

# POWER-SCALABLE MULTI-LAYER HALFTONE VIDEO DISPLAY FOR ELECTRONIC PAPER

Chao-Yung Hsu,<sup>1,2</sup> Chun-Shien Lu,<sup>1,\*</sup> and Soo-Chang Pei<sup>2</sup>

<sup>1</sup>Institute of Information Science, Academia Sinica, Taipei, Taiwan, ROC

<sup>2</sup>Graduate Institute of Communication Eng., National Taiwan University, Taipei, Taiwan, ROC

## ABSTRACT

*Video halftoning is a key technology for use in the new display device, electronic paper (e-paper). One challenging issue is how to save the limited power of mobile e-paper device when a halftone video is displayed with various frame rates. In this paper, we propose a power-scalable multi-layer halftone video display scheme, which is composed of layer coding, non-uniform sampling, and flicker rate reduction. Our method not only efficiently save power over the state of the art video halftoning technology but also keep the quality of halftone video nearly unchanged when power saving is additionally considered. Experimental results demonstrate the effectiveness of the proposed method.*

**Keywords:** Electronic paper, Flicker, Video halftoning

## 1. INTRODUCTION

### 1.1. Background

In the film, "Minority Report," which was directed by Stephen Spielberg and published by DreamWorks in 2002. There is a clip showing John Anderton (acted by Tom Cruise) boarding a train to hide himself in the crowd in order to escape from his partner. Before he enters the train, he is captured and identified by a surveillance camera. Soon, the focus of picture moves from John's face to the newspaper on a passenger who sits opposite John. As shown in Fig. 1, we can see that some columns of the newspaper are changed from "Molecular nano-technology?" and "Medical nanodevice triumphs!" (Fig. 1(a)) to "Breaking News! Precrime Hunts its Own!" (Fig. 1(b)). It was just a movie in 2002, but it is becoming real now.



**Fig. 1.** The content in e-paper can be dynamically and instantaneously changed.

The display industry is actively pursuing new display technologies to make the monitor lighter, thinner, and more portable. The electronic paper (e-paper) or smart paper [3, 11, 12] is exactly the advanced and emerging display technology, which enables one to get a digital file form from a personal computer through I/O port

and keep it on the screen without needing power. Many evidence has shown the promotion of the use of e-paper. A recent example is that Samsung has developed a color e-paper device with large screen size that can be applied on public signboard. The power consumption of the Samsung's device is 300mw per second that is only 1/500 of LCD. Compared with traditional monitors, e-paper, indeed, consumes little power and is advantageous to the environment in that no trees are hewn during the generation process.

### 1.2. Related Work

Video halftoning is a means of transferring a general video sequence into a format that can be displayed on devices with limited intensity resolutions and color palettes. In the literature, only few video halftoning techniques have been discussed. A three-dimensional (3D) error diffusion algorithm that is used to mitigate the "flicker" flaw was proposed in [5]. Gotsman [4] applied an iterative image halftoning algorithm to deal with the problem of temporal flickers existing between video frames. This algorithm achieves visual results better than those obtained from classic halftoning algorithms at the expense of increasing computational load. In [1], Atkins *et al.* investigated the display of color image sequences using a model-based approach for multilevel error diffusion. Their algorithm achieves an improvement in image quality over that yielded by frame-independent quantization when the frame rate (frames per second, fps) is sufficiently high to support temporal averaging by the human visual system. In [6], a direct-binary-search algorithm was applied to 3-D error diffusion. In [10], Sun presented a motion adaptive gain control-based 3-D error diffusion method to enhance the temporal consistency of the visual patterns by minimizing the flicker flaws. However, flicker flaws are still not efficiently reduced by means of the existing methods. Recently, we have proposed a new video halftoning method [7], which is composed of spatial error diffusion and inter-frame reference error diffusion. When compared with traditional 2D and 3D error diffusion techniques, experimental results show that our method can significantly reduce the average flicker rates under the constraint of comparable video quality.

### 1.3. Motivation

Once flicker reduction is achieved, this will lead to efficient mitigation of the probability of state change (e.g., power consumption) and the extensive use of e-paper in mobile environments, where power supply is limited and batteries are common power source, becomes possible. Under the constraint of power supply, the frame rate of halftone video to be displayed on e-paper should be properly adjusted no matter whether the higher frame rate is

\*Corresponding Author: Dr. C. S. Lu (lcs@iis.sinica.edu.tw)

supported by the architecture of e-paper or not. In view of this, we aim to propose a power-scalable multi-layer halftone video display technique that provides flexible video display with different frame rates on e-paper. We study and compare uniform sampling and non-uniform sampling of original halftone video for display with different frame rates. The halftone video is generated using our previous video halftoning method [7], which achieves temporal consistency with efficient flicker reduction. Power saving is measured in terms of number of flickers per second since power is consumed when halftone values are changed. To our knowledge, this is still a rather unexplored field.

## 2. VIDEO HALFTONING PRESERVING TEMPORAL CONSISTENCY

### 2.1. Flicker Flaws in Video Halftoning

A general video halftoning method consists of spatial error diffusion and temporal error diffusion, both of which create the flicker phenomenon. Flicker means the change of halftone values (either from black to white or from white to black) during the display of consecutive video frames that will be easily perceived by human eyes. For temporal error diffusion, this procedure will cause the pixels located at the same positions of neighboring video frames to have different halftone values due to the introduction of diffused temporal errors in particular when the pixels have the same or similar gray values. For spatial error diffusion, the diffused spatial errors will affect the halftoning results of the subsequent gray-scale pixels. If the area of gray-scale pixels located at the same positions of neighboring video frames is affected by different diffused spatial errors, then the resultant halftone values may be different, leading to flicker flaws. This situation occurs with higher probabilities in particular with pixels with gray-scale values close to the quantization threshold (e.g., 128) of halftoning. In our study, the flicker flaws caused by either temporal error diffusion or spatial error diffusion are called illegal flicker, which must be eliminated.

### 2.2. GOP-based Inter-Frame Reference Video Halftoning

In [7], we presented a new video halftoning method, which is composed of intra-frame error diffusion and inter-frame reference, to reduce flicker flaws in halftone video. In our method, the video frames are divided into many groups of pictures (GOPs) for video halftoning. Each GOP consists of an I frame and a number of P frames, where P frames refer to I frame during the halftoning process. During halftoning, the first video frame,  $I(x,y,z1)$ , is halftoned by spatial error diffusion [4] to generate the first halftone frame  $H(x,y,z1)$ . To reduce the flicker flaws, the remaining halftone video frames are not generated by means of spatial error diffusion. On the contrary, each of them is compared with its previous frame during the subsequent halftoning process. For example, the second frame  $I(x,y,z2)$  is compared with the first one for inter-frame error diffusion. If the gray level difference between two pixels located at the same position is smaller than a so-called flicker sensitivity threshold, then the halftone value of the pixel in the second frame is assigned to be the same as that in the first frame; otherwise, the halftone value is determined via a quantization process in error diffusion. No matter which condition occurs, the quantization errors still need to be diffused. In fact, the former is similar in principle to block-based compression with skip mode for bit-rate saving. Therefore, our halftoning processing also has potential to

benefit halftone video compression. In the proposed video halftoning method, video frames are halftoned sequentially.

## 3. PROPOSED METHOD

In this section, a power-scalable multi-layer halftone video display scheme is proposed for display over e-paper. In order to serve to display under different frame rates, the original halftone video obtained using [7] is divided into 1 base layer and several enhancement layers. Basically, the base layer contains the most important information that must be displayed first while the enhancement layers are used to improve video quality if higher frame rate is permitted. The idea is similar to traditional layer video coding in that the base layer and more/fewer enhancement layers are incorporated for display if high/low frame rate is considered in an e-paper device.

An intuitive method for arranging the halftone video frames into a multi-layer structure is to apply uniform sampling. More specifically, a video sequence is divided into many segments of the same size. Next, all the first frames in the intervals are collected to form the base layer, all the second frames in the segments are collected to form the first enhancement layer, and so on. Actually, uniform sampling is a simple way to achieve the goal of displaying a halftone video with different frame rates. However, uniform sampling will lead to severe flicker flaws because the relationship between neighboring frames is broken due to sampling. In addition, the base layer may not contain the most important information in a video. That is, uniform sampling produces a multi-layer structure that violate the spirit of video shot change detection in video summarization, where important video frames are preserved with higher priority.

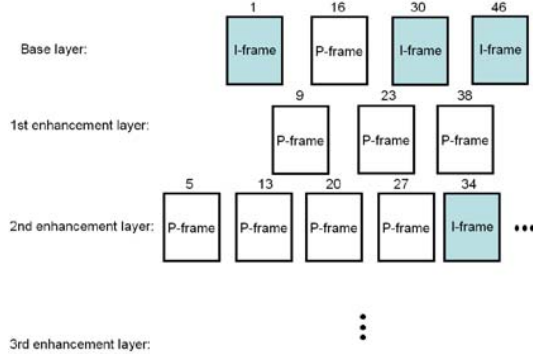
To deal with these two problems, non-uniform sampling is adopted in our method.

### 3.1. Encoding Multi-Layer Halftone Video based on Non-Uniform Sampling

In order to non-uniform sample a halftone video into one base layer and a number of enhancement layers, we take advantage of the GOP structure generated in our video halftoning method [7]. For generation of the base layer, we first uniformly set the positions of the video frames in the base layer according to the frame rate. Suppose the frame rate is  $f$  fps (frames per second). The integer multiplier of  $f$  plus 1, i.e.,  $f\alpha + 1$  (where  $\alpha \geq 0$  is an integer), will be defined as the positions of video frames in base layer. Here, the I frames of GOPs, which have positions close to  $f\alpha + 1$ , are selected as elements of the base layer for a good video summarization in a low frame rate situation. If no I frames are available, then the P frames located at the positions of  $f\alpha + 1$  are selected. An example of the base layer generation is shown in Fig. 2, where there are 60 frames in a video clip and the frame rate is 15 fps. In this example, the I frames with indices 1, 30, 46, and the p frame with index 16 are chosen to constitute the base layer. Among them, although frame 16 is a P frame, it must be chosen since there is no I frame close to position 16. However, each P frame selected to be in the base layer will be reconstructed by means of a logical XOR process because P frames refer to their corresponding I frame in a GOP during the halftoning process [7]. Under this circumstance, a P frame is only used to store the different pixels between itself and its corresponding I frame. Therefore, a P frame can be fully

reconstructed by means of logical XORing this P frame and its corresponding I frame. Since the reconstructed P frame contains the original full information, it becomes an I frame finally. Once the base layer is decided, the enhancement layers are easy to obtain, as described as follows.

The first enhancement layer is composed of the frames that are located between two neighboring frames in the base layer. The second enhancement layer is composed of the frames that are located between two neighboring frames in the first enhancement layer. This process is repeated until all frames are assigned to either a base layer or an enhancement layer.



**Fig. 2.** An example of encoding multi-layer halftone video based on non-uniform sampling.

### 3.2. Decoding Multi-Layer Halftone Video based on Frame Insertion

The decoder can decide the frame rate according to the power supply capability of e-paper. For a desk device, more enhancement layers can be included to improve the quality of displayed halftone video. For a mobile device, fewer enhancement layers must be included in order to save power consumption. Since a halftone video is encoded into a multi-layer structure, the decoding process is done by simply inserting the frames of enhancement layers back into the previous layer. For example, the frames in the first enhancement layer are first inserted back into the base layer to form a temporarily decoded halftone video. Then, the remaining enhancement layers are subsequently inserted back into the temporarily decoded video to accomplish decoding under the constraint of frame rate.

However, a problem caused by non-uniform sampling is that the number of frames in each enhancement layer is different. More specifically, the total number of frames in an enhancement layer (say  $j$ ) may be smaller than the number of intervals in the  $(j-1)$ -th enhancement layer. This problem will lead to the case where a video frame at the  $j$ -th enhancement layer cannot trivially find its location at the  $(j-1)$ -th enhancement layer. Two solutions can be used to cope with this problem.

First, if the number of frames in a lower layer (say  $j$ ) is equal to that of its corresponding upper layer (say  $j-1$ ), then the frames of the lower layer can be subsequently inserted back into the interval between two neighboring frames in the upper layer. Second, if the number of frames in a lower layer is different from or smaller than that of its corresponding upper layer, then for each frame to be inserted, video frame matching based on computationally complex

image perceptual hashing [8] or computationally uncomplex interpolation/extrapolation is adopted to choose a similar video frame from a previously decoded halftone video to accomplish insertion. Details are omitted here due to limited space.

## 4. EXPERIMENTAL RESULTS

In the experiments, a number of 24-bit color video sequences were used. Among them, the results obtained from the “Vassar” video [9], “Secretary” video, which was excerpted from the movie “The devil wears the Prada,” and “Dolphin” video, which was captured by a hand-held camera, were reported here. The three video sequences, respectively, contain small, moderate, and large motion contents. The frame size ranges from  $320 \times 240$  to  $436 \times 1024$ . The performance of our method was evaluated in terms of power consumption and halftone video quality.

### 4.1. Power Consumption

In order to evaluate whether our method can efficiently save power under various frame rates, the average power consumption measured in terms of number of flickers per second is a good indicator. In this aspect, our method and Sun’s video halftoning method [10], which can be recognized as a state-of-the-art video halftoning technology, were used for comparison. Fig. 3 shows the average power consumption for the “Vassar,” “Secretary,” and “Dolphin” video sequences, respectively, where the x axis denotes the frame rate (frames per second, fps) and the y axis denotes the power consumption (number of flickers per second).

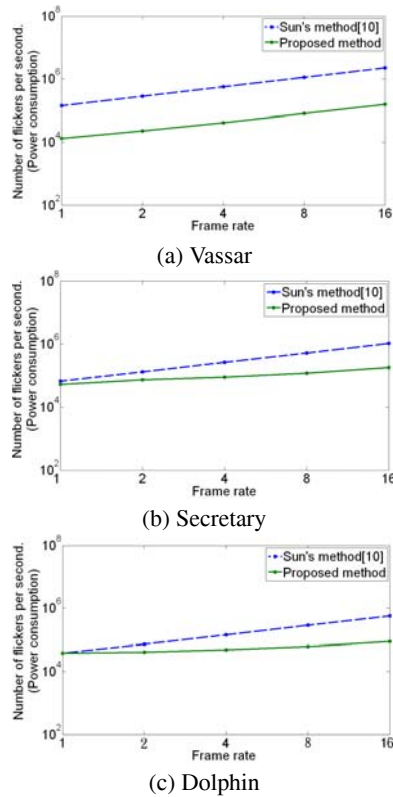
It can be observed from these results that our method consistently consumes lower power under different frame rates. We can find from Fig. 3(a) that the power consumption of our method is smaller than that of Sun’s method when the frame rate is 1 fps. The reason is that “Vassar” is a slow-motion video, which has the GOP of “IPPP...” in the original encoded video stream. Therefore, the base layer contains the only I frame and other P frames. In addition, our video halftoning method, which preserves temporal consistency, reveals its strength in reducing flickers (corresponding to reducing power consumption). On the other hand, Figs. 3(b) and (c) show that our method consumes nearly the same power as Sun’s method [10] when the frame rate is 1 fps. The reason is that the halftone videos generated from the “Secretary” and “Dolphin” video sequences, which possess moderate-to-high motions, are almost completely composed of I frames when the frame rate is 1 fps. The differences between a pair of I frames are as high as the differences between two independent halftoned frames that are generated from error diffusion. Under this condition, our video halftoning method does not get any benefit in reducing flickers. However, when frame rates are larger than 1 fps, Figs. 3(b) and (c) show that our method consumes less power than Sun’s method.

### 4.2. Halftone Video quality

To evaluate the quality of halftone video obtained using our multi-layer display method, halftone PSNR (HTPSNR) was adopted as the criterion in this paper. HTPSNR is defined between two (halftone) video frames,  $X$  and  $Y$  as:

$$HTPSNR(X, Y) = PSNR(HVS(X), HVS(Y)), \quad (1)$$

where  $HVS()$  denotes a contrast sensitivity function of the human visual system [2]. The qualities of the original non-layered



**Fig. 3.** Comparisons of power consumption for halftone videos, Vassar (a), Secretary (b), and Dolphin (c).

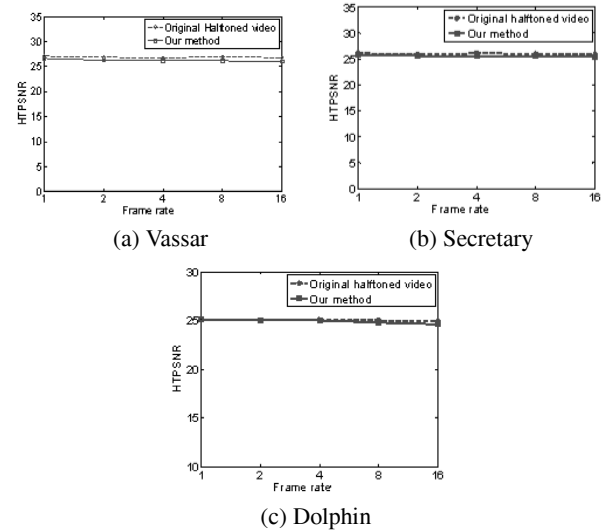
halftone video and our multi-layer halftone video were compared in terms of HTPSNR. In Eq. (1), the video frame  $X$  is either a non-layered halftone video frame or a multi-layer halftone video frame while  $Y$  is an original gray-scale video frame. The HTPSNR values measured for the three video sequences, “Vassar,” “Secretary,” and “Dolphin,” are shown in Fig. 4. We can observe from Fig. 4 that the HTPSNR values for the non-layer halftone video and our multi-layer halftone video are almost the same. This implies that our method does not degrade the quality of halftone videos.

## 5. CONCLUSION

In view of the fact that video halftoning will play an important role in supporting the use of the emerging display device, e-paper, this paper proposes a power-scalable multi-layer halftone video display technique that can support halftone video display under various frame rates. The main contribution is that our method can save more power over the state of the art video halftoning method and nearly does not introduce quality degradation. To our knowledge, this is the first work that can achieve flicker rate reduction and power saving. Future work will study halftone video compression.

## 6. REFERENCES

[1] C. B. Atkins, T. J. Flohr, D. P. Hilgenberg, C. A. Bouman, and J. P. Allebach, “Model-based color image sequence quantization,” in *Proc. SPIE/IST Conf. Human Vision, Visual*



**Fig. 4.** Comparisons of halftone video quality for halftone videos, Vassar (a), Secretary (b), and Dolphin (c).

*Processing, and Digital Display V*, Vol. 2179, pp. 310-317, Feb. 1994.

- [2] P. J. Barten, “Physical model for the contrast sensitivity of the human eye,” in *Proc. IS&T/SPIE Int. Symp. on Electronic Imaging Science and Technology*, Vol. 1666, San Jose, CA, Feb. 9-14, pp. 57-74, 1992.
- [3] G. P. Crawford, “A bright new page in portable displays,” *IEEE Spectrum*, Vol. 37, pp. 40-46, Oct. 2000.
- [4] C. Gotsman, “Halftoning of image sequence,” *Vis. Comput.*, Vol. 9, No. 5, pp. 255-266, 1993.
- [5] 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-Proceedings of Computer Animation*. Springer-Verlag, pp. 181-190, 1989.
- [6] D. P. Hilgenberg, T. J. Flohr, C. B. Atkins, J. P. Allebach, and C.A. Bouman, “Least-squares model-based video halftoning,” in *Proc.SPIE/IST Conf. Human Vision, Visual Processing, and Digital Display V*, Vol. 2179, pp. 7-10, 1994.
- [7] C. Y. Hsu, C. S. Lu, and S. C. Pei, “Video halftoning preserving temporal consistency,” *Proc. IEEE Int. Conf. on Multimedia and Expo*, Beijing, China, 2007.
- [8] C. S. Lu and C. Y. Hsu, “Geometric distortion-resilient image hashing scheme and its applications on copy detection and authentication,” *ACM Multimedia Systems Journal*, special issue on Multimedia and Security, Vol. 11, No. 2, pp. 159-173, December 2005.
- [9] Mitsubishi Electric Research Laboratories, “MERL multiview video sequences,” <ftp://ftp.merl.com/pub/avetro/mvctestseq>.
- [10] Z. Sun, “Video halftoning,” *IEEE Trans. on Image Processing*, Vol. 15, No. 3, pp. 678- 686, 2006.
- [11] <http://www.eink.com>
- [12] <http://www.sipix.com.tw>