Journal of Electronic Imaging 1(4), 388–395 (October 1992).

Efficient fax transmission of halftone image sequences

Kevin J. Parker
Hei Tao Fung
University of Rochester
Department of Electrical Engineering
Rochester, New York 14627

Abstract. *We propose an algorithm in which a sequence of digital halftone images is efficiently transmitted using encoding compatible with conventional facsimile devices. To enhance the CCITT standardized coding schemes for Group 3 and Group 4 facsimile apparatus, pre-encoding is done on each image in the sequence: The image is either pre-encoded as a combination of bit representation of block means and an "error" image, according to the ToneFac algorithm, or it is pre-encoded as an "interimage" in which interframe redundancy is converted into spatial redundancy, and the least "busy" of the above images is encoded and transmitted. The approach is referred to as the ToneSec algorithm.*

1 Introduction

The efficient transmission of a sequence of images should take advantage of the spatial redundancy of each frame as well as the temporal redundancy of consecutive frames. Intensive work on interframe coding of continuous-tone images has been accomplished.[1] The Moving Picture Experts Group (MPEG) has set a video compression standard for multimedia applications.[2] However, less attention has been paid to the efficient transmission of halftone image sequences. For example, individuals may wish to transmit multiple image frames from one personal computer (PC) to another. If both PCs contain facsimile modem boards, conventional halftoning and transmission of each frame would be inefficient, time-consuming, and expensive if long-distance phone lines were used. Coding schemes for Group 3 and Group 4 fax devices have been standardized by the CCITT.[3–5] According to these schemes, long run-lengths and line-to-line redundancy within an image result in faster transmission. The problem is that since conventional fax printers are only capable of producing black and white dots, each gray-scale image needs to be halftoned. A halftone image is typically composed of a mosaic of black and white pixels, resulting in short run-lengths, little redundancy from line to line, and inefficient encoding. Efficiency is improved by using the ToneFac algorithm in which an image is precompressed into block means and an "error" image before being encoded using the coding standard.[6] However, the fact that temporal redundancy occurs in image sequences should prompt us to ask the following question: Is it possible to take advantage of the temporal redundancy while keeping the advantages of the CCITT coding standard?

Paper 92-024 received June 18, 1992; revised manuscript received Aug. 5, 1992; accepted for publication Oct. 7, 1992.
1017-9909/92/$2.00.

Predictive coding of interframe images has been considered by many.[7–9] Considerable gains in performance are achieved by introducing motion compensation in predictors.[2,10,11] In MPEG's compression standard, motion-compensated prediction is used to exploit the temporal redundancy of video signals.[2] Motion compensation generally requires motion estimation. Motion estimation covers a set of techniques used to extract the motion information from an image sequence.[11–14] The problems associated with the techniques are the computational burden and the extra bits required for transmitting motion vectors. Motion compensation is computationally intensive and could be costly to implement, and it is not oriented toward binary halftone images. Note also that when a gray-scale object is translated, the image features are constant but merely displaced. However, when a continuous-tone image is halftoned, the image features in the halftone depend on the image features and the halftone cell position. Thus, the motion estimation problem is different for halftone image sequences. So we attempt to find a simpler method that can send halftone image sequences and take advantage of temporal redundancy wherever possible.

2 ToneSec Algorithm

By noting that a halftone image is bilevel, if pixel-to-pixel exclusive-or (XOR) operation is performed on two halftone images, and the result is called the *interimage,* then one image can be recovered from another pixel-to-pixel XOR operation on the interimage with the other original image. (We call the preceding image in a sequence the *reference image*.) In the case of large frame-to-frame redundancy, the interimage will be mostly black (0's) with long black pixel runs and large line-to-line redundancy. In other words, frame-to-frame redundancy is transformed into spatial redundancy in a sparse interimage, which allows for the exact recovery of the original halftone image. Therefore, the CCITT schemes can be employed to encode the interimage efficiently. In the case of small frame-to-frame redundancy, it may be worse to encode the interimage than the original image. So at some points, there may be a need to switch to the ToneFac algorithm. The decision is based on some criteria.

The advantage of the proposed algorithm is that its lossless operation allows the receiver to recover the exact halftone images as rendered in the transmitter. The second advantage is its simplicity in the required temporal-to-spatial redundancy transformation. The overall approach is termed the ToneSec (half*tone se*quence *c*ompression) algorithm.

Halftoning of gray-scale images can be accomplished in various ways.[15] However, we assume that a blue noise mask (a halftone screen) is stored in transmitters and receivers in order to render rapidly a halftone image with an aperiodic and isotropic pattern.[16] Nevertheless, the following algorithm works with any halftone screen.

Transmission steps

The algorithm uses the following transmission steps (see Fig. 1):

1. Initialize $k=0$, $rl_n=rd_n=0$; determine I, J, and I', where k denotes the sequential order, I the number of rows in the original image, J the number of pixels per line, and I' the number of rows on the overall image (the error image plus the block means in cases where ToneFac is used).
2. Get a gray-scale image, $g_k(i,j)$.
3. If the picture is the first image of a sequence, i.e., $k=0$, go to step 9.
4. Halftone $g_k(i,j)$ with blue noise mask $b(i,j)$ to get the halftone image $h_k(i,j)$. For all (i,j) in J,

   $$\text{if } g_k(i,j) \geqslant b(i,j),\ h_k(i,j)=1 \ ,$$

   $$\text{if } g_k(i,j) < b(i,j),\ h_k(i,j)=0 \ .$$

5. Calculate the interimage $n(i,j)$ between $h_k(i,j)$ and $h_{k-1}(i,j)$:

   $$n(i,j)=\mathrm{XOR}(h_k(i,j),\ h_{k-1}(i,j)) \text{ for all } (i,j) \ .$$

6. Calculate the mean run-length of $n(i,j)$:

   $$rl_n = I \times J/Nr \ ,$$

   where $Nr=$ the total number of runs.
7. Calculate the line-to-line redundancy, rd_n. The line-to-line redundancy is defined as the average percentage of pixels in each line that have the same value as the preceding line.
8. If $I' \times rl_n \times rd_n \geqslant I \times rl_o \times rd_o$, then transmit $n(i,j)$ and code for interimage, and go to step 13. Step 8 compares the number of lines, run-lengths, and redundancies of the two images (interimage versus overall image) in order to select the most efficient coding.
9. According to the ToneFac algorithm, find block means $G(K)$, where $G(r)$ is the mean gray value over the r'th rectangular subregion of arbitrary size, and the error image $e(i,j)$. For all (i,j) in K,

   $$\text{if } G(K) \geqslant b(i,j),\ e(i,j)=\mathrm{NOT}(h(i,j)) \ ,$$

   $$\text{if } G(K) < b(i,j),\ e(i,j)=h(i,j) \ .$$

   Functionally this is equivalent to the definition of error image in Ref. 6, although the form of the equation is simplified here.
10. Combine the error image $e(i,j)$ and the bit representation of the block means $G(K)$ as the overall image $o(i',j)$.
11. Calculate the mean run-length, $rl_o=I' \times J/Nr_o$, where Nr_o is the total number of rows in $o(i',j)$, and line-to-line redundancy, rd_o of $o(i',j)$.
12. If $I' \times rl_n \times rd_n \geqslant I \times rl_o \times rd_o$, then transmit $n(i,j)$ and code for interimage. Otherwise, transmit $o(i',j)$ and code for ToneFac.
13. $k=k+1$.
14. If there is another image to transmit, go to step 2.
15. End.

In step 10, we assume that the bit representation of $G(K)$ is an $m \times J$ image. For an 8-bit gray-scale image, each block mean is represented by 8 bits. Therefore, for a 416×512 image and subregion of size 8×8, $m=512 \times 416 \times (8/512)/$

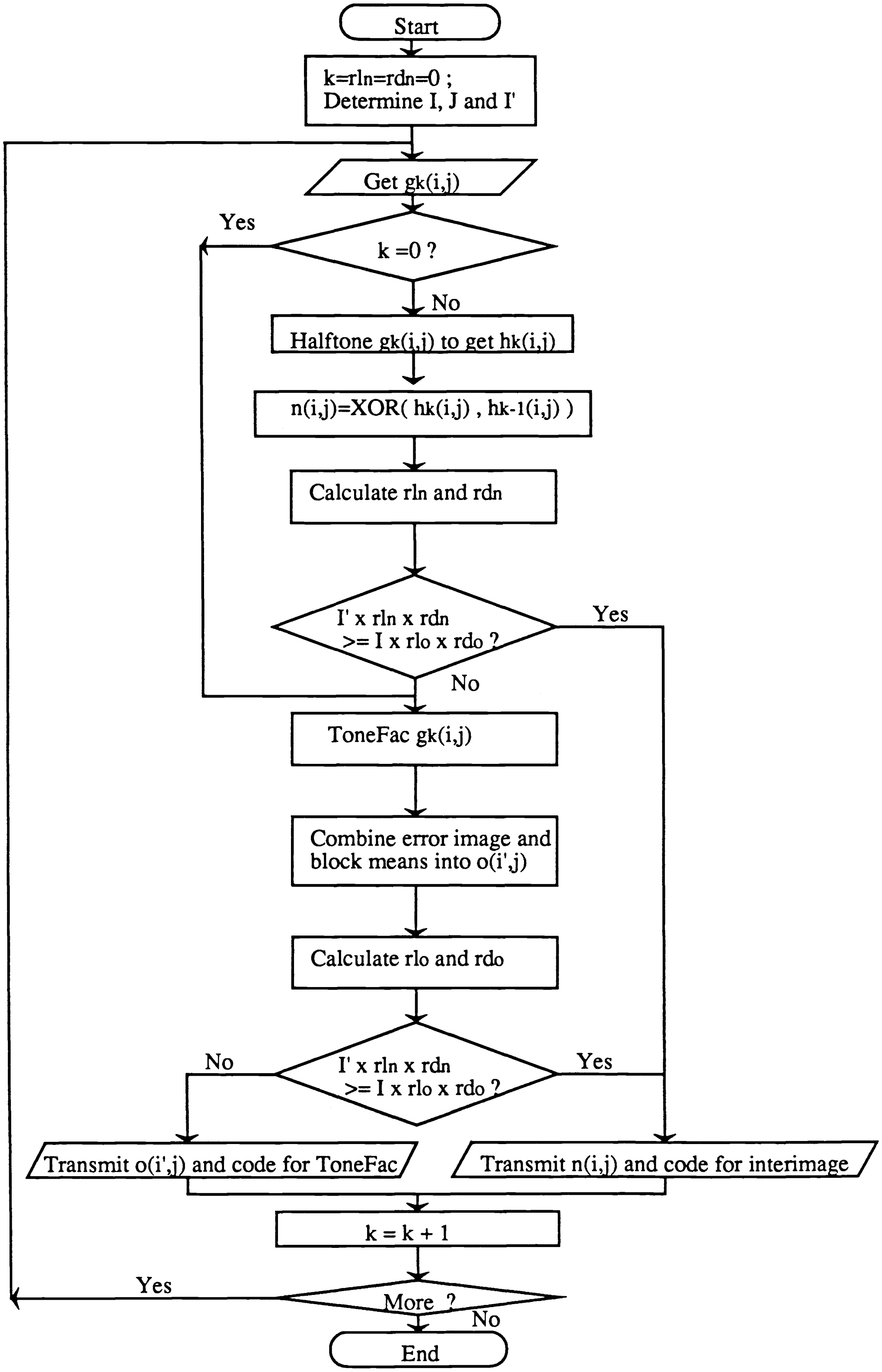

Fig. 1 Flowchart of algorithm transmission steps.

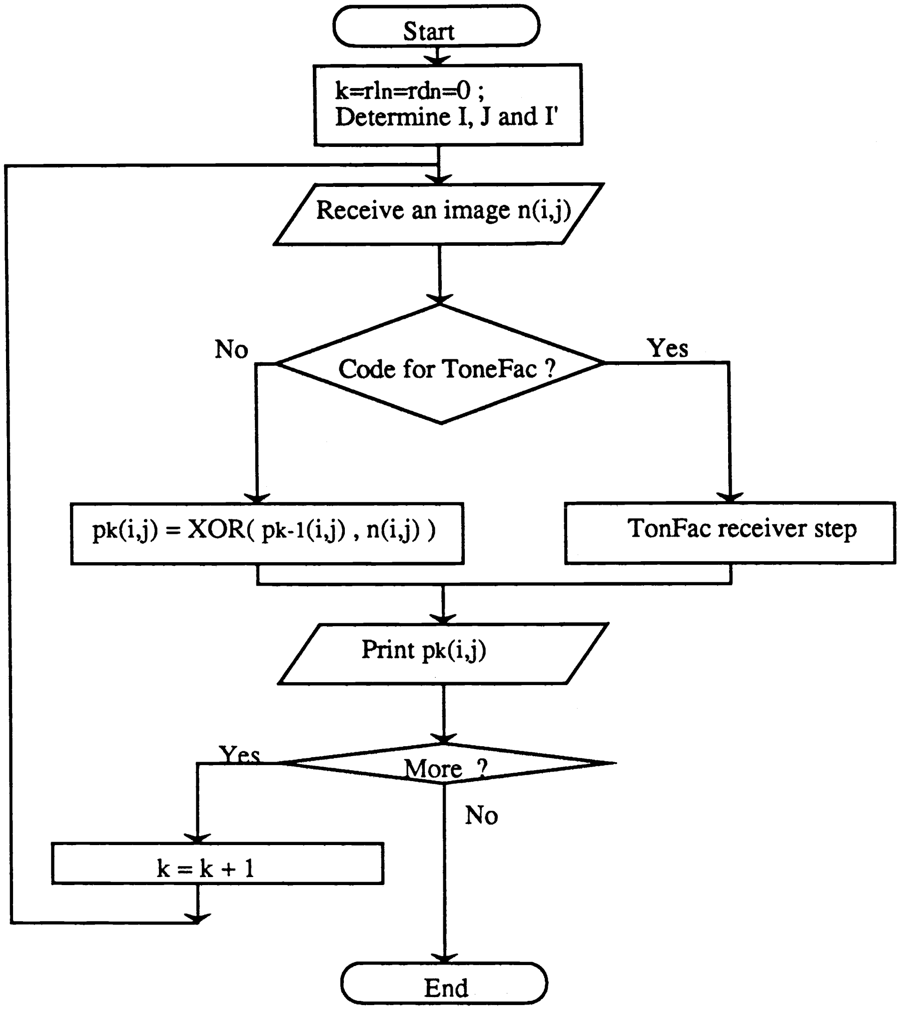

Fig. 2 Flowchart of algorithm receiving steps.

$(8\times 8)=52$. That is, 52 additional lines of data are required to transmit the block mean values in primitive form. The criteria in steps 8 and 12 are primitive where transmission efficiency is estimated to be proportional to the mean run-length and line-to-line redundancy, and the ratio of the size of the overall image to that of the interimage is taken into account.

Receiving steps

The following receiving steps are used (see Fig. 2):

1. Initialize $k=0$; determine I, J, and I'. All symbols are defined as before.
2. Receive an image $n(i,j)$ from the transmitter.
3. If code for ToneFac is received, follow the ToneFac receiver steps to recover the image $p_k(i,j)$. Decompose $n(i,j)$ into the error image $e(i,j)$ and block means $b(i,j)$. For all (i,j) in K,

   if $G(K)\geqslant b(i,j)$, $p_k(i,j)=\mathrm{NOT}(e(i,j))$,

   if $G(K)<b(i,j)$, $p_k(i,j)=e(i,j)$.

   Print $p_k(i,j)$.
4. If code for interimage is received, recover the image as follows:

   $p_k(i,j)=\mathrm{XOR}(p_{k-1}(i,j),\ n(i,j))$.

   Print $p_k(i,j)$.
5. If there is another image to receive, increment k and go to step 2.
6. End.

3 Simulations and Discussions

Three sequences of six images each are extracted from videotapes using a frame grabber. They are tailored to 416×512 pixels with 256 gray levels. The size of the subregion is 8×8. Block means can be coded differently, for example, in gray codes or their bits can be regrouped to fit the CCITT schemes. However, in this paper they are represented simply as their binary numbers. The "girl" sequence shows the movement of the girl's mouth. The first and last frames are shown in Figs. 3 and 4, respectively. The "chicken" sequence demonstrates translation. The first and last frames are shown in Figs. 5 and 6, respectively. The "teacher" sequence shows head rotation (not rotation

Fig. 3 First frame of the "girl" sequence. Six television image frames were obtained using a frame grabber and a VCR. The images were halftoned using the blue noise mask.

Fig. 5 First frame of the "chicken" halftone sequence.

Fig. 4 Sixth and last frame of the "girl" halftone sequence.

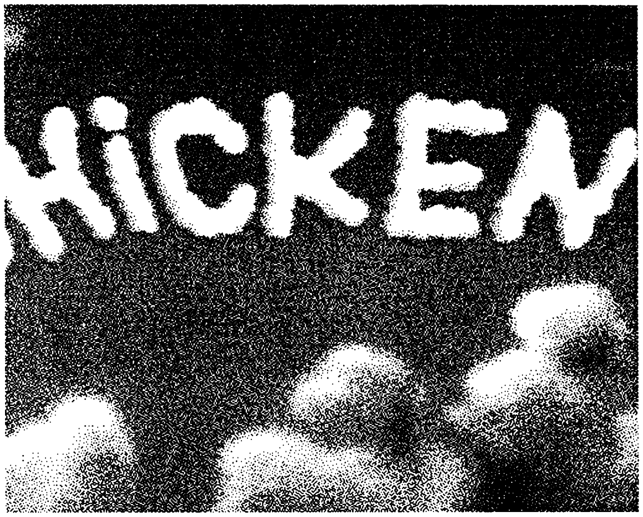

Fig. 6 Sixth and last frame of the "chicken" halftone sequence.

of a block of pixels). The first and last frames are shown in Figs. 7 and 8, respectively.

The results are given in Tables 1, 2, and 3. In the tables, "TF" stands for ToneFac, "IN" for interimage, "rl" for mean run-length, and "rd" for line-to-line redundancy. Interimages exist only if there is an image previously transmitted, so they are not available for frame 1. Table 1, the results for test image "girl," shows that it is more efficient to transmit interimages for frames 2 to 6 due to the fact that the moving area is only a "small" area (Figs. 3 and 4). However, Table 2, the results for "chicken," demonstrates that for another image, transmitting the overall images derived from the ToneFac algorithm is more efficient. Figures 5 and 6 show a rapid translation of the word "CHICKEN" over six frames. Table 3 illustrates a more complicated case. Figures 9(a) through 9(f) show the actual information chosen at each step for the "teacher" series. Note the extra lines displayed at the bottom for ToneFac steps. This information is the simple binary representation of the block means. If the criteria in steps 8 and 12 are different, the ToneSec choice may be different.

In a very primitive simulation test using transmission of the "teacher" sequence over conventional facsimile devices, we estimate that the original halftone sequence would take 424 s, whereas the ToneSec sequence requires 138 s, a compression factor of 3.

Intuitively and qualitatively, the temporal-to-spatial redundancy transformation works best for images with slow motion and rapid motion involving a small area. In contrast, motion-compensated prediction is good for images with slow rotation and translation. Once motion reaches a certain point of complexity, the images concerned are better transmitted using the ToneFac algorithm without interframe coding.

Fig. 7 First frame of the "teacher" halftone sequence.

Fig. 8 Sixth and last frame of the "teacher" halftone sequence.

Table 1 Comparison of halftone images for "girl" test image.

	Original halftone			ToneFac			Interimage			ToneSec choice
Frame #	rl(pels)	rd(%)	entropy	rl(pels)	rd(%)	entropy	rl(pels)	rd(%)	entropy	
1	4.94	79.47	0.5010	11.17	91.33	0.2848	N/A	N/A	N/A	TF
2	4.93	79.39	0.5044	10.95	91.30	0.2881	17.80	94.03	0.1900	IN
3	4.97	79.53	0.5010	11.04	91.30	0.2873	16.69	93.61	0.2011	IN
4	4.89	79.24	0.5062	10.87	91.16	0.2897	14.13	92.49	0.2307	IN
5	4.92	79.43	0.5020	11.15	91.42	0.2844	16.43	93.52	0.2055	IN
6	4.88	79.29	0.5069	11.00	91.29	0.2871	13.52	92.25	0.2416	IN

Table 2 Comparison of halftone images for "chicken" test image.

	Original halftone			ToneFac			Interimage			ToneSec choice
Frame #	rl(pels)	rd(%)	entropy	rl(pels)	rd(%)	entropy	rlpels)	rd(%)	entropy	
1	2.22	54.41	0.8336	6.54	86.80	0.4265	N/A	N/A	N/A	TF
2	2.27	55.38	0.8278	6.49	86.73	0.4277	4.57	77.87	0.5921	TF
3	2.32	56.45	0.8231	6.52	86.54	0.4271	4.32	76.52	0.6139	TF
4	2.33	56.63	0.8222	6.61	86.81	0.4264	4.16	75.50	0.6236	TF
5	2.32	56.41	0.8296	6.60	86.68	0.4281	5.69	82.08	0.4907	TF
6	2.31	56.20	0.8304	6.61	86.68	0.4298	4.02	74.79	0.6401	TF

Table 3 Comparison of halftone images for "teacher" test image.

	Original halftone			ToneFac			Interimage			ToneSec choice
Frame #	rl(pels)	rd(%)	entropy	rl(pels)	rd(%)	entropy	rl(pels)	rd(%)	entropy	
1	2.88	64.74	0.7617	8.38	89.63	0.3466	N/A	N/A	N/A	TF
2	2.93	65.23	0.7576	8.22	89.47	0.3506	8.63	87.99	0.3420	IN
3	2.91	65.06	0.7580	8.21	89.43	0.3495	8.47	87.76	0.3461	IN
4	2.96	65.56	0.7429	8.64	90.02	0.3358	7.41	86.02	0.3960	TF
5	2.89	64.58	0.7499	8.25	89.27	0.3490	7.61	86.36	0.3735	IN
6	2.94	65.35	0.7385	8.59	89.82	0.3374	8.12	87.15	0.3543	IN

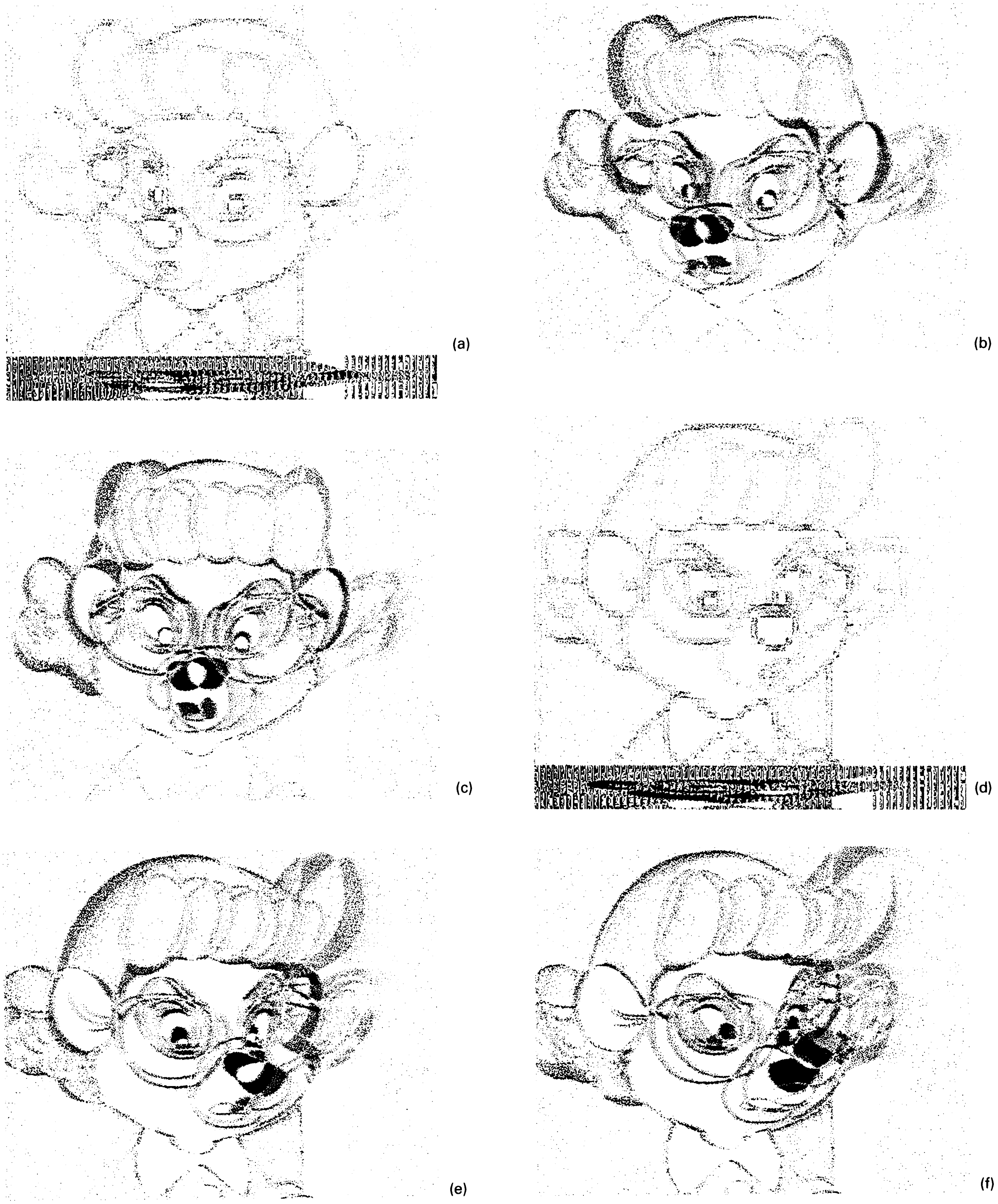

Fig. 9 The compressed information for the "teacher" halftone sequence: (a) The ToneFac information for the first frame. The extra lines at the bottom carry the simple bit representation of the block means of the original image. (b) Interframe information for the second halftone image in the sequence. (c) Interframe information for the third halftone image in the sequence. (d) A ToneFac representation of the fourth frame, where this compression was chosen instead of an interframe because of the statistics shown in Table 3. (e) Interframe information required to reconstruct the fifth frame. (f) Interframe information required to reconstruct the sixth frame.

4 Conclusion

The transmission of a sequence of halftone images can utilize conventional fax protocols while both the spatial redundancy and interframe redundancy are exploited. This paper presents the ToneSec algorithm, which combines the ToneFac algorithm with the technique by which the frame-to-frame redundancy is converted to spatial redundancy in an interimage that can be efficiently coded by the CCITT scheme. Considerable gain in transmission speed is achieved for images with slow motion or small-area rapid motion. The price paid is more computational steps for fax devices.

Acknowledgment

The authors acknowledge the helpful software tips from the graduate students in the Department of Electrical Engineering at the University of Rochester.

References

1. A. K. Jain, "Image data compression: a review," *Proc. IEEE* **69,** 349–389 (Mar. 1981).
2. D. LeGall, "MPEG: a video compression standard for multimedia applications," *Comm. ACM* **34,** 46–58 (1991).
3. R. Hunter and A. H. Robinson, "International digital facsimile coding standards," *Proc. IEEE* **68**(7), 854–867 (1980).
4. CCITT, "Recommendation t.6, facsimile coding schemes and coding control functions for Group 4 facsimile apparatus" (1984).
5. Y. Yasuda, T. Kamae, K. Kobayashi, "Advances in fax," *Proc. IEEE* **73**(4), 706–730 (Apr. 1985).
6. K. J. Parker and A. C. Cheung, "Efficient fax transmission of halftone images," *J. Electron. Imaging* **1**(2), 203–208 (Apr. 1992).
7. F. W. Mounts, "A video encoding system with conditional picture-element replenishment," *Bell Syst. Tech. J.* **48,** 2545–2554 (Sep. 1969).
8. R. F. W. Pease and J. O. Limb, "Exchange of spatial and temporal resolution in television coding," *Bell Syst. Tech. J.* **50,** 191–200 (Jan. 1971).
9. J. C. Candy et al., "Transmitting television as clusters of frame-to-frame differences," *Bell Syst. Tech. J.* **50,** 1889–1917 (Aug. 1971).
10. S. Broferrio and F. Rocca, "Interframe redundancy reduction of video signals generated by translating objects," *IEEE Trans. Comm.*, 448–455 (Apr. 1977).
11. J. R. Jain and A. K. Jain, "Displacement measurement and its application in interframe image coding," *IEEE Trans. Comm.* **29,** 1799–1808 (1981).
12. C. D. Kuglin and D. C. Hines, "The phase correlation image alignment method," *Proc. IEEE Int. Conf. Cybernetics and Society,* San Francisco, California, September 23–25, 1975.
13. C. Caffario and F. Rocca, "The differential method for image motion estimation," *IEEE Trans. Info. Theory* **IT-22,** 573–579 (Sep. 1976).
14. C. Hsieh, P. Lu, and J. Shyn, "Motion estimation using interblock correlation," *IEEE Int. Symp. Circuits and Systems,* Vol. 2, 995–998 (1990).
15. R. A. Ulichney, *Digital Halftoning,* MIT Press, Cambridge, MA (1987).
16. T. Mitsa and K. J. Parker, "Digital halftoning using a blue noise mask," *Proc. SPIE* **1452,** 47–56 (Feb. 1991).

Kevin J. Parker received the BS degree in engineering science, Summa Cum Laude, from SUNY at Buffalo in 1976. Graduate work in electrical engineering was done at MIT, with MS and PhD degrees received in 1978 and 1981. From 1981 to 1985 he was an assistant professor of electrical engineering at the University of Rochester; currently he holds the title of professor of electrical engineering and radiology. Dr. Parker has received awards from the National Institute of General Medical Sciences (1979), the Lilly Teaching Endowment (1982), the IBM Supercomputing Competition (1989), and the World Federation of Ultrasound in Medicine and Biology (1991). He is a member of the IEEE Sonics and Ultrasonics Symposium Technical Committee and serves as reviewer and consultant for a number of journals and institutions. He is also a member of the IEEE, the Acoustical Society of America, and the American Institute of Ultrasound in Medicine. Dr. Parker's research interests are in medical imaging, linear and nonlinear acoustics, and digital halftoning.

Hei Tao Fung received his BS and MS degrees in electrical engineering from the University of Rochester, New York, in 1992. His interests include image processing and integrated circuit design. Mr. Fung is a member of Tau Beta Pi and Phi Beta Kappa.