# Flicker Assessment of Low-to-Medium Frame-rate Binary Video Halftones


Hamood-Ur Rehman and Brian L. Evans
*Department of Electrical and Computer Engineering*
*Wireless Networking and Communications Group*
*The University of Texas at Austin, Austin, Texas, USA*
{*rehman, bevans*}*@ece.utexas.edu*



***Abstract*—Video display on devices with limited bit-depth capabilities requires conversion of the original (continuous-tone) video having a higher bit-depth to one having a lower bit-depth. Video halftoning refers to a process that attempts to perform this conversion such that the resulting halftone video is perceptually similar to the continuous-tone video. This quantization process, however, results in artifacts. A video halftoning algorithm can be assessed by quantifying the artifacts introduced by it. In this paper, we discuss an approach to assess flicker, a typical temporal artifact, in low-to-medium frame-rate binary video halftones produced from grayscale continuous-tone videos.**



***Keywords*-video halftoning, flicker assessment, temporal artifact.**


## I. Introduction

Display devices need to reduce the bit-depth of a video, if the original bit-depth of the video data is higher than the bit-depth supported by the display device. Halftoning is a process to perform this bit-depth reduction. The original full bit-depth video is called the continuous-tone video and the reduced bit-depth video is called the halftone video. Halftone videos can have both spatial and temporal artifacts. Flicker is a temporal artifact that is typically observed in binary halftone videos produced from grayscale continuous-tone videos. It is possible that the pixel grayscale values within a region of a continuous-tone frame do not significantly change in the next continuous-tone frame. If such is the case, flicker can usually be observed, if the corresponding halftone pixels (in the *same* region) toggle values between the (successive) halftone frames. Halftone flicker is referred to as high frequency temporal noise in [1].

The type of display can have an impact on the perception of halftone flicker. For example, on many Liquid Crystal Display (LCD) screens, halftone flicker appears as full field flicker. On Cathode Ray Tube (CRT) displays, halftone flicker may appear as scintillations. Halftone flicker can be undesirable for several reasons. It could cause discomfort to the viewer's eyes [2], and hence degrade his or her video viewing experience. On some devices, reduction of flicker might correspond to reduction of power consumption [3]. Flicker reduction could also facilitate efficient compression of halftone videos [1], [2]. For one or more of these reasons, researchers have proposed video halftoning algorithms that attempt to reduce flicker in halftone videos [1], [2], [4], [5].

Successful design of flicker reduction algorithms depends on appropriate evaluation of flicker. An important factor that affects the perception of flicker is the frame-rate at which the halftone video is viewed. At low-to-medium frame rates (30 frames-per-second or less), flicker between successive halftone frames will correspond to temporal frequencies at which the human visual system (HVS) is sensitive [6]. Evaluation of perceived flicker at these frame rates is critical in determining the quality of halftone videos. In this paper, we present a technique that attempts to improve upon exisiting halftone flicker assessment techniques. Our proposed assessment technique is designed to assess *perceived* flicker in low-to-medium frame-rate halftone videos.

The rest of the paper is organized as follows. Commonly reported halftone flicker assessment techniques are discussed in Section II. Section III presents the proposed technique. An implementation of our proposed technique along with the results is discussed in Section IV. The paper concludes with a summary of our findings in Section V.

## II. Halftone Flicker Assessment

Flicker assessment has typically been done by evaluating difference images [1], [5]. The absolute pixel-by-pixel difference between two successive halftone frames is evaluated. The "on" pixels in the resulting binary image, the difference image, show locations where pixels toggled their values. This difference image represents the flicker between the successive frames used to form the difference image. For example, Figure 1 illustrates flicker in two successive frames of a halftone video.

The difference image technique, discussed above, has a few drawbacks. This technique is feasible for evaluating flicker, if only a few difference images are to be looked at. This might be the case, if the flicker performance on these (few) difference images is considered to be representative of the entire video. However, if the entire video is to be evaluated for the actual flicker present, this technique will prove to be not feasible, especially for videos with large number of frames. Since this technique requires visual


This research was supported by an equipment donation by Intel Corp.

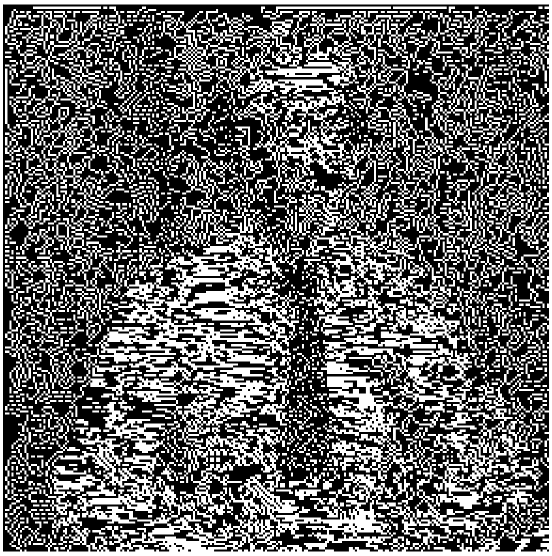

Figure 1. Absolute value of the difference of the Floyd-Steinberg error diffusion halftones for frames 36 and 37 of the Trevor sequence. A white dot indicates a toggling from either black to white or white to black at that pixel. A black dot indicates that no toggling occurred. Note the significant number of toggled bits in the relatively constant background.

inspection of the difference image, it is not entirely objective. This technique is also prone to false positives in the sense that at a scene change, although the binary patterns of successive halftone frames are expected to be quite different due to the scene change, the difference image would report higher flicker. Also, at scene changes, the perception of flicker might be lower due to the temporal masking effects of the HVS [7].

A method based on the difference image technique that evaluates flicker for the entire sequence has been used in [2]. This method computes average flicker per adjacent pair of halftone frames by adding the "on" pixels in their absolute difference image and then dividing the resulting sum by the total number of pixels in a frame. This measure, called average flicker rate (AFR), is calculated for all adjacent pairs of halftone frames. AFR plotted as a function of frame number gives the flicker performance of the entire video in a single plot. Thus, it is better than the difference image approach. However, since the measure is based solely on the difference images, it still suffers from most of the shortcomings of the difference image approach: it can give false positives, and is not a perceptual measure. Figure 2 shows the AFR, as dotted line (top curve), for the concatenated video halftone produced by independently halftoning each frame using Floyd-Steinberg [8] error diffusion. The sequence was formed by concatenating 30 frames from each of the following five standard sequences: Table Tennis, Garden, Susie, Cage, Caltrain.

## III. Proposed Technique

If flicker is present in a halftone video, its perception will depend on temporal as well as spatial characteristics of the video. Halftone flicker is thus a local phenonmenon. For the case of halftone images, it has been observed by Ulichney [9] that the nature of dither is most important in the flat regions of an image. In our observation on perception of flicker, this holds true for low-to-medium frame-rate halftone videos as well. We have observed that perception of flicker is higher in spatial regions that stay relatively constant temporally. It is even higher if such a region also has relatively flat graylevel. This implies that low spatial frequency regions exhibit higher flicker compared to high spatial frequency regions. This observation is consistent with spatial masking effects that typically reduce the perception of noise present in high frequency and textured regions of an image. To incorporate these observations into the evaluation of perceptual flicker present in a halftone video, we propose a new approach to evaluate halftone flicker. In the evaluation of flicker in a halftone video, our method utilizes the corresponding continuous-tone video. Thus, our proposed measure is a full-reference (FR) evaluation measure. To facilitate the clarity of presentation, we introduce the following notation:

- $V_c$ = continuous-tone (contone) video;
- $V_d$ = the corresponding halftone video;
- $C_i$ = $i^{th}$ frame of $V_c$;
- $C_i\,(m,n)$ = pixel located at $m^{th}$ row and $n^{th}$ column of the contone frame $C_i$;
- $D_i$ = $i^{th}$ frame of $V_d$;
- $D_i\,(m,n)$ = pixel located at $m^{th}$ row and $n^{th}$ column of the halftone frame $D_i$;
- $C_{s,i,j}(m,n)$ = local similarity measure between contone frames $C_i$ and $C_j$ at pixel location $(m,n)$;
- $C_{s,i,j}$ = similarity map/image between contone frames $C_i$ and $C_j$;
- $D_{d,i,j}(m,n)$ = local dissimilarity measure between halftone frames $D_i$ and $D_j$ at pixel location $(m,n)$;
- $D_{d,i,j}$ = dissimilarity map/image between halftone frames $D_i$ and $D_j$;
- $W_i(m,n)$ = local frequency measure at pixel location $(m,n)$ in the $i^{th}$ continuous-tone frame;
- $W_i$ = spatial frequency map/image of $C_i$;
- $F_i(m,n)$ = local perceived flicker measure at pixel location $(m,n)$ in the $i^{th}$ halftone frame ($i \geq 2$);
- $F_i$ = perceived flicker map/image at the $i^{th}$ halftone frame ($i \geq 2$);
- $\widehat{F_i}$ = perceived average flicker observed at the $i^{th}$ halftone frame ($i \geq 2$).

Based on the notation introduced above, we now propose a framework for the evaluation of perceived flicker. Let $I$ be the total number of frames in $V_c$. Let $M$ be the total number of pixel rows in each frame of $V_c$, and $N$ be the total number of pixel columns in each frame of $V_c$. Based on our discussion on flicker, earlier in this section, we note that $F_i(m,n)$ is a function of $C_{s,i,i-1}(m,n)$, $D_{d,i,i-1}(m,n)$, and $W_i(m,n)$. Therefore,

$$F_i(m,n) = f(C_{s,i,i-1}(m,n), D_{d,i,i-1}(m,n), W_i(m,n)) \tag{1}$$

For the $i^{th}$ halftone frame, we also define Perceived Average Flicker as:

$$\widehat{F_i} = \frac{\sum_m \sum_n F_i(m,n)}{M \cdot N} \tag{2}$$

$\widehat{F_i}$ is analogous to AFR of [2], discussed in Section II. However, the Perceived Average Flicker $\widehat{F_i}$ aims to evaluate *perceptual* flicker. Perceptual Flicker Index $F$ of a halftone video $V_d$ is defined as:

$$F = \frac{\sum_i \widehat{F_i}}{(I-1)} \tag{3}$$

Perceived Average Flicker $\widehat{F_i}$ can be plotted (against frame number) to evaluate flicker performance of individual halftone frames. Perceptual Flicker Index $F$ gives a single number to evaluate flicker of the entire halftone video. In the next section, we present a particular instantiation of the framework introduced in this section.

## IV. Experimental Results

Note that $F_i(m,n)$, $C_{s,i,i-1}(m,n)$, $D_{d,i,i-1}(m,n)$, and $W_i(m,n)$ constitute the maps/images $F_i$, $C_{s,i,i-1}$, $D_{d,i,i-1}$, and $W_i$ respectively. Therefore, to evaluate $F_i(m,n)$ in (1), we need the spatial frequency map of $C_i$, $W_i$, similarity map between contone frames $C_i$ and $C_{i-1}$, $C_{s,i,i-1}$, and the dissimilarity map between the successive halftone frames $D_i$ and $D_{i-1}$, $D_{d,i,i-1}$. In this particular instantiation of our proposed framework, we set $C_{s,i,i-1}$ to be the Structural Similarity (SSIM) Index Map [10] evaluated between the continuous-tone frames $C_i$ and $C_{i-1}$. We will denote it by $SSIM\{C_i, C_{i-1}\}$. We scale $SSIM\{C_i, C_{i-1}\}$ to have its pixels take values between 0 and 1 inclusive. For the dissimilarity map, we set

$$D_{d,i,i-1} = (|D_i - D_{i-1}|) * \widetilde{p} \tag{4}$$

where $\widetilde{p}$ represents the point spread function (PSF) of the HVS. We are assuming that the HVS can be represented by a linear shift-invariant system [11] represented by $\widetilde{p}$. Note that the pixel values of the map $D_{d,i,i-1}$ are between 0 and 1 inclusive. To evaluate $W_i$, we set each of its pixels as:

$$W_i(m,n) = \frac{\sum_k HPF(\mathbf{F}\{K\})}{K} \tag{5}$$

where $\mathbf{F}\{K\}$ represents the Fourier transform of a window of *K* pixels around and including $W_i(m,n)$. *HPF* refers to a high pass filtering operation. $W_i$ is normalized to have pixel values between 0 and 1 inclusive. Now, with these maps defined, we define (1) as:

$$F_i(m,n) = SSIM\{C_i, C_{i-1}\}(m,n) \cdot D_{d,i,i-1}(m,n) \cdot (1 - W_i(m,n)) \tag{6}$$

Note that $F_i(m,n) \in [0,1]$. We now analyze the form of $F_i(m,n)$ to see how it reflects our observations on perceived flicker. $F_i(m,n)$ is a product of three terms. At pixel location $(m,n)$, the first term measures the local similarity between the successive continous-tone frames. A higher value of the first term, $SSIM\{C_i, C_{i-1}\}(m,n)$, will mean that the successive frames have a higher structural similarity in a local neighborhood of pixels centered at pixel location $(m,n)$. This will in turn assign a higher weight to any flicker observed. This is desired because if the "local" scene does not change, perception of any flicker would be higher. The second term, $D_{d,i,i-1}(m,n)$, depends on the number of pixels that toggled in a neighborhood around (and including) pixel location $(m,n)$. It gives us a measure of perceived flicker due to HVS filtering. Since the HVS is modelled as a low pass filter in this experiment, $D_{d,i,i-1}(m,n)$ will have a higher value, if the pixel toggles form a cluster as opposed to being dispersed. This is aligned with our observations on flicker perception. The third term, $(1 - W_i(m,n))$, measures the *low* frequency content in a local neighborhood centered at $C_i(m,n)$. A higher value of this term will result in higher value of perceived flicker. Finally, we incorporate the effect of scene changes by setting $SSIM\{C_i, C_{i-1}\}(m,n)$ to a low value (*zero* in this experiment), if a scene change is detected between continuous-tone frames $C_{i-1}$ and $C_i$. This, in turn, reduces the value of $F_i(m,n)$ at scene cuts lowering the value of perceptual flicker to account for temporal masking effects. In this work, we utlized $SSIM\{C_i, C_{i-1}\}$ for scene cut detection. Note that between successive continuous-tone frames $C_{i-1}$ and $C_i$, a very low *average* value of $SSIM\{C_i, C_{i-1}\}$ can indicate a change of scene.

### *A. Discussion of Results*

Since a window of size 11x11 was used to perform processing to form $F_i$, we ignore 5 rows and 5 columns on the boundaries of $F_i$ to avoid boundary artifacts due to padding done during processing steps (discussed above). We now discuss our results depicted in Figure 2 as the solid line (the bottom curve). Concatenated sequence discussed in Section II was evaluated for flicker. Using (6), we calculate and plot Perceived Average Flicker, $\widehat{F_i}$, as a function of frame no. In Figure 2, scene changes are indicated on the horizontal axis using arrows. Observe that the solid line curve is not merely a scaled down version of the dotted line curve that depicts the AFR measure discussed in Section II. Note that the proposed measure (solid line) reports zero perceived average flicker at scene changes, while the one used in [2] shows distinct peaks (very high flicker) at scene changes. As already discussed in Section II, at a scene change the binary pattern of the halftone is expected to change, and this change does not imply higher perceived flicker. The standard sequences Garden (segment 2) and Caltrain (segment 5) both have high spatial frequency content in each of their constituent frames and there is also translational motion. The proposed measure shows lowest flicker for these sequences, as opposed to the dotted curve that reports the Cage (segment 4) sequence to have the lowest flicker. The Cage sequence has some rotational motion and the background in the video is relatively flat (low spatial frequency). For the Cage video segment, the observed flicker is very high owing to the large low spatial frequency content in the video. Although the average number of pixel toggles per frame pair as measured by AFR (dotted curve) is the lowest for Cage video, the proposed measure (solid line) takes low spatial frequency content into consideration and accordingly reports a flicker rate higher than that of Caltrain and Garden sequences. Susie sequence (segment 3) has the largest number of pixel toggles (as depicted by the dotted line) and relatively low frequency content. This sequence segment has the highest flicker. The concatenated sequence was viewed indoors under normal lighting conditions on a Dell U2410 LCD monitor. The sequence was viewed at 30 frames per second (fps). We found the performance of our measure reasonably consistent with our subjective observations.

## V. Conclusion

In this paper, we proposed a *perceptual* evaluation measure of flicker in binary halftone videos played back at low-to-medium frame rates. An instantiation of the proposed framework for objective evaluation of perceived flicker was discussed to evaluate the validity of the proposed framework. No attempt was made to optimally implement the framework. The framework is general and, therefore, different instantiations are possible. While the framework has been designed considering some perceptual mechanisms, more rigorous testing is recommended to establish its validity. Further testing might also help in further improving the proposed measure.

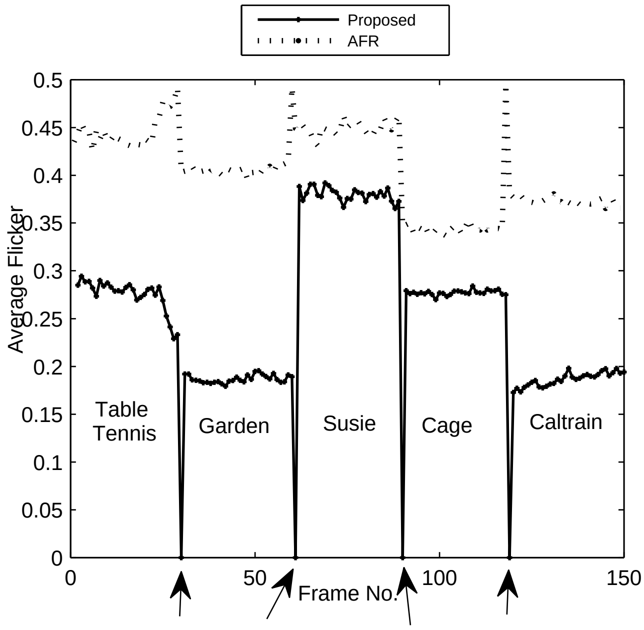


Figure 2. Flicker performance of the concatenated sequence. Average Flicker Rate (AFR), as computed by Hsu *et al.* 2007 [2] is shown as dotted line (top curve). Perceived Average Flicker, the proposed measure, is shown as solid line. Scene changes are indicated by arrows below the horizontal axis.

## References


[1] C. Gotsman, "Halftoning of image sequences," *The Visual Computer*, vol. 9, no. 5, pp. 255–266, 1993.

[2] C. Hsu, C. Lu, and S. Pei, "Video halftoning preserving temporal consistency," in *Proc. IEEE Int. Conf. on Multimedia and Expo*, 2007, pp. 1938–1941.

[3] ——, "Power-scalable multi-layer halftone video display for electronic paper," in *Proc. IEEE International Conference on Multimedia and Expo*, 2008, pp. 1445–1448.

[4] H. Hild and M. Pins, "A 3-d error diffusion dither algorithm for half-tone animation on bitmap screens," in *State-of-the-Art in Computer Animation*. Berlin, Germany: Springer-Verlag, 1989, p. 181190.

[5] Z. Sun, "Video halftoning," *IEEE Trans. on Image Processing*, vol. 15, no. 3, pp. 678–686, 2006.

[6] J. Robson, "Spatial and temporal contrast-sensitivity functions of the visual system," *Journal of the Optical Society of America*, vol. 56, no. 8, pp. 1141–1142, 1966.

[7] W. J. Tam, L. B. Stelmach, L. Wang, D. Lauzon, and P. Gray, "Visual masking at video scene cuts," in *Proc. SPIE Conf. on Human Vision, Visual Processing, and Digital Display*, vol. 2411, no. 1, 1995, pp. 111–119.

[8] R. Floyd and L. Steinberg, "An adaptive algorithm for spatial grayscale," in *Proc. SID Int. Symp., Dig. Tech. Papers*, 1976, p. 3637.

[9] R. A. Ulichney, "Review of halftoning techniques," in *Proc. SPIE Conf. on Color Imaging: Device-Independent Color, Color Hardcopy, and Graphic Arts*, vol. 3963, no. 1, 1999, pp. 378–391.

[10] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, 2004.

[11] T. Pappas, J. Allebach, and D. Neuhoff, "Model-based digital halftoning," *IEEE Signal Processing Magazine*, vol. 20, no. 4, pp. 14–27, July 2003.