Effective Flicker Detection Technique Using Artificial Neural Network for Video

Ripal Patel[1],Shubha Pandey[1],Chirag I Patel[2] and Robinson Paul[1]

[1] B.V.M College of Engineering, Gujarat, India.
[2] Nirma University, Gujarat, India.

Abstract

Flicker is the most common and an intolerable blemish present in signal processing world that leads to distortion in the transmitted frame of a video string. To dodge such misinterpretations a technique for detection of the flickering frame in a video is depicted in this research. Earlier methods were based on removing flicker by calculating the threshold of the consecutive frame difference and then finding the flickering frame. The proposed method in this research includes finding flickering frame using neural network concept. Therefore, the advantage of the practice disclosed here is that it removes the tedious calculation part of the threshold value and thereby the computational part becomes easier with added accurate result.

1 Introduction

Visible distortion between displaying of two frames on raster called flicker. It is the unnecessary fluctuations present in pixels of a frame which leads to distortion of the particular frame and thereby displaying erroneous video. The prime reason of production of this flicker is lower value of refresh rate, which drops the intensity for a noticeably long time, hence heed by human visual system. Due to flickering television or any video display devices, detrimental causes are generated in a human-like migraine or any hazardous effects. Therefore it is necessary to detect and remove flicker. In this paper, we have presented the approach for flicker detection from video frames.

As described in (Keesey, 1972) to see the variation in flicker and a flickering line as a function of frequency variation the threshold value of light should have different values. Sensitivity for flickering decreases at higher frequencies and therefore after using discrete cosine transform the higher magnitude value should be removed.

2 Related Work

The author proposed a method in (Nguyen, 2014) to eradicate motion induced flicker in a multifield per frame image display. It is performed by calculating gradient characteristics between a first field and subsequent fields in a multifield per frame display. The advantage of this method is that motion induced flicker is suppressed and data transmission requirements are reduced through the image processing method. In this flicker detection is performed using motion compensation method. It includes the steps of calculating a luminance difference between neighboring pixels; summing up the luminance difference in a specific direction if the difference is greater than a threshold value, performing motion estimation, calculating motion displacement; performing a motion compensation to a current frame; calculating a difference between a previous frame data and the compensated current frame data; and detecting the flicker by using the difference.

The advantage of the techniques for the flicker detection in (Nguyen, 2014) is that the flicker from fluctuating light resources is correctly perceived and decreased irrespective of the frequency of the fluctuating light. The flicker detection and correction engine chooses the first channel from a first image frame for dealing out and then subtracts each pixel value in the first channel from a corresponding pixel value in a prior image frame to generate a difference image frame. After that, the first alternating current component based on a discrete cosine transform associated with the difference image frame is calculated. After that, it reduces flicker that is present in the first image frame based on the first AC component.

Methods for detection in image flicker are cited in (Kinrot, 2011).An imaging device senses distorted the image and remove the distortion for the period of processing can use an analysis module to find statistics suggestive of image flicker earlier to eliminating the distortion. An imaging device has a diode for the illuminating field of view can use the diode as a photo sensor to decide one or more flicker. A digital imaging device may use imaging elements to produce an image having an acknowledged distortion, such as that commenced by one or more lenses or another optical element, and then create an exact image, in which some or all of the distortion has been removed.

Xiaojun has presented the method to detect fire from video sequences using flicker detection phenomenon (Qi, 2009). Flicker is assumed as high frequency component in the temporal domain, hence, the author used cumulative time derivative matrix. Firstly, the difference of luminance of consecutive frames is calculated. Then, the difference of luminance is added cumulatively. Finally, the frame considers flicker or not, that is based on the thresholding for converting in terms of binary image.

Another technique for flicker detection has developed for fire detection application (Töreyin, 2006). This method used temporal wavelet analysis for frequency and time localization of flicker. Author has developed high frequency analysis using wavelet transform of moving pixels. The track record of moving pixels has been stored and analyzed using two stage filter bank.

3 Proposed Approach

Figure 1 shows the proposed framework of flicker detection system.
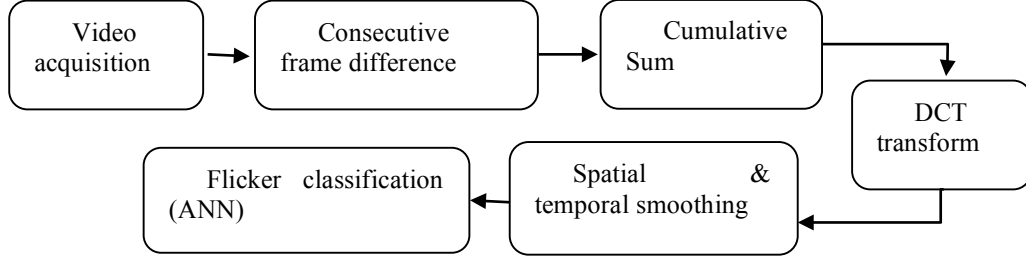
Figure 1: Proposed flicker detection approach

3.1   Frame Difference

Suppose that there is a change only in some part of a region in the frame and rest of the background remains the same. Therefore, since everything is same except for the change we can say that if we subtract a frame from its previous frame we will only get those areas in which change occurs. Therefore, the frame difference concept is used (Patel, 2016). In this, we subtract consecutive frame to get only the desired area of flickering.

Let $I_{present\_frame}$ be the present frame of a video and $I_{previous\_frame}$ is just its previous frame, then the difference frame $I_{difference\_frame}$ can be given as,

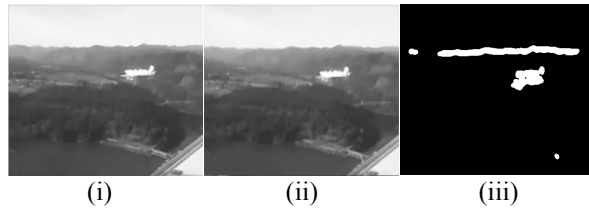$$I_{difference\_frame}=| I_{present\_frame}-I_{previous\_frame} | \qquad (1)$$

Figure 2: frame difference method (i)previous frame (ii)current frame (iii)frame difference

3.2   Cumulative Sum

Cumulative sum is the addition of different values in an array so that if there are n number of elements in an array they all can be sum together to give a single value in the array.

This process was used because if a number will have greater value in a n*n matrix then after summation the value across row or column will be enhanced only. Therefore, to have a single value in row or column in an array summation was taken. It can be defined using the concept as that in equation 2.

$$I_{sum}=sum(I_{difference\_frame}) \qquad (2)$$

The method was further subdivided into two parts.

a)   *Method 1*
      After the frame difference, the sum of all values was taken along the row.

b)   *Method 2*

After the frame difference, the sum of all values was taken along the column.

Figure 3 :Image showing the cumulative sum of a frame

3.3 Discrete Cosine Transform

It is related to Fourier transform which helps to find the Fourier series coefficients associated with the periodically extended sequence (Ahmed, 1974). They are used in various fields of science and engineering including audio and video compression. We have used this method so that the value of frame difference coefficient after its transformation can be known and based on that we can say that higher magnitude corresponds to the flickering region.

$$y(k) = w(k) \sum_{n=1}^{N} x(n)\cos\left(\left(\frac{\pi}{2N}\right)(2n-1)(k-1)\right), \qquad (3)$$

$k = 1,2,..N$

where

$$w(\text{k}) = \frac{1}{\sqrt{N}}, k = 1$$

$$w(\text{k}) = \sqrt{\frac{2}{N}}, 2 \leq k \leq N$$

Figure 4:Image showing the DCT of a frame

3.4   Temporal & Spatial Smoothing

The motive of smoothing is to give an idea of relatively slow changes of value with slight attention paid to the close identical of data values, while curve fitting concentrates on achieving as close a match as possible.

Temporal smoothing

For temporal smoothing, the moving average filter was used with three frames taken and then creating its average value.

The video sequence is taken and then its three frames are taken after that average is taken.

Let the current value of the pixel defined as I(x,y), where x and y are the coordinates. Then its previous pixel will have value I(x,y-1) and more like I(x,y-2).

$$I = \frac{[I(x,y,t-2)+I(x,y,t)+I(x,y,t-1)]}{3} \qquad (4)$$

Spatial smoothing

Spatial smoothing means that pixels are averaged with their neighboring pixel. This leads to removal of higher frequencies component which means it behaves like a low pass filter. As a result, sharp edges of the images are. The benefit of using this method is that it improves the signal to noise ratio (SNR) and thereby increasing sensitivity.

Moving average filter was used for spatial smoothing in which the average of three pixels value was calculated and its average was taken.

$$I = \frac{[I(x,y-2)+I(x,y)+I(x,y-1)]}{3} \qquad (5)$$

3.5   Artificial Neural Network

The purpose of the neural network is to curb problems in the same way as the human brain would (Patel, 2011) (Patel, 2016). A neural network is a computational approach, which is analogous to that of biological brain's neurons for solving a difficulty. Instead of programming the system, it is a self-learning and trained system. As human neurons adapt and learn from the surrounding slowly and gradually. Similarly, a neural network functions. Out of the various frames provided to the system certain are used for training purpose, other for validation purpose and the left for testing. The input to a neural network is always a real number and the output can be seen in a graphical manner as confusion matrix.

For this proposes algorithm the number of input nodes 240 (i.e. Number of rows in video sequence) for method 1 and 320 (i.e. Number of columns in video sequence) for method 2. The number of output nodes for both the methods is 2 to check whether flicker is present or not. The number of hidden layers used in this proposed approach is 10. Iterations are used 5000 in this experiment.

4   Experimental Setup

The proposed approach was performed on 2.30Ghz Intel core i3 processor of a 32bit operating system having RAM of 2GB, MATLAB Version-8.3.0.532 (R2014a).

5 Dataset

There was no predefined dataset available for flicker frame detection, therefore for the experiments to be performed videos were taken from YouTube and the snapshots of same are shown in Figure 5. It shows the video sequence extracted from the video where the depicted method was performed. It consists of two videos which are shown in two rows of Figure 5.
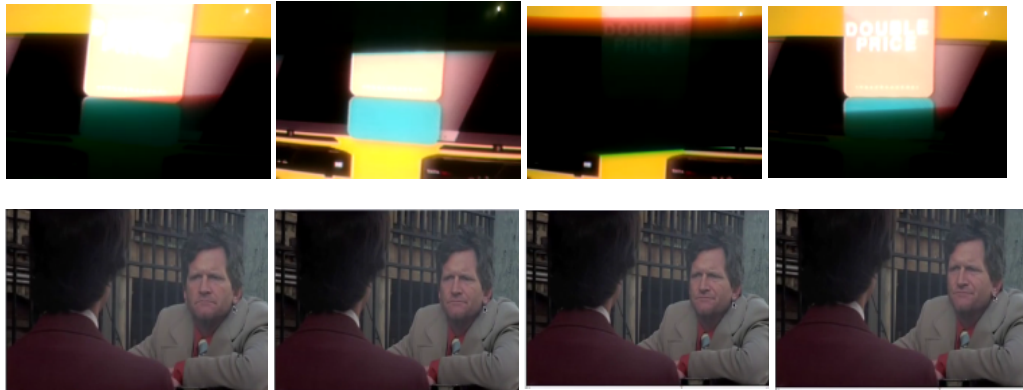

Figure 5: Video sequence from different video where flicker is present

The numbers of frames used in proposed experiment are 60 in which a minimal number of frames considered flicker. 70% of total frames used for training of the neural network and 15% of total frame used for validation stage in which the overall parameters of neural networks are calculated. The rest of the frames are used in the testing stage.

The research was performed using two different ways whose output are shown in figure 6 and figure 7. In one method the sum was taken along the row and then its DCT was taken while in other the sum was taken along its column.

6 Results

It is observed that the output was having 93.3% accuracy when the summation for the pixel values was taken along the row while for column it was 85.0%. The reason for greater accuracy in case of a row might be because of the sudden changes in value across the x-axis in coordinates than y-axis coordinates.

The confusion matrix for method 1 has been shown in Figure 6. The result shows that the proposed method give 93.3% accuracy for classifying frames whether the flicker is present or not in that particular frame.

Figure 7 on other side describes the confusion matrix for Method 2. In this case, the accuracy for detecting the flickering frame from a video sequence was 85.0%.

Figure 6: Confusion matrix for Method 1

Figure 7: Confusion matrix for method 2

7   Conclusion

In algorithm for flicker detection in the video is developed. The combination of 2-D Discrete cosine transform and an artificial neural network has been used in proposed technique. The experimental results depict that presented technique performed effectively.

In future, the method can be improved by incorporating temporal wavelet transform and support vector machine classifier.

References

Keesey, Ulker Tulunay.(1972) "Flicker and pattern detection: a comparison of thresholds." *JOSA* 62, no. 3: 446-448.

Jang, Sung-Kyu, and Pyeong-Woo Lee.(September 15, 2015) "Method for flicker detection in image signal." U.S. Patent 9,137,424.

Nguyen, Hugh Phu, Yining Deng, Eric Francois Xavier Dujardin, and Abhinav Sinha.(January 22, 2014.) "Dct based flicker detection." U.S. Patent Application 14/161,584.

Kinrot, Uri (March 18, 2011). "Methods and Systems for Flicker Correction." U.S. Patent Application 13/051,233.

Qi, Xiaojun, and Jessica Ebert. (2009)"A computer vision based method for fire detection in color videos." *International journal of imaging* 2, no. S09: 22-34.

Töreyin, B. Uğur, Yiğithan Dedeoğlu, Uğur Güdükbay, and A. Enis Cetin.((2006)) "Computer vision based method for real-time fire and flame detection." *Pattern recognition letters* 27, no. 1: 49-58.

M. Patel, Chirag I., Sanjay Garg, Tanish Zaveri, Asim Banerjee, and Ripal Patel. (2016) "Human action recognition using fusion of features for unconstrained video sequences." *Computers & Electrical Engineering*.

Ahmed, Nasir, T_ Natarajan, and Kamisetty R. Rao. (1974) "Discrete cosine transform." *IEEE transactions on Computers* 100, no. 1: 90-93.

Patel, Chirag I., Ripal Patel, and Palak Patel. (2011) "Handwritten character recognition using neural network." *International Journal of Scientific & Engineering Research* 2, no. 5: 1-6.

Patel, Chirag I., Sanjay Garg, Tanish Zaveri, Asim Banerjee, and Ripal Patel. (2016)"Human action recognition using fusion of features for unconstrained video sequences." *Computers & Electrical Engineering*.

Chirag Patel, Ripal Patel and Palak Patel (4, April 2011)**.** Handwritten Character Recognition using Neural Network, *International Journal of Scientific and Engineering Research,* Volume 2.

Zheng, Shuai, Junge Zhang, Kaiqi Huang, Ran He, and Tieniu Tan.(2011) "Robust view transformation model for gait recognition." In *Image Processing (ICIP), 2011 18th IEEE International Conference on*, pp. 2073-2076. IEEE.

Chirag I Patel and Ripal Patel(2011) 'Goal Detection from unsupervised Video Surveillance' in proceeding of ACITY-2011, Chennai, India, and July 2011 published in ***Springer (LNCS) in communications in computer and information science (CCIS) series*.** ISSN:1865-0929

Patel, C. I., and Ripal Patel. (2011) "Gaussian mixture model based moving object detection from video sequence." In *Proceedings of the International Conference & Workshop on Emerging Trends in Technology*, pp. 698-702. ACM.

Patel, Chirag I., and Sanjay Garg.(2012) "Robust Face Detection using Fusion of Haar and Daubechies Orthogonal Wavelet Template." *International Journal of Computer Applications (0975–8887)* 46, no. 6 (2012): 38.

Thakkar, Ankit, Nimeshkumar Jivani, Jigneshkumar Padasumbiya, and Chirag I. Patel.(2013) "A new hybrid method for face recognition." In *Engineering (NUiCONE), 2013 Nirma University International Conference on*, pp. 1-9. IEEE.

Patel, Chirag I., and Ripal Patel. (2013) "Illumination invariant moving object detection." *International Journal of Computer and Electrical Engineering* 5, no. 1: 73.

Patel, Chirag I., Sanjay Garg, Tanish Zaveri, and Asim Banerjee. (2014)"Top-down and bottom-up cues based moving object detection for varied background video sequences." *Advances in Multimedia* 2014: 13.

Vijay Ukani, Sanjay Garg, Chirag I Patel, and Hetali Tank.(November 2016) Efficient Vehicle Detection and Classification for Traffic Surveillance System.International *Conference on Advanced Computing and Data Science*, Volume: 1.

Patel, Chirag I., Sanjay Garg, Tanish Zaveri, and Asim Banerjee.(2014) "Top-down and bottom-up cues based moving object detection for varied background video sequences." *Advances in Multimedia* 2014: 13.

Patel, Ripal, Chirag I. Patel, and Ankit Thakkar (2012). "Aggregate features approach for texture analysis." In *Engineering (NUiCONE), 2012 Nirma University International Conference on*, pp. 1-5. IEEE.