

DYNAMIC TEXTURE BASED TRAFFIC VEHICLE MONITORING SYSTEM

Amit Y Kurane¹, Priti P Rege²

¹College of Engineering, Pune, India

²College of Engineering, Pune, India

Abstract

Dynamic Textures are function of both space and time which exhibit spatio-temporal stationary property. They have spatially repetitive patterns which vary with time. In this paper, importance of phase spectrum in the signals is utilized and a novel method for vehicle monitoring is proposed with the help of Fourier analysis, synthesis and image processing. Methods like Doppler radar or GPS navigation are used commonly for tracking. The proposed image based approach has an added advantage that the clear image of the object (vehicle) can be used for future reference like proof of incidence, identification of owner and registration number.

Keywords-Fourier Transform, dynamic texture, phase spectrum

1. INTRODUCTION

Dynamic textures are function of both space and time which exhibit spatio-temporal stationary property. They have spatially repetitive patterns which vary with time. In image synthesis, textures are often used to designate things like candle flame, bicycle wheel, traffic on road, rafting, flames coming out of chimney. Nowadays, traffic congestion is becoming very severe problem in the developed and developing countries. Different types of roads are used for different purposes. The roads within the cities or small roads cannot be used for container like oversized heavy vehicles which can damage the normal roads. Also, there are express ways on which only four wheelers and other heavy vehicles are allowed on the road. There are other cases where two wheelers are not allowed. In this paper, we propose an image processing based scheme to detect various sized vehicles. The scheme can be used to ban a particular of vehicles on specific road. The scheme is extended to calculate speed of the vehicle.

The rest of the paper is consolidated as follows. The related work is reviewed in Section 2. The proposed approach is presented in Section 3 and 4 experimental results are demonstrated in Section 5. Final conclusions are summarized in Section 6.

2. LITERATURE REVIEW

Fourier transform is useful for an overall picture of multidimensional content. Although phase spectrum obtained using Fourier transform is complex and difficult to understand, it contains more information about signal than magnitude and it is shown that using iterative algorithm [4], original multidimensional signal can be reconstructed from its phase only. Motion analysis performed using optical flow algorithm is best suited only in case of sequences which have continuous, local and smooth motion. Ghanem in [1]

justified dynamic textures as certain class of video signals where the phase values captures complex, rapidly varying global motion.

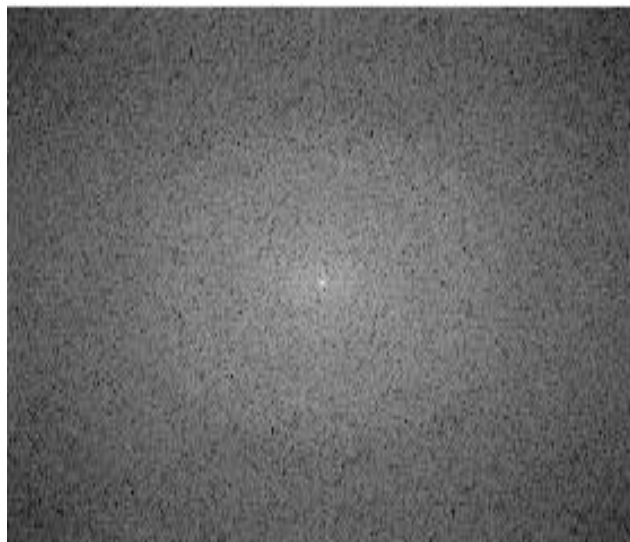
In this paper, 3D Fourier spectrum of road video is used for dynamic texture based segmentation of various vehicles. Vehicles are classified as two wheelers / four wheelers / and oversized heavy vehicles.

3. 3D FOURIER TRANSFORM BASED DYNAMIC TEXTURE SEGMENTATION

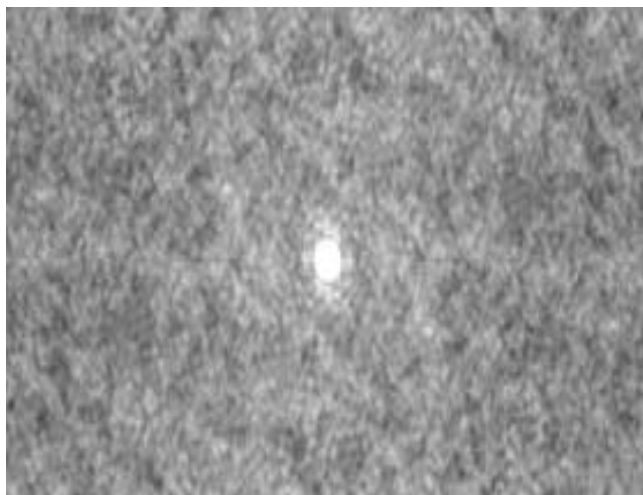
As shown in Figure 1(c) and (d), image is reconstructed from magnitude spectrum and phase spectrum respectively using test image shown in Figure 1(a). The phase values determine the shift in the sinusoid components of the image. With zero phase, all the sinusoids are centered at the same location and a symmetric image is obtained whose structure has no real correlation with the original image at all. Being centered at the same location means that the sinusoids are a maximum at that location, and is why there is a big white patch in the middle of Figure 1(c).



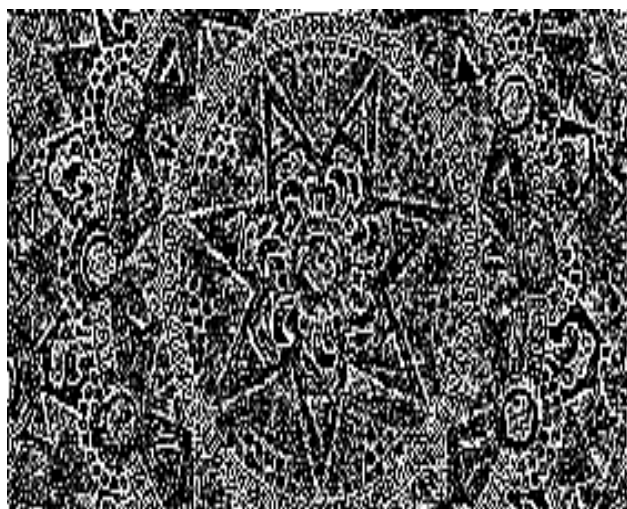
(a)



(b)



(c)



(d)

Fig 1: Image reconstruction (a)Test image, (b) Fourier spectrum, (c) Image reconstruction from Magnitude only, and (d) Image reconstruction from Phase only

Consider a case as shown in Figure 1(d) where an image is reconstructed from magnitude only. For an image with dimensions X and Y, the 2D DFTF (0, 0) at zero frequency value is given as

$$F(0,0) = XY \frac{1}{XY} \sum_{x=0}^{X-1} \sum_{y=0}^{Y-1} f(x,y) \quad \text{---(I)}$$

i.e. $|F(0,0)| = (XY) * \text{average value of } f(x,y)$

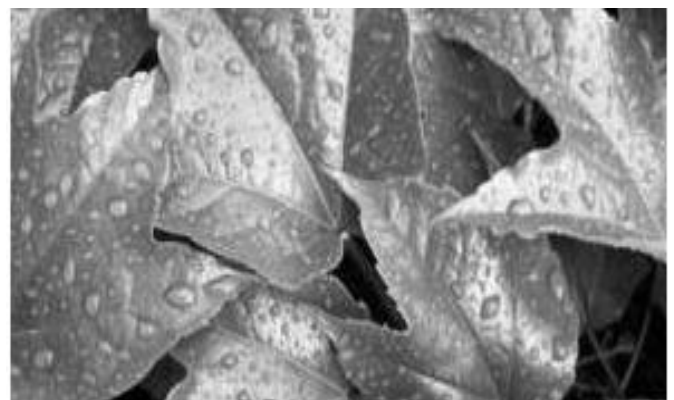
Here, as proportionality constant XY is very large, the $|F(0,0)|$ term i.e. dc component the largest component of spectrum has magnitude that can be several orders larger than the other terms. In another case i.e. when the image is reconstructed from phase only, imagefunction $f(x,y)$ which has relation with $F(u,v)$ can be given as

$$f(x,y) = \frac{1}{|F(u,v)|} \sum_{u=0}^{U-1} \sum_{v=0}^{V-1} F(u,v) e^{\frac{1j*2\pi*ux}{U}} e^{\frac{1j*2\pi*vy}{V}} \quad \text{---(II)}$$

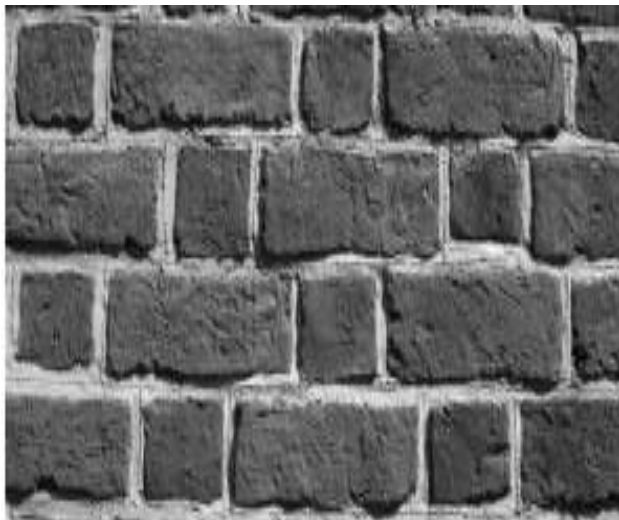
Here, all the component sinusoids are given the same magnitude. This basically normalizes the brightness everywhere in the image. Since most of the signals have low intensity values at high frequencies the phase only reconstruction will give emphasis to high frequency components like texture, line or edges. Also, it determines the shape of the image.

The phase-only reconstruction preserve features because of the principle of phase congruency. At the location of edges and lines, most of the sinusoid components have the same phase. This properly alone can be used to detect lines and edges without considering magnitude value [9]. Thus, even if the phase spectrum seems to be complex one, the phase information is most important.

As shown in Figure2, Magnitude and Phase values are interchanged among two test texture images Figure2 (a) and (b) to obtain two new texture images Figure2(c) and (d). It can be observed from the Figure 2(c) and (d) that in each case phase has dominance in determining the feature content of an image.



(a)



(b)



(c)



Fig 2: Phase and Magnitude exchange (a) Texture 1, (b) Texture 2 (c) Texture reconstruction from magnitude of Texture 1 and phase of Texture 2, (d) Texture reconstruction from magnitude of Texture 2 and phase of Texture 1

Changing the magnitude of the various component sinusoids changes the shape of the feature. When a phase-only reconstruction is performed, all the magnitudes are set to unity magnitude value, which changes the shape of the features, but not their location. In many images, the low frequency components have a magnitude higher than the high frequency components, so phase-only reconstruction does look like a high-pass filter. In short, phase contains the information about the locations of features.

The proposed approach is motivated by [3], which used the important characteristics of phase in case of dynamic texture and performed dynamic texture segmentation. As shown by Oppenheim's experiment [1], the phase spectrum contains most of the structural information about the image. In 2D, these are reflected as lines and edges. In 3D, the phase also represents lines and edges but also movement. Instead of 2D frames and time, we imagine the video as a 3D solid where the z-axis is the frame number. If a slice is taken along the z-axis (3rd dimension), movement appears like an edge in the signal. When reconstruction is performed using just the phase spectrum, all the component sinusoids obtained are of the same magnitude. This basically normalizes the brightness everywhere in the image. In dynamic texture video, this normalization is also in the time axis. Hence, smaller amplitude moving parts becoming greater in amplitude (and the opposite for high amplitude parts).

The 3-D Fourier transform of the input sequence is given by,

$$I(u, v, s) = F\{I(x, y, n)\} \\ = \frac{1}{XYN} \sum_{x=0}^{X-1} \sum_{y=0}^{Y-1} \sum_{n=0}^{N-1} I(x, y, n) \cdot e^{-j2\pi(ux/X+vy/Y+sn/N)} \quad \text{---(III)}$$

where X and Y are the width and height of each frame respectively and N is the total number of frames.

Similarly, the 3-D inverse DFT using only phase spectrum is obtained as

$$\hat{I}(x, y, n) = |F^{-1}\{e^{iP[I(u,v,s)]}\}| \quad \text{--- (IV)}$$

Hence, using the significance of the phase property the segmentation of the video sequence is performed.

The proposed work is extended and used for two application which are discussed below.

4. PROPOSED TRAFFIC MONITORING SYSTEM OF DYNAMIC TEXTURE

In this work, we are proposing two novel schemes for image based monitoring of vehicles.

- 1) Classification based on size of the Vehicles
- 2) Speed calculation of detected vehicle

4.1 Classification based on Size of the Vehicles

Algorithm is developed to classify the vehicles based on their sizes.

Algorithm -1

1. For each frame in a input video, perform step 2 and append step 2 result in resultant array 'I(x,y,n)'
2. Smoothen the current frame using 2D Gaussian filter
3. Perform 3D FFT for the whole sequence I(x, y, n) using (Eq.(III))
4. Calculate the phase spectrum using the real and imaginary parts of 3D DFT
5. Calculate the reconstructed sequence $\hat{I}(x, y, n)$ using (Eq.(IV))
6. For each frame in a input video, perform step 7 to step 10 to get segmentation mask for each frame and append step 10 result in resultant segmentation mask array 'S(x, y, n)'
7. Smooth the reconstructed frame of $\hat{I}(x, y, n)$ using the averaging filter.
8. Compute the mean value of the current frame
9. Convert the current frame into binary image using mean value as the threshold
10. Perform morphological processing, i.e., filling and closing, to obtain segmentation mask for the current frame
11. Select the predefined region from the segmented frame and perform labelling operation on all k detected vehicles
12. Count number of pixels N_k in each k^{th} detected vehicle in current frame

13. Set two threshold values TH1 for minimum value and TH2 for maximum value
14. For each k^{th} detected vehicles in current frame repeat step 15 to step 19
15. If $(N_k < TH1)$ or $(N_k > TH2)$
16. Mark the k^{th} vehicle as banned
17. Otherwise
18. Mark the k^{th} vehicle as allowed
19. End if
20. End Algorithm.

With the proposed approach, number of applications can be developed such as entry for particular vehicle can be restricted or number of vehicles passed through given time can be calculated. It can be further modified for continuous tracking of single vehicle. There are classic methods like Doppler radar or GPS navigation which are used for tracking. Using the proposed approach enjoy advantage that a clear image of that object can be used for future reference, for further formalities.

4.2 Speed Calculation of Detected Vehicle

Suppose a car is moving on the road with certain speed. The side view and the front view of the car positions captured by the CCTV camera are shown in the Figure 3 and 4 respectively at different frames taken at certain interval. If one wants to calculate the speed of vehicle, then there is requirement of two quantities i.e. distance covered by the vehicle and time elapsed to travel the same distance.

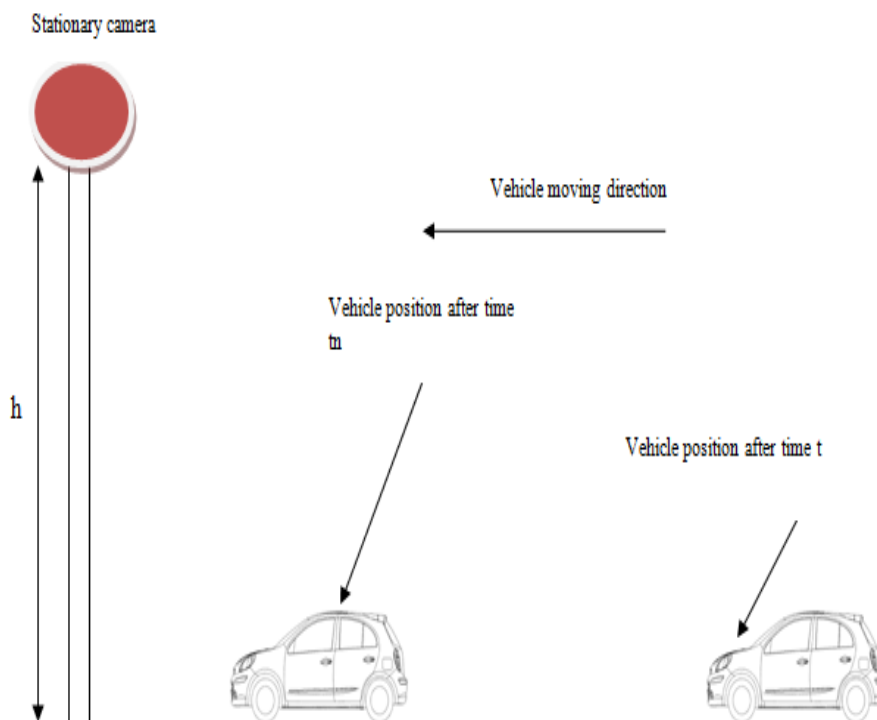


Fig 3: Side view of car positions at different frames

The total time elapsed by vehicle for moving from one position say x1 at frame 1 to next position say x2 at frame N is calculated using frame rate.

$$\text{Total time elapsed} = \frac{\text{No of Frames over which vehicle motion is captured}}{\text{Frame rate}} \quad \text{-- (V)}$$

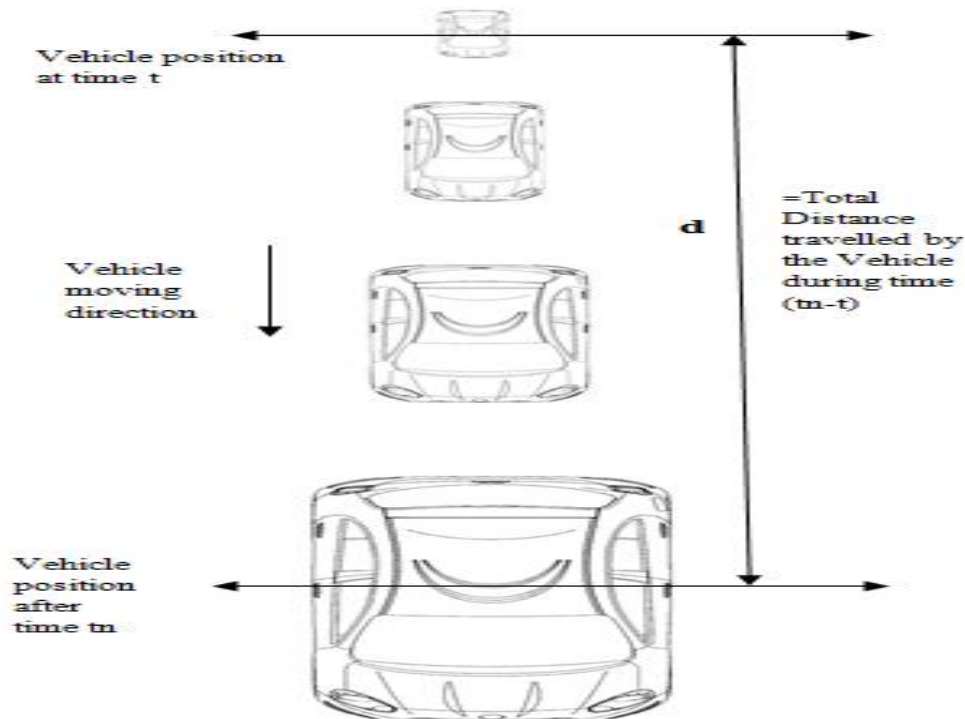


Fig 4: Front view of vehicle positions at different frames

The real work is to compute the distance. As the vehicle comes towards camera or moves away from camera, its size or shape goes on increasing or decreasing respectively and so there is need to use nonlinear scale to compute the distance.

$$\text{Speed of vehicle} = \frac{\text{Total Distance traversed by vehicle}}{\text{Total Time elapsed}} \quad \text{-- (VI)}$$

The second Algorithm proposed by us calculates the speed of vehicle using the dynamic textures depicting the vehicle. Centroid distance among two different positions of a vehicle in two different frames over which vehicle motion is captured is calculated and projected to actual physical distance on road. Using this distance and by knowing time between frames, speed of vehicle is calculated.

Algorithm -2

1. For each frame in a input video perform step 2 to step 8
2. Perform labelling operation on all k detected vehicles in the segmented frame obtained from Algorithm 1
3. Calculate centroid of vehicle in each frame for which speed being calculated
4. Calculate the pixel distance travelled by vehicle from the centroid points of vehicle between two different frames

5. Convert the pixel distance(2D) into actual pixel distance(3D) by using Homographic transformation and Inverse Perspective mapping
6. Convert the actual pixel distance into physical distance on road
7. Calculate the time elapsed from Eq. (V)
8. Calculate the speed of vehicle from Eq. (VI)
9. End Algorithm.

5. EXPERIMENTAL RESULTS

In this section, proposed approach is used to perform segmentation of dynamic texture sequences. The size of each image sequence is 480 × 360. The format of sequence can be chosen to be avi or mp4. The radius of the circular averaging filter is set to 9 for each frame of the reconstructed sequence. The proposed approach is implemented in MATLAB and all the experiments are conducted on a core i3 laptop with 3GB of RAM. Figure 5 demonstrates the segmentation results achieved using the proposed approach to dynamic texture segmentation. The images on the left side of Figure 5(a), (b), (c), (d) and (e) show the original image, while the images on the right side shows the original image overlaid with the corresponding segmentation mask. Figure 6 shows segmentation results obtained at different frames of same video sequences.

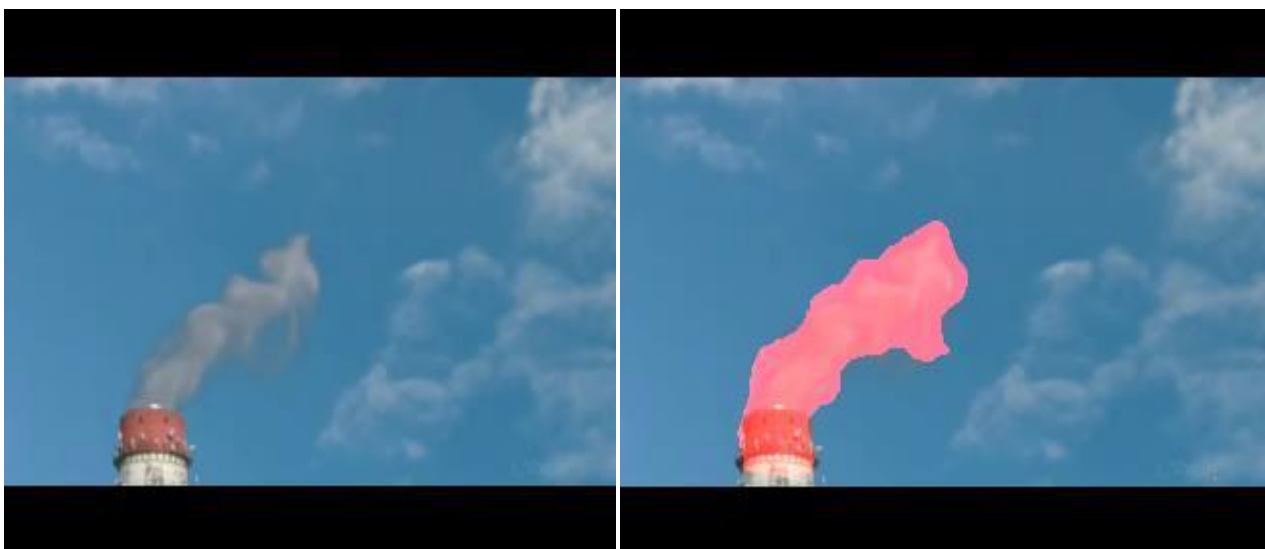
As shown in the Figure 7, experiment is performed on the videos of moving vehicles which are taken in different views. Using the proposed approach and setting predefined threshold value, the oversized vehicles are banned. Following results are obtained. In Figure 7(a) and (b), the left image is the video frame captured and the right one is the result image using Algorithm 1. Car is allowed vehicle so it is marked in Blue line and container is the banned vehicle so marked in Green. It can be seen that the proposed approach can achieve reasonably good segmentation for dynamic textures especially when the camera motion is small or the camera is static.

6. CONCLUSION AND FUTURE SCOPE

Phase conveys more information regarding signal structure than magnitude does, especially in the case of images or multidimensional signals. It is therefore imperative to use phase information in various signal/image processing schemes. In this paper, segmentation of dynamic textures is performed using 3-D Fourier transform. Further, we have developed two applications which are giving successful results. Future work can be developing an Algorithm which will try to bring out interdependence between neighboring video frames so that method can be generalized for more complex applications.

REFERENCES

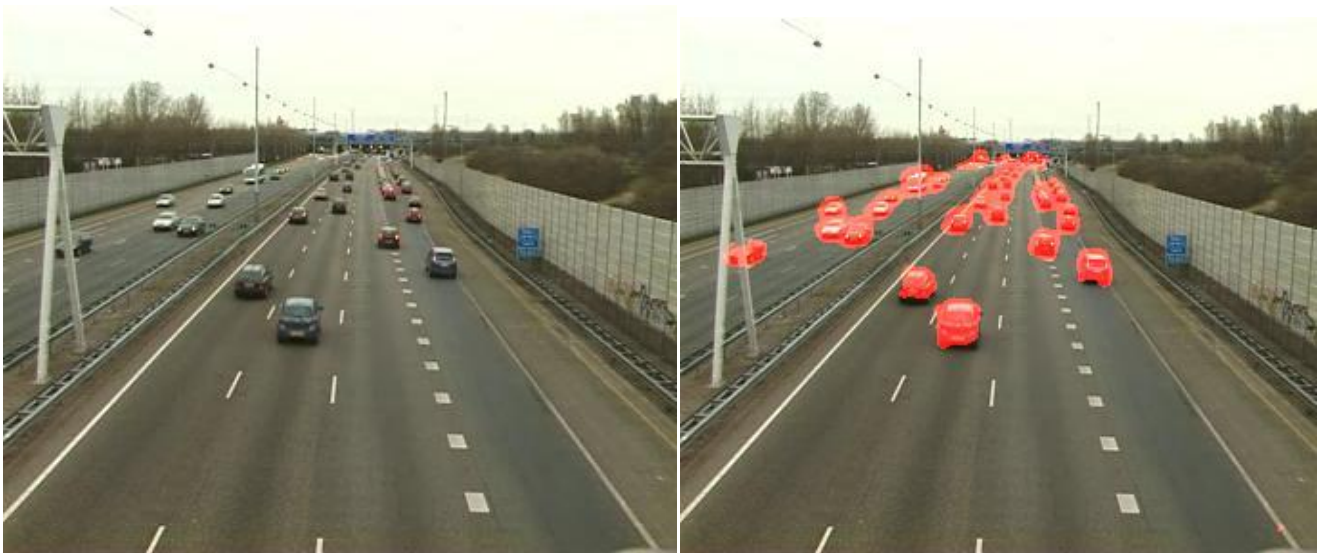
- [1] B. Ghanem and N. Ahuja, "Phase based modeling of dynamic textures", *IEEE 11th International Conference on Computer Vision*, 2007.
- [2] Alan V Oppenheim, Jae S LIM, "The Importance of Phase in Signals", *proc. of the IEEE*, vol. 69, no. 5, May 1981.
- [3] Jianghong Li, Liang Chen and Yuanhu Cai, "Dynamic Texture Segmentation Using Fourier Transform", *IEEE Fifth International Conference on Image and Graphics*, 2009.
- [4] M. Hayes, "The reconstruction of a multidimensional sequence from the phase or magnitude of its Fourier transform", *IEEE Trans. on Acoustics, Speech, and Signal Processing*, 30(2), 1982.
- [5] B. Karasulu and S. Korukoglu, "Moving Object Detection and Tracking in Videos", *Performance Evaluation Software, Springer Briefs 7 in Computer Science*, DOI: 10.1007/978-1-4614-6534-8_2, 2013.
- [6] B. Abraham, O. I. Camps, and M. Sznaier, "Dynamic texture with Fourier descriptors", *Proc. of the 4th Int. Workshop on Texture Analysis and Synthesis*, pp. 53–58, 2005.
- [7] Amiaz, S. Fazekas, D. Chetverikov and N. Kiryati, "Detecting regions of dynamic texture.", *Lecture Notes in Computer Science*, 4485:848, 2007.
- [8] A. B. Chan and N. Vasconcelos, "Mixtures of dynamic textures", *Tenth IEEE International Conference on Computer Vision*, 2005.
- [9] Peter Kovesi, "Phase Congruency Detects Corners and Edges", *School of Computer Science & Software Engineering, The University of Western Australia*, Crawley, W.A. 6009, 2002.



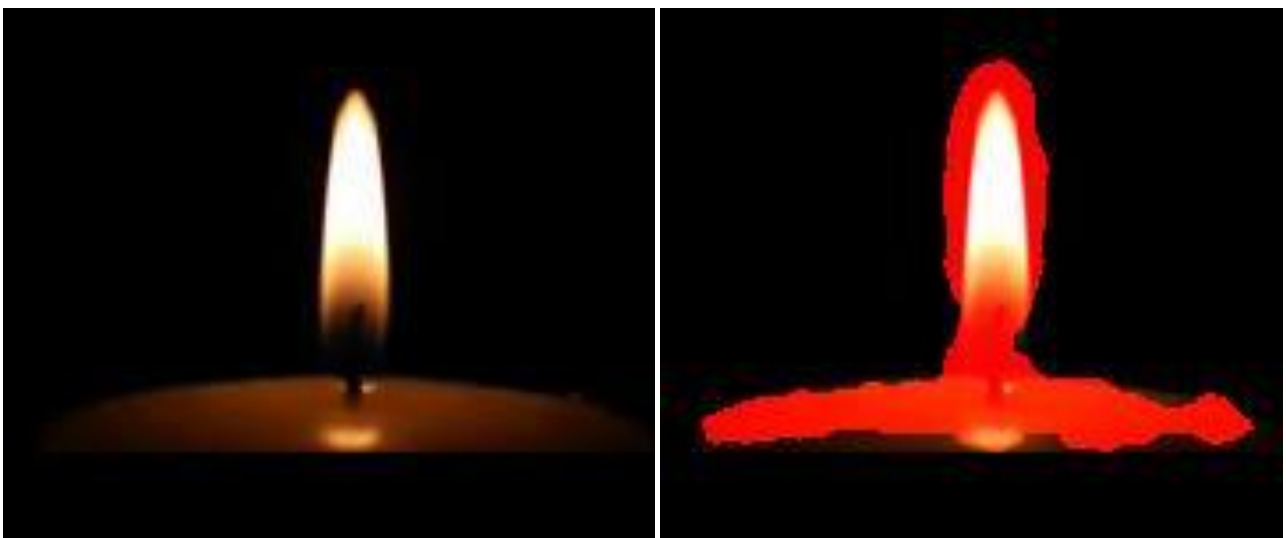
(a) Flames coming out of chimney.avi



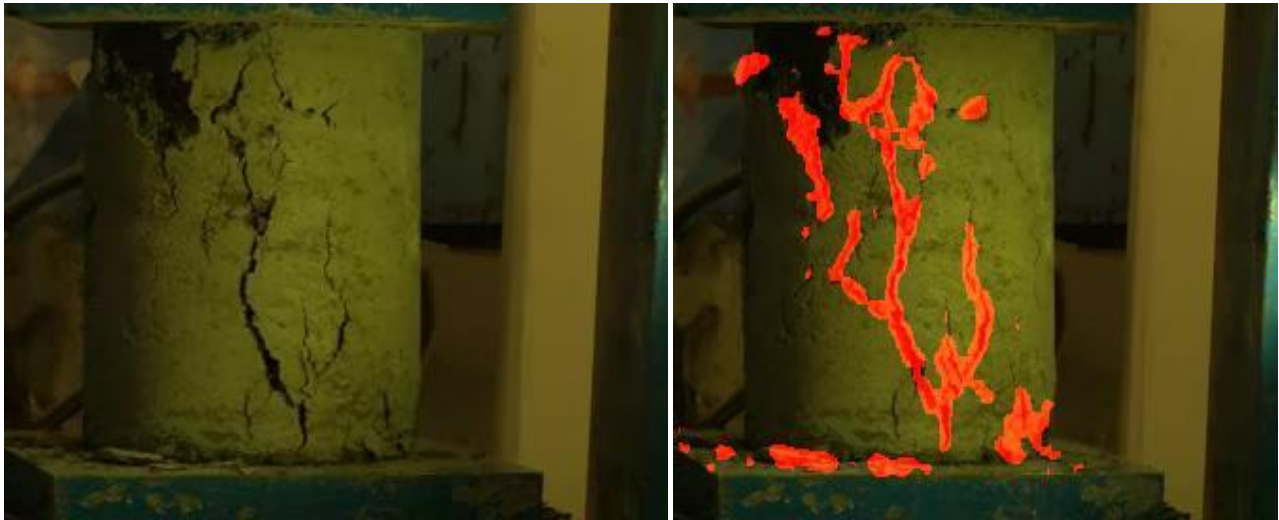
(b) Bicycle wheel.avi



(c) Traffic on road.avi



(d) Candle flame.avi



(e) Cylinder block put under pressure.avi

Fig 5: Results of Dynamic Texture Segmentation. Segmentation results obtained by using the proposed approach for dynamic texture sequences. The images on the left and right sides of (a), (b), (c), (d), (e) depict the original image and the corresponding mask with red colour, respectively. This figure is best viewed in colour.

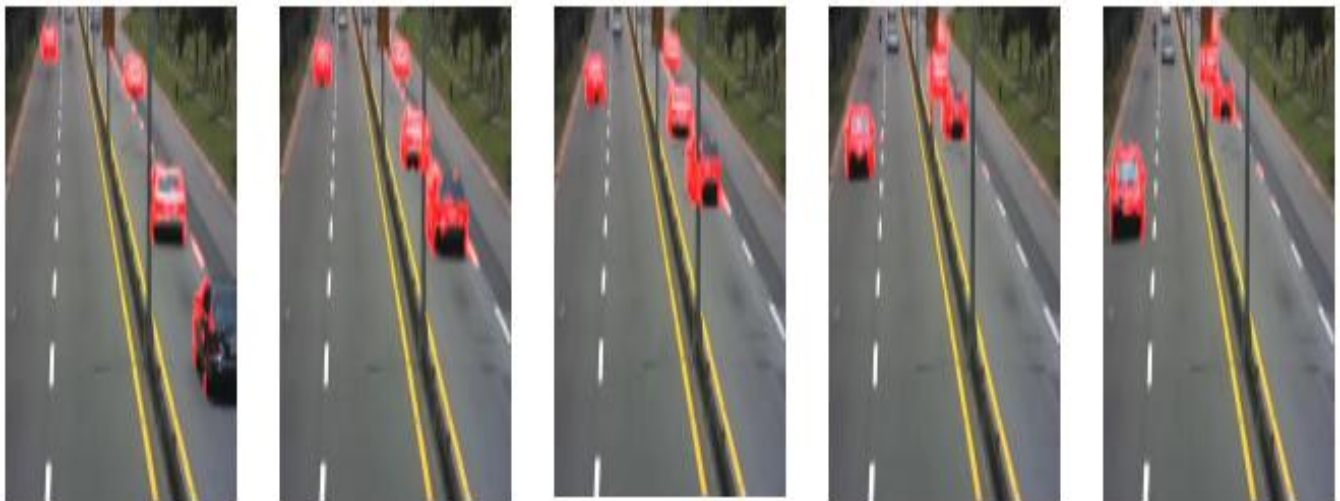


Fig 6: Segmentation results obtained for different frames of video of traffic on road



(a) vehicle moving on road_sample1.avi



(b) vehicle moving on road_sample2.avi

Fig 7: Results obtained by using the proposed approach for two traffic videos. The images on the left and right sides of (a) and (b) depicts the original image with vehicles moving on road and the corresponding banned vehicles marked with green and allowed vehicles marked in blue respectively.