THE BENEFITS OF FREQUENCY MODULATION SCREENING

Erwin Widmer, Kurt Schläpfer, Veronika Humbel, Serdar Persiev *

Abstract

In a previous project (1) the different algorithms for frequency modulation (FM) screening have been compared and analyzed. Several algorithms were found to give on low-resolution output devices a similar image quality as conventional halftone images on high-resolution systems. The present project had the aim to apply these algorithms to B&W images on laser printers with 600 dpi and to four color printing. The results confirm that FM screening is able to be an alternative to conventional screening, especially if low-resolution output systems are being used. A computer program was written to apply FM screening algorithms to Macintosh computers.

1. Introduction

The first publication on frequency modulation screening in the graphic arts has been published by the Technical University of Darmstadt in 1983 (2). To be accurate, it should be added that the principles of the process have already earlier been described in the computer technology (3). The term "frequency modulation", however, was new and hitherto only used in communication technology to describe wave signals. If the terms frequency and amplitude are applied to halftone screening, frequency means the number of halftone dots per cm which is identical with the screen ruling, whereas the amplitude represents the size of the halftone dots. This is illustrated in figure 1.

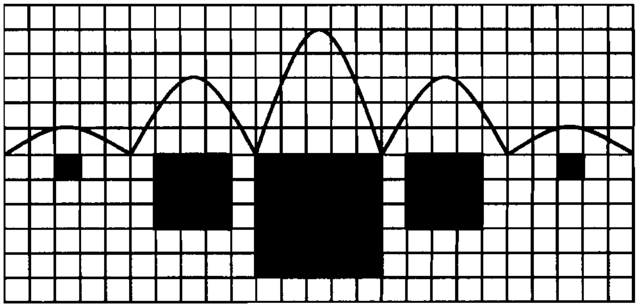
The conventional halftone screening process can be therefore referred to as an amplitude modulation screening process.

FM screening is not possible with conventional photographic processes and can only be produced electronically.

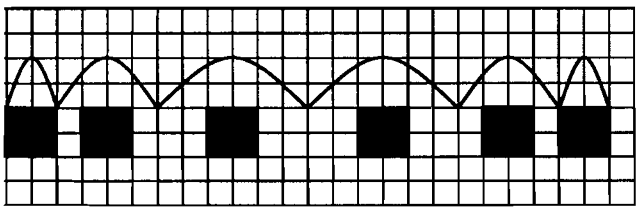
In the present paper, the mathematical algorithms for FM screening are studied and judged with respect to their resulting image quality. Other than at the Technical University of Darmstadt, FM screening will be examined only with simple algorithms.

To better understand FM screening, first the conventional electronic screening process will be briefly described.

* EMPA/UGRA St. Gallen, Switzerland



a) amplitude modulation



b) frequency modulation

Figure 1: **Principle of amplitude and frequency modulation**

2. Principle of amplitude modulation screening

In electronic screening, a halftone dot is usually generated from four scanned pixels. For screening, a pixel is divided into a matrix of single recording dots. Their number is variable and depends on the used imagesetter. A matrix of 6 x 6 single dots is a typical value for one pixel. Then a halftone dot consists of 4 x 6 x 6 = 144 single dots. The matrix of 12 x 12 dots allows to render 144 different dot area values (and in addition the dot area of 0% for the paper white). For each of these 144 dot areas the dot arrangement can be freely chosen what permits to define the shape of halftone dots. Figure 2 shows how a tone value of 25% can be written. The question whether chain dots or round dots are written, depends on the arrangement of the recording dots in the matrix. Such an arrangement is called bitmap. To render 144 values, the same amount of bitmaps is required. They are stored in the data memory. The principle of electronic screening consists, therefore, of reading the corresponding bitmap for each gray value being scanned from the original and recording it on film.

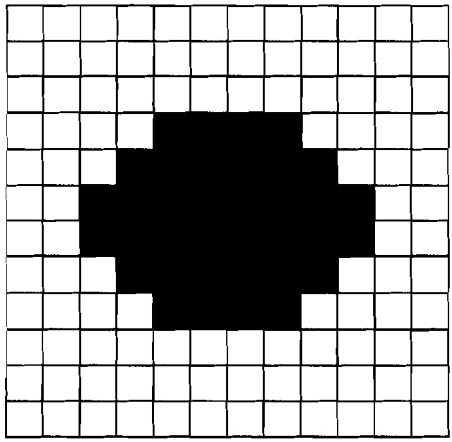


Figure 2: **Arrangement of recording dots in a bitmap for a conventional halftone screening**

3. Principle of frequency modulation screening

As already mentioned, FM screening produces dots of identical size but of different distance or frequency. The dots are not arranged to form a coherent area in the matrix; they are spread up completely irregularly. Figure 3 shows how a tone value of 25% dot area could be written. In higher tone values, however, coherent areas are inevitable. How can this irregular dot arrangement be achieved? In the simpliest case, a dot arrangement in form of a bitmap can be defined for each tone value and put into the data memory. The only difference to amplitude modulation screening is then that the dots in the bitmap are arranged differently, i.e. as isolated small dots rather than as a coherent area. If an image were screened according to that simple procedure, the irregularly arranged dots, would produce an objectionable pattern. To avoid this disadvantage, the dot arrangement should be varied even for the same gray value. Theoretically, each gray value could be related to several bitmaps and addressed alternatingly, but in this way, the data memory would be overloaded. Another possibility is to calculate the dot arrangement each time individually. This happens by means of algorithms which are explained in the following chapter.

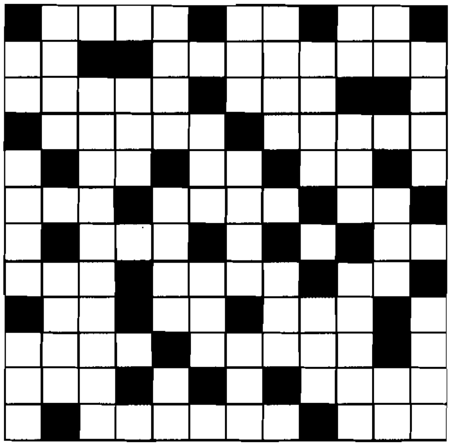


Figure 3: **Arrangement of recording dots in a bitmap for a frequency modulation screening**

4. Algorithms for frequency modulation screening

If an image is scanned, the image information is converted into gray values, normally between 0 (solid-tone) and 255 (white). For the resulting halftone film, however, there are only two possible states: The image area is either black (i.e. printing) or white (i.e. non-printing). The continuous-tone signal produced by the scanner (between 0 and 255) must therefore be transformed into a binary value (1 or 0). The simplest possibility is to code all gray values above a certain threshold value with 1 and the remainder with 0. It is obvious that, in this way, a large part of the image information is lost, especially in the mid-tones of an image.

To obtain a better reproduction result, either the scanned gray values or the threshold value can be modified. The most popular methods based on these principles are the error diffusion process and the dither process.

4.1 Error diffusion

The term itself explains the process: If a gray value is compared with the threshold value, and if the binary value is set 1 or 0, a greater or lesser error will result. This error will now be considered in the further calculations what leads to a diffusion of the error on the following pixels. In the simpliest case, the resulting error is only transferred to the following pixel. For example: If a gray value is 113 and the threshold value is 127, the resulting dot in the binary matrix gets the value 0, i.e. it is a printing dot. The produced error is then 113. If the next gray value is 120, this error is added, yielding a gray value of 120 + 113 = 233 which becomes a non-printing dot. The resulting error is now 255 – 233 = 22. This error is again added to the next gray value. According to this principle, the whole image is coded. However, this

relatively simple one-dimensional process leads to visible structures.
A considerable improvement is achieved with the two-dimensional error diffusion. The error is not only transferred to the following pixels, but also to a number of neighbouring pixels. As an example the error may be distributed to 4 neighbouring pixels as follows:

	X	7
3	5	1

X is the pixel producing the gray value error to be distributed. The numbers represent fractions of 16.
Below is shown another error distribution on 12 neighbouring pixels where the numbers denote fractions of 48:

		X	7	5
3	5	7	5	3
1	3	5	3	1

If the error is distributed to only 4 pixels, the resulting picture still shows an objectionable pattern which can be overcome with the inclusion of further neighbouring pixels. Tests have shown, however, that the inclusion of more than 12 neighbouring pixels give no further improvement of the image structure. Instead of this, the calculation time is considerably extended.

4.2 Dither process

Compared to the error diffusion process, the dither process works with different threshold values which are arranged in a quadratic matrix. The smallest possible arrangement is a 2 x 2 matrix:

0	2
3	1

If the matrix is related to the range of 256 gray levels, the following threshold values result:

1	128
192	64

For the binary coding, each gray value is now compared with the threshold values in the matrix:

112	94	33	142	
109	56	212		

Gray values of the picture

+

0	128
192	64

Threshold matrix

→

1	0	1	1	
0	0	1		

Binary coded picture

The first pixel of the first line ist compared with the threshold value 0. Since the value 112 is larger than the threshold value, the corresponding pixel is coded in the binary matrix with 1. Then, the second gray value 94 is compared with the threshold value 128. This value is smaller, yielding zero in the binary matrix. The third pixel is again compared with zero, the fourth with 128 etc. The procedure for the second line of the picture matrix is the same. With the size of the threshold value matrix, the number of obtainable tone values is given. In a 2 x 2 matrix only five tone values are possible. It is obvious that five tone values are not sufficient for a good reproduction. An increase of the number of tone values can be obtained by enlarging the threshold value matrix. Often, a 8 x 8 matrix is used which can produce 65 tone values.
A typical 8 x 8 dither matrix is shown below:

0	128	32	160	8	136	40	168
192	64	224	96	200	72	232	104
48	176	16	144	56	184	24	152
240	112	208	80	248	120	216	88
12	140	55	172	4	132	36	164
204	76	236	108	196	68	228	100
60	188	28	156	52	180	20	148
252	124	220	92	244	116	121	84

Dispersed-dot dither matrix

If this matrix is used, the process is called dispersed-dot dither. "Dither" refers in this context to the threshold value which is alternating. "Dispersed" describes the regularity of the arrangement. As can be seen, the difference between two subsequent threshold values in one line is always 128. The difference in the columns between two values is alternatingly 192 or 64.
Another possibility of arranging the threshold values is shown below. They are arranged in such a way, that subsequent values are as close as possible what is referred to as clustered-dot dither.

4	44	76	140	136	56	24	12
20	84	116	196	168	104	88	36
52	100	180	228	216	184	120	68
132	164	121	244	248	232	200	148
144	204	236	252	240	208	160	128
64	124	188	220	224	176	96	48
32	92	108	172	192	112	80	16
8	28	60	156	152	72	40	0

Clustered-dot dither matrix

Both processes are compared in figure 4 which shows the resulting dot arrangement for four different gray values. The dot arrangement obtained with the dispersed-dot dither matrix shows the expected picture associated with the term "frequency modulation". In contrast to this, the clustered-dot dither matrix shows a dot arrangement which is more typical for amplitude modulation screening.
This proves that amplitude modulation screening is equally possible with the dither process.

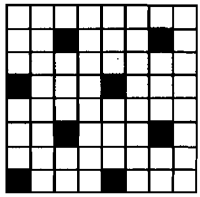
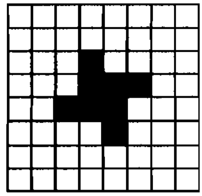
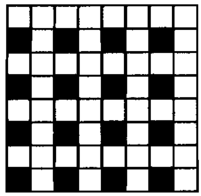
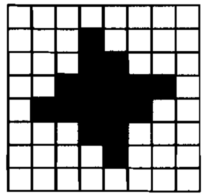
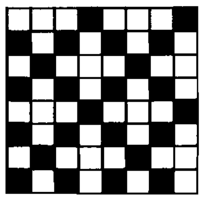
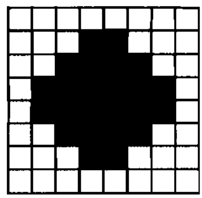
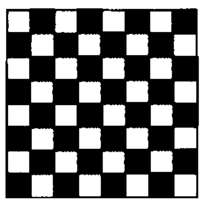
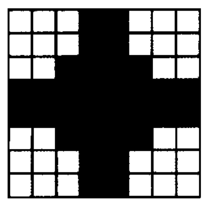
Gray value	Dot area	Number of recording dots	Dispersed-dot dither	Clustered-dot dither
223	12.5	8		
191	25	16		
159	37.5	24		
127	50	32		

Figure 4: Comparison of dispersed-dot dither and clustered-dot dither

5. Application of frequency modulation screening

In offset printing the usual screen ruling is 60 lines/cm (150 lines/inch). If a conventional halftone dot is generated from a matrix of 12 x 12 dots, the recording resolution for the film exposure has then to be 720 dots per cm (1830 dpi). The 720 recording dots are necessary to make sure that the edges of the halftone dot can be plotted without disturbing structures. If FM and AM screening are compared at this high recording resolution, FM screening shows the following advantages:

- Smoother transitions in tone rendering: In particular, the effect of corner link-up occuring with conventional halftone dots is not possible.
- Better detail rendering, because the size of the scanned pixel and not the chosen screen ruling determines the resolution.
- No moiré pattern in multicolor printing and no need for screen angles.

This has already been shown in the studies of the Technical University of Darmstadt. But the aim of the present study was not to compare the two screening processes at a recording resolution of 720 dots/cm (1830 dpi) or higher. The idea was rather to examine what reproduction quality could be obtained with FM screening in systems having lower recording densities. The following explains why amplitude modulation screening leads at low recording densities to unsufficient results.

If the recording resolution is for instance 600 dpi, the single dot has a diameter of 42 micron. In case of a screen ruling of 60 lines/cm (150 lines/inch), this dot size is equivalent to a dot area of about 6%. Therefore, the problem of amplitude modulation screening at low recording resolution is the limitation of gray values (see table 1). This can be overcome with coarse screen rulings, however, at the expense of resolution.

The relationship between the number N of gray values, the screen ruling F and the recording resolution L (in dot/cm) is given by the following equation:

$$N = \left(\frac{L}{F}\right)^2 + 1$$

In contrast to this, the FM screening shows no trade-off between resolution and the number of reproducible gray values.

Recording resolution		**Number of reproducible gray values at a screening ruling of**		
dot/cm	**dpi**	**30 lines/cm (75 lines/inch)**	**40 lines/cm (100 lines/inch)**	**60 lines/cm (150 lines/inch)**
120	300	17	10	5
160	400	26	17	10
240	600	65	37	17
400	1000	170	101	50
500	1250	257	145	65
1000	2500	1090	626	257

Table 1: **Number of gray values vs screen ruling for different recording densities**

6. Experimental work

Different algorithms based on the principles of error diffusion and dithering were tested as to their ability to render gray tones without an objectionable pattern.
Altogether 12 algorithms have been tested, using a specially designed test image.
The binary coded images were recorded with 125 dots/cm (318 dpi) and 250 dots/cm (635 dpi) on a Linotronic laser imagesetter.
In a first evaluation, the images recorded with 12 algorithms were visually assessed by pair comparison. As a result, one algorithm based on a modification of the error diffusion principle was chosen for further tests.
Using this algorithm, a series of four test images, a step wedge and a continuous wedge were reproduced with 250 dots/cm (635 dpi) and 500 dots/cm (1270 dpi).
For a comparison, the same pictures and test elements were reproduced with a conventional halftone screen at a screen ruling of 60 lines/cm (150 lines/inch).
Appendices 1, 2 and 3 show the same picture, as conventional halftone image (recorded with 2540 dpi) and with FM screening (recorded at 635 dpi and with 1270 dpi). At first glance, there is no quality difference evident between the two screening principles. A closer look, however, reveals a very slight graininess in the FM picture obtained with 635 dpi which can be overcome by increasing the recording resolution to 1270 dpi.
Appendix 4 illustrates that FM screening permits to reproduce a stepless gray wedge without discontinuities if recorded with 635 dpi. In contrast to this, a conventional halftone wedge recorded with 635 dpi shows a limited number of gray values visible as steps in a continuous wedge. To avoid this, a recording resolution of 2540 dpi is necessary.
While the test images shown in the appendices were recorded on film, a further investigation was aimed at testing how FM screened images are rendered on low resolution printers. In a first test series a laser printer with 600 dpi was investigated. Using a test image generated with 600 dpi, it was found that the resolution of the toner particles is not sufficient to correctly render the tonal values of the image. However, an image generated with 300 dpi can be satisfyingly reproduced on a 600 dpi printer. A finer particle size of the toner would clearly contribute to a better image quality, an aspect which will be investigated in future tests.
The conversion of the continuous-tone data into an FM screened image has so far been performed on a mainframe computer based on a FORTRAN program. With the availability of FORTRAN compilers for Macintosh Computers the program will be integrated in a DTP system and made available for interested users of Macintosh based DTP equipment.

A further series of tests dealt with the question of what dot gain values result compared with conventional halftone dots.
Figure 5 shows the dot gain curve of a 635 dpi gray wedge printed on coated paper compared with a conventional gray wedge screened at 60 lines/cm (150 lines/inch).
As can be seen, FM screened dots give higher dot gain values which, however, can be compensated by a preadjustment of the gradation.

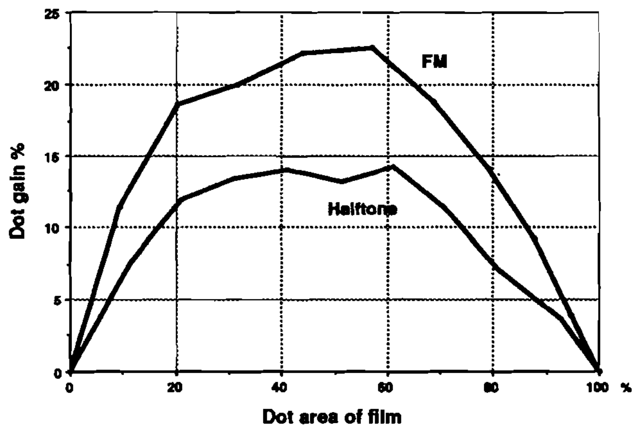


Figure 5: **Dot gain curves for FM screen dots recorded with 635 dpi and conventional halftone dots of 60 lines/cm printed on coated paper**

7. Conclusion

This study is not aimed at proposing a substitution for the conventional screening process. What has been shown is that with relatively simple algorithms and at low recording densities an image quality is achievable which closely matches the quality of a conventional halftone image screened with 60 lines/cm (150lines/inch).

The benefits of FM screening compared with conventional halftone screening are:

- No limitation of the number of reproducible tone values, not even at very low recording densities such as 300 dpi and 600 dpi.
- Better detail rendering, because the recording resolution can be equal to or higher than the scanning resolution.
- Smoother transitions of tone values, because discontinuities caused by merging halftone dots cannot occur.
- No moiré patterns in multicolor printing, because FM screening requires no screen angles.

To conclude, some minor drawbacks of FM screening cannot be fully denied, as for instance:

- higher dot gain values in printing
- a tendency to graininess at low recording densities (below 600 dpi)
- a slightly increased calculation time for the conversion from continuous-tone data to binary coded data.
- a sensitivity to losses in dot area when duplicated from film to film or printed down on plates, if the recording resolution exceeds 1200 dpi.

8. References

(1) Humbel, V.: Frequenzmodulierte Rasterverfahren und ihre Eignung für niedrig auflösende Wiedergabesysteme. UGRA-Bericht 89/1 , St. Gallen, 1990.

(2) Fischer, G.: Der frequenzmodulierte Bildaufbau – ein Beitrag zum Optimieren der Druckqualität. Dissertation. Darmstadt, 1986.

(3) Ulichney, R.: Digital Halftoning. The MIT Press. Cambridge, Massachusetts, 1987.

(4) Riedel, K.: Datenreduzierende Bildcodierung: Kapitel 6: Pseudo-Grauton-Codierung. Francis Ingenieur Elektronik. München, 1986.

(5) Stucky, P.: MECCA – A Multiple-Error Correction, Computation Algorithm for Bilevel Image Hardcopy Reproduction. Research Report RZ 1060, IBM Research Laboratory. Zürich, 1987.

Appendix 1

Frequency modulation screening
635 dpi
Screening program EMPA/UGRA

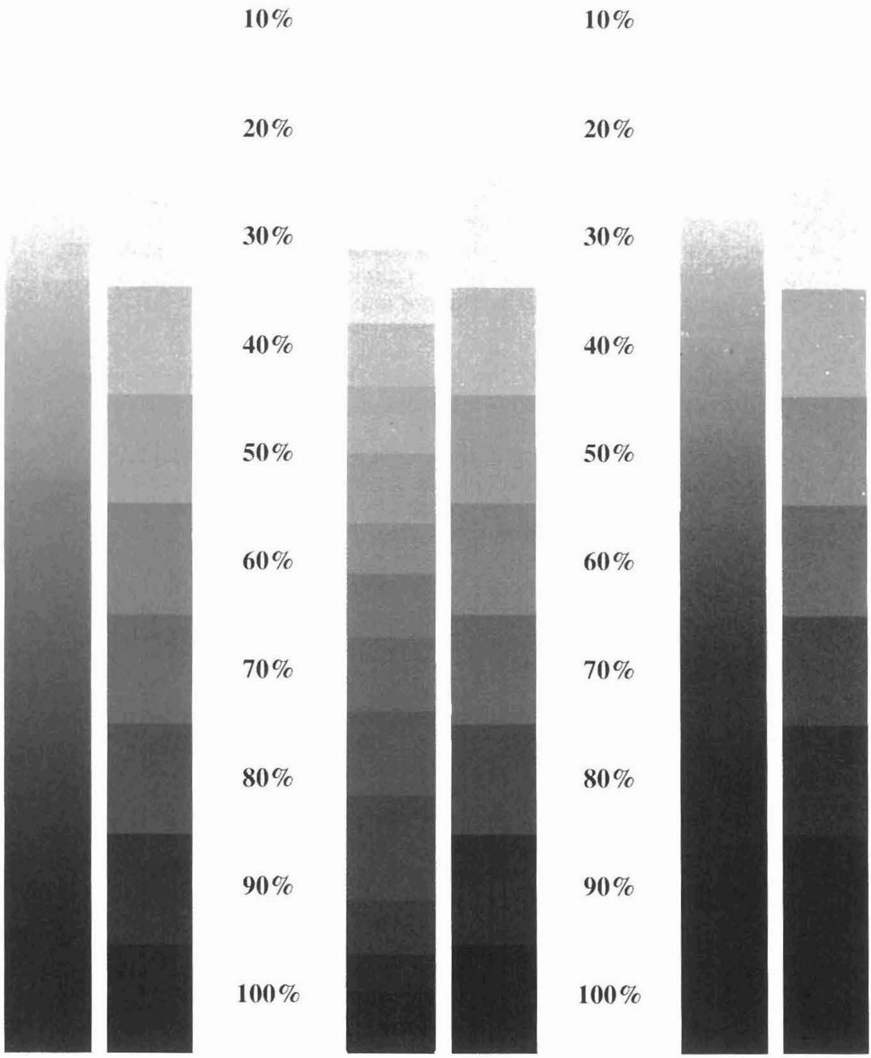
Appendix 2

Frequency modulation screening
1270 dpi
Screening program EMPA/UGRA

Appendix 3

Conventional halftone screening
60 lines/cm (150 lines/inch)
recorded with 2540 dpi

Appendix 4



Conventional halftone screening
60 lines/cm
(150 lines/inch)
recorded with 2540 dpi

Conventional halftone screening
60 lines/cm
(150 lines/inch)
recorded with 635 dpi

Frequency modulation screening
635 dpi
Screening program
EMPA/UGRA