Fast Retinex Computation for Video Sequences

Jayanta Mukherjee and R.Parthasarathi and S.Goyal
Department of CSE, IIT Kharagpur, INDIA 721302
E-mail:jay@cse.iitkgp.ernet.in

Abstract— **In this work we propose a fast algorithm for retinex computation of a video sequence. Using traditional multi-scale retinex computation one may process each frame of the video sequence individually. But this requires large computation time as it involves convolution of individual frames with large Gaussian masks. To reduce the computation time, we have performed convolution over a single frame and used these values for the subsequent frames. We have observed the results are similar with those obtained by the individual frame processing. However, the results greatly deteriorate when the scene-changes occur in the video sequence. Hence we have also detected the scene changes and made the process adaptive.**

Keywords— **Video Processing, Color Enhancement, Retinex Theory,**

I. INTRODUCTION

In recent years a lot of interests has been shown in using the retinex theory for color enhancement and for obtaining color constancy. Retinex computation based on the center/surround model proposed by E.H. Land [1] has been used for color enhancement by several researchers [2] [3] [4] [5]. Recently Jobson, Rahman and Woodell [2] have developed computation techniques for color enhancement. Initially they used Single Scale Retinex (SSR) for achieving color constancy. Later they used Multi Scale Retinex (MSR) [3] which provides superior results to those obtained by SSR. They have also performed a post processing of the retinex computed images for obtaining good color rendition. This operation is termed by them as color restoration filtering (CRF). One can easily see that the same method could easily be extended for video enhancement by processing individual frames using the same algorithm. However the computation overhead is large as each frame requires convolution operations with large Gaussian masks. In our work we have proposed a fast algorithm for computing the retinex values for a video sequence. To reduce the computation, we have performed convolution over a single frame and used those values for the subsequent frames. We have observed the results are similar with those obtained by the individual frame processings. However, the results greatly deteriorate when the scene-changes occur in the video sequence. Hence we have also detected the scene changes and made the process adaptive. In our work we have also modified the retinex computation methodology as adopted by Jobson et al. Here for convolving an image with a Gaussian mask, the image is subjected to an isotropic diffusion process. We have also replaced the logarithmic function by a sigmoidal function. It has been observed that in our case no color restoration operation is required. This also greatly improves the speed of the algorithm.

II. RETINEX COMPUTATION

Consider a color image having pixel values in three spectral bands. Let us denote each spectral band as $I_i, i = 1, 2, 3$. The pixel value for the ith spectral band at (x, y) location is denoted as $I_i(x, y)$. The single scale retinex is given by the following expression [2]

$$R_i(x,y) = log(I_i(x,y)) - log(G(x,y) * I_i(x,y)) \tag{1}$$

where $R_i(x, y)$ is the retinex output for the ith spectral band at the (x, y) pixel location in the image space. It may be noted that '$*$' denotes the convolution operation in the above expression. $G(x, y)$ is the surround function which is a 2D Gaussian mask with the standard deviation σ (uniform in both the principal directions). The MSR output is then given by the weighted sum of the outputs of the several different SSR outputs using Gaussian masks of different values of σ. Jobson et al [3] have used three surround functions and given equal weightage to each of them for retinex computation. Typically the values of σ's taken by them are 15 ,80 and 250. As MSR computation produced a grayish appearance of the processed images, Jobson et al [3] have used a color restoration function(CRF) as a post processing of the retinex computation. Finally the values are translated and scaled to be brought under the displayable range of the display device.

III. Isotropic Diffusion and Replacement of Logarithmic Function by a Sigmoidal Function

An image is convolved with a Gaussian mask (space surrounding function $G(x,y)$ in MSR computation) when it is fed as an initial input to the solution of the 2D heat diffusion equation. For a spectral band I_i this may be written as:

$$\frac{\partial I_i}{\partial t} = w.(\frac{\partial^2 I_i}{\partial x^2} + \frac{\partial^2 I_i}{\partial y^2}), \text{where } w \text{ is a constant} \quad (2)$$

The well known discrete formulation of this equation is given below:

$$\begin{aligned} I_i^{(t+1)}(x,y) \;=\; & I_i^{(t)}(x,y) + w.(I_i^{(t)}(x-1,y) \\ & +I_i^{(t)}(x+1,y) + I_i^{(t)}(x,y-1) \\ & +I_i^{(t)}(x,y+1) - 4 * I_i^{(t)}(x,y)) \end{aligned} \quad (3)$$

The above equation shows that as the number of iterations increases the amount of smoothing will increase. In fact at the nth iteration, the resulting image is same as the convolved one with a Gaussian mask of an equivalent σ of $\sqrt{2.n}$. It may also be noted that for the convergence of the above equation the value of w should lie in between 0 and 0.25 [6]. Typically at $w = 0.25$ the equation assumes the following simple form:

$$\begin{aligned} I_i^{(t+1)}(x,y) \;=\; & \tfrac{1}{4}(I_i^{(t)}(x-1,y) + I_i^{(t)}(x+1,y) \\ & +I_i^{(t)}(x,y-1) + I_i^{(t)}(x,y+1)) \end{aligned} \quad (4)$$

Given these formulations one is capable of continuously varying the σ of the surrounding function in MSR and accumulate the resulting retinex values after each iteration of the diffusion process.

We have obtained similar results as obtained previously [3] by applying similar color restoration operations with a different set of values for the respective parameters. In our case we have kept the value of N as 10000. For lower values of N, the processed image appear more grayish and less bright. However, we have noticed that even if we increase the number of iteration to a great extent (say N=30000), the results do not improve much. We also propose here to replace the logarithmic function by a sigmoidal function, which takes the following form:

$$Sigmoid(x) = \frac{1}{(1 + e^{(-(x+\theta)/T)})} \quad (5)$$

Here the parameters θ and T denote the threshold and the scale for setting up of the activation level of the sigmoid function. The value of the threshold is taken as 128 where as scale has been fixed at 100 in our computations. We have observed that the use of sigmoidal function eliminates the necessity of color restoration operation. So the modified retinex computational model is given below:

$$\begin{aligned} R_i(x,y) \;=\; & \tfrac{1}{N}.\Sigma_{t=1}^{N}(Sigmoid(I_i^{(0)}(x,y)) - \\ & Sigmoid(I_i^{(t)}(x,y))) \end{aligned} \quad (6)$$

In the above equation N is the total number of iterations and $I_i^{(0)}(x,y)$ is the original pixel value of the ith spectral band at (x,y) pixel location. However as we find that it is not necessary to consider each and every iteration for accumulating retinex values, we have accumulated these values after a fixed interval (typically, 50 iterations in our implementation). The average of the accumulated values provide R_i for each spectral band.

IV. Fast Retinex Computation for Video Sequences

Let $I_i(x,y,k)$ denote the pixel value for ith spectral component at the location (x,y) for the kth frame in the sequence. Our objective is to compute the retinex value for each and individual frame to achieve the color constancy. It may be observed that for a single static frame the retinex computation is performed in three stages. The first stage deals with the formation of the surround and here this is done through the 2D heat diffusion equation, the second part is the excitation stage where we are using sigmoidal function instead of logarithmic function and finally the retinex computation stage or integration stage. It was found that the computation of the surround takes about 85% of the computation time in our processing. Hence in order to speed up the computation, the temporal redundancy of the video sequence is made use of in the computation of the diffusion layer. We propose to carry out diffusion only through the first frame and subsequently the respective diffused pixel values are used in the successive frames. We have also observed that instead of the first frame one may consider the middle frame as the base for the retinex computation. The performance is improved for the later strategy. That is why in our implementation we have considered the middle frame (i.e. $\frac{M}{2}$th frame) as our base frame in the computation.

One may consider different strategies for defining the group of frames for which retinex computation is performed based on the diffusion over a single frame (which is also included in the group). These groups can be pre partitioned by a fixed number of frames or they may be formed by detecting the scene changes so that each video segment form a group. The first one is referred here as *fixed frame strategy* and while the later one as *adaptive strategy.* It may be noted that we have used the Twin Comparison method [7] for detecting the scene changes.

V. Results

We have carried out our experimentation's on a Silicon Graphics Platform (Octane with MIPS R10000 processor, 175 MHz clock, 128 MB RAM and IRIX 64 release 6.4 OS). Individual frames are of sizes of 240 × 180 pixels. It is interesting to note that the computation time reduces to a great extent (almost hyperbolically) as we increase the number of frames in our group (refer Figure 1). But if there is any scene change the quality of the retinex constructed images are degraded greatly. This we can observe in the Figure 2. Here we have plotted the root mean square (RMS) error between the retinex images formed by individual processing and the group processing respectively. It may be noted that the results of Figure 2 are obtained by processing a movie clip named as "CAVE.AVI" (obtained from *http://www.compuware.com*) from its 26th frame to 125th frame. We can also observe the improvement in the quality if we adaptively form the groups by detecting the scene changes (Figure 2). It may be noted that the overhead of scene change detection is negligible compared to the computation of the diffused pixel values. A typical example of retinex processed images are given in Figure 3 (a) to (d). The results of individually processed retinex images and those obtained from adaptive grouping strategy are almost similar. But, the later one takes significantly less time (almost one third of the individually processed one) for processing.

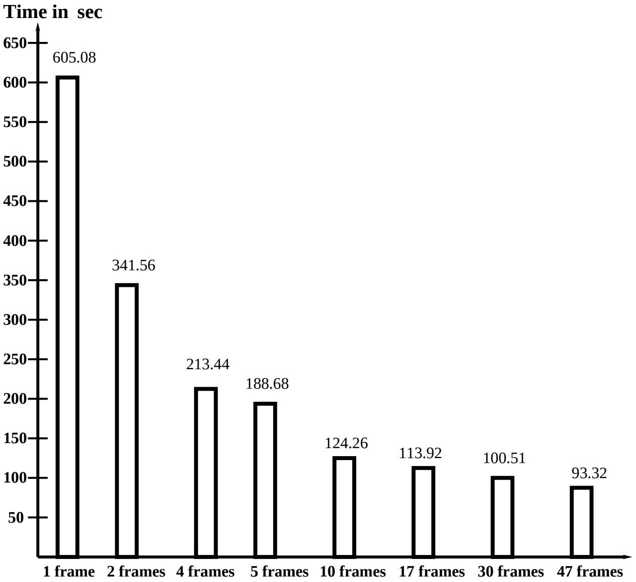

Figure 1. Average computation time vs. Number of frames

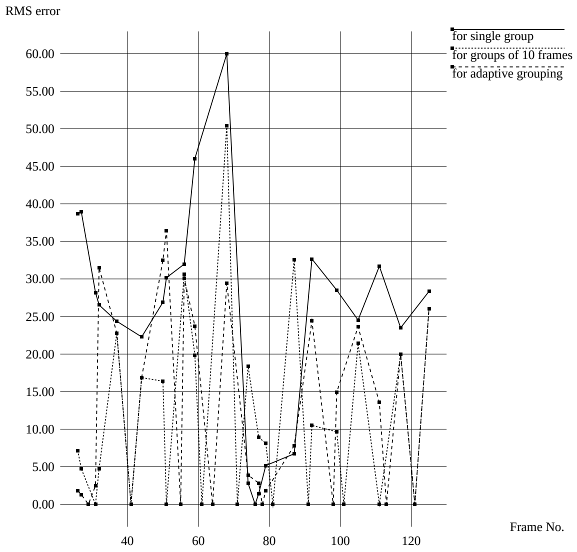

Figure 2. RMS error vs. Processing Strategies

(a) original

(b) Individually Processed

(c) Processed with a single group

(d) Processed with adaptive grouping

Figure 3. A Typical Example of Retinex Processing

VI. Conclusion

In our work we have demonstrated how the retinex computation for the video sequences can be enhanced exploiting the temporal redundancy present in the sequence. Here we have processed a group of frames instead of individual processing. We have observed that similar results are obtained even in the case of group processing. But it has the advantage of reducing the computation time to a great extent which is typically one third of the time taken for the individual processing.

References

[1] Edwin H. Land, "Recent advances in retinex theory.," *Vision Research*, vol. 26, pp. 7–21, 1986.

[2] Daniel J. Jobson, Z.Rahman, and Glenn A. Woodell, "Properties and performance of a center/surround retinex," *IEEE Transactions on Image Processing*, vol. 6, no. 3, pp. 451 – 462, March 1997.

[3] Daniel J. Jobson, Z.Rahman, and Glenn A. Woodell, "A multiscale retinex for bridging the gap between color images and the human observation of scenes," *IEEE Transactions on Image Processing: Special Issue on Color Processing*, vol. 6, no. 7, pp. 965 – 976, July 1997.

[4] A. C. Hurlbert, "Formal connections between lightness algorithms," *Journal of the Optical Society of America A*, vol. 3, no. 10, pp. 1684 – 1693, Oct. 1986.

[5] Kobus Barnard and Brian Funt, "Investigations into multiscale retinex," *Colour Imaging in Multimedia '98, Derby, UK*, pp. 9–17, March 1998.

[6] Harry Wechsler, *Computational Vision*, Academic Press, Inc., 1990.

[7] H.J.Zhang, A.Kankanhalli, and S.W.Smoliar, "Automatic partitioning of full-motion video," *ACM/Springer Multimedia Systems*, vol. 1, no. 1, pp. 10–28, 1993.