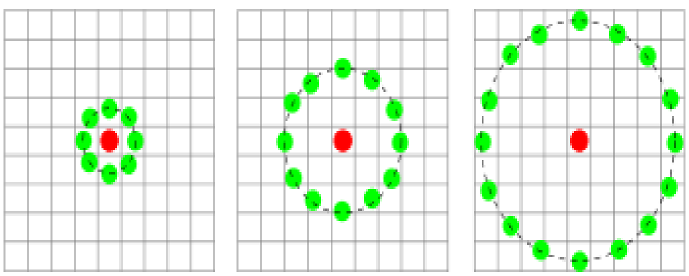


Fig 3 Local Binary Pattern

B. Optical Flow

Optical flow is the distribution of apparent velocities of movement of brightness patterns in an image. Optical flow can arise from relative motion of objects and also that of the viewer. This implies that, optical flow can give important information regarding the spatial arrangement of the objects viewed and also the rate of change of this arrangement. Discontinuities in the optical flow can help in segmenting images into regions that correspond to different objects. To compute the optical flow between two images, the optical flow constraint equation has to be solved. This equation is general for many of the methods that are used to compute optical flow.

$$I(x,y,t) = I(x+\Delta x, y+\Delta y, t+\Delta t)$$

C. Steps for Segmentation

Image segmentation is the process of partitioning a digital image into multiple segments. There are mainly 3 steps involved in segmentation: splitting, agglomerative merging and pixel-wise classification. When an input video is given, the whole video is split into multiple equal sections and features like LBP, WLD and OF are computed for each and every section independently. Using these features calculated from each of these split sections the dynamic texture segmentation is done. The segmentation procedure can be mainly classified into three distinct steps and they are: splitting, agglomerative merging and pixel-wise classification.

A. Splitting

The input video is read as frames and the frames are split equally into four parts or sixteen parts or sixty four parts and so on. The number of sections to which the video is split depends on the video of interest. For each of these split sections of the frames of video, the texture descriptors are calculated in XY, XT and YT plane. Both of these texture descriptors are used together and each of them are given equal weights of 0.5. The texture descriptors are calculated for each of these planes and each plane holds different weights. The XY, XT and YT planes are given weights of 0.2, 0.4 and 0.4 respectively. This ensures that temporal information is more prominent in the resultant histograms.

B. Pixel-wise Classification

Pixel-wise classification is done to better localize the boundaries of the roughly segmented section of a video. This operation is done on the boundary pixel-set of the dynamic texture. For each pixel in the boundary pixel-set, the LBP_{TOP}, WLD_{TOP} and OF is computed over an 8 x 8 neighborhood. In the same way as before, the Cumulative is calculated and compared with the threshold. After the pixels are classified as dynamic texture or otherwise, the boundary of the pixels that were classified as dynamic texture is further analyzed to refine the boundary.

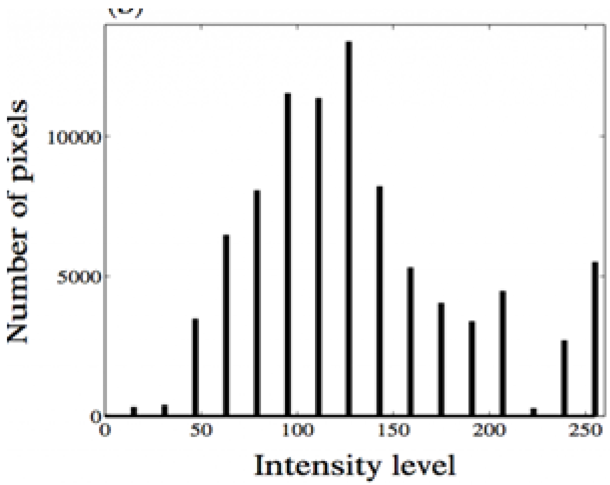
c. Agglomerative Merging

After having computed the Cumulative from various histograms of patches of roughly uniform texture, the Cumulative having same or nearly same values are kept together and those with completely different set of values are taken as another.

IJISET - International Journal of Innovative Science, Engineering & Technology, Vol. 1 Issue 2, April 2014.

www.ijiset.com

ISSN 2348 - 7968



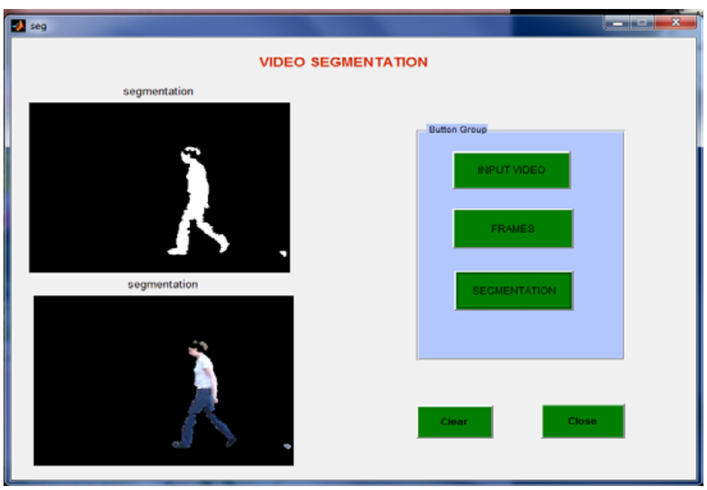


Fig.4 Video Segmentation

7. CONCLUSION

Thus a new method has been proposed for dynamic texture segmentation using spatial and temporal descriptors and optical flow of pixels. This involves computing histograms for spatial mode and temporal mode using LBP and WLD. Also optical flow is used in temporal mode as it is the natural method for segmentation. From the computed histograms the Cumulative is calculated and this is used as a measure for comparison of the computed histograms. Experimental results suggest that this method produces better results with less overhead.

REFERENCES

[1] Jie Chen, Guoying Zhao, Mikko Salo, Esa Rahtu, and Matti Pietikäinen, Automatic Dynamic Texture Segmentation using Local Descriptors and Optical Flow., *IEEE Trans. On Image Processing*, Vol. 22(1), pp. 326-339, 2013.

[2] G. Doretto, A. Chiuso, Y. N. Wu, and S. Soatto, “Dynamic textures ,”Int. J. Comput. Vis., vol. 51, no. 2, pp. 91–109, 2003.

[3] M. Szummer and R. W. Picard, “Temporal texture modeling,” in Proc. IEEE Int. Conf. Image Process., Sep. 1996, pp. 823–826.

[4] A. B. Chan and N. Vasconcelos, “Modeling, clustering, and segmenting video with mixtures of dynamic textures,” IEEE Trans. Pattern Anal.Mach. Intell., vol. 30, no. 5, pp. 909–926, May 2008.

[5] T. Amiaz, S. Fazekas, D. Chetverikov, and N. Kiryati, “Detecting regions of dynamic texture,” in Proc. Conf. Scale Space Variat. Methods Comput.Vis., 2007, pp. 848–859.

[6] D.W.Murray and B. F. Buxton, “Scene segmentation from visual motionusing global optimization,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 9, no. 2, pp. 220–228, Mar. 1987.

[7] R. Polana and R. Nelson, “Temporal texture and activity recognition,”in Motion-Based Recognition. Norwell, MA: Kluwer, 1997.

[8] T. Ojala, M. Pietikäinen, and T. Mäenpää, “Multiresolution gray-scale and rotation invariant texture classification with local binary patterns,”IEEE Trans. Pattern Anal. Machine Intell., vol. 24, no. 7, pp. 971–987, Jul. 2002.

[9] J. Chen, G. Zhao, and M. Pietikäinen, “An improved local descriptor and threshold learning for unsupervised dynamic texture segmentation,” in Proc. 12th IEEE Int. Conf. Computer. Vis. Workshop, Oct. 2009, pp. 460–467.