METHODS FOR SPATIOTEMPORAL DITHERING

Jeffrey B. Mulligan
NASA Ames Research Center

1. Introduction

The problem of representing gray-level images on a binary display device is known as "dithering" or "halftoning." (A good survey can be found in Ulichney[1]). Dithering relies upon the fact that the human visual system integrates information over spatial regions, so that a spatial pattern of light and dark can evoke a sensation approximating that of a uniform gray area even when the individual display elements can be resolved. Electronic image display devices, such as cathode ray tubes and flat panel displays offer the additional possibility of exploiting the visual system's integration in the time domain to increase a display's gray scale resolution; additionally, it may be possible to exploit the visual system's spatiotemporal sensitivity to make dynamic dithering noise which is less visible than the corresponding static noise. In this paper we examine how a number of existing two-dimensional dithering methods may be generalized to three dimensions.

2. Temporal properties of vision

The potential power of these techniques is rooted in the fact that the visual system has different temporal responses to different image features. Halftoning algorithms may exploit this by hiding high frequency noise in the perceptual bands which have a low-pass response. One of the best examples is color: when two colored lights are exchanged or flickered, the color will appear to alternate at low flicker rates, but when the frequency is raised to 15-20 Hz, *color flicker fusion* occurs, where the flicker is seen as a variation of intensity only. The subject can eliminate all sensation of flicker by balancing the intensities of the two lights (at which point the lights are said to be *equiluminant).* When the intensities are not balanced, the luminance flicker can be seen at frequencies as high at 50-60 Hz.

The differential *spatial* properties of the chromatic and achromatic systems have been exploited in several halftoning algorithms[2,3]. Mulligan[2] proposed negatively correlating the dither matrices in multi-component ordered dither to reduce spatial luminance variation at the expense of added chromatic noise. This can be done in the temporal domain equally well. Mulligan and Ahumada[4] , and Balasubramanian, Carrara and Allebach[3] used different spatial error filters for achromatic and chromatic quantization errors in iterative algorithms designed to find the visually optimal halftone.

Other parameters besides color affect temporal sensitivity. For example, the spatial frequency of an achromatic pattern influences the temporal sensitivity[5-11]. For low to medium spatial frequencies, temporal sensitivity is bandpass with peak sensitivity between 5-10 Hz, while for patterns above 4 cycles per degree temporal sensitivity becomes lowpass.

3. Algorithms

3.1. Dispersed dot ordered dither

The term ordered dither refers to a general class of algorithms in which the array of desired input values is compared with a corresponding array of threshold values. The output pixels are set or cleared depending on the outcome of this comparison. "Dispersed dot" ordered dither usually refers to the work of Bayer[12] who demonstrated an algorithm for generating matrices which are optimal under a certain set of assumptions. These matrices insure that for any gray level, pixels are set in a uniform way across the cell.

Large dither matrices can be constructed using a recursive algorithm. We start with the smallest non-trivial matrix, in this case 2x2. We can generate two new matrices with twice the linear size by a) enlarging the original matrix by replicating each entry; or b) replicating the entire matrix. We then multiply the entries of the second matrix by 4, and form the final matrix by adding the values from the first (enlarged) matrix. This process is illustrated in figure 1. The procedure can be applied recursively until the desired size of matrix is obtained.

Figure 1: Illustrates the generation of a 4 by 4 dither matrix from a 2 by 2 seed matrix. The process of replication and rescaling may be applied recursively to generate arbitrarily large matrices, and is easily extended to more dimensions (see text).

The matrices generated by this procedure are completely determined by the starting matrix of size 2. If we ignore differences due to rotations and reflections, then there are three possibilities for the two-dimensional case, one of which generates the Bayer matrices. For any choice of the seed matrix, the algorithm insures that for any given level, the number of set bits will differ by at most one between any pair of quadrants, and this principle holds recursively for subquadrants. The seed matrix corresponding to the Bayer matrix produces a checkerboard pattern when a density of 1/2 is requested, while the other two seed matrices produce patterns of horizontal or vertical stripes. Empirically, the checkerboard pattern is less visible, so it is the pattern of choice.

The matrix enlargement algorithm is easily generalized to three dimensions by replicating in three dimensions instead of two, multiplying the entries of the input matrix by 8 instead of 4, and adding this to each octant of the enlarged three dimensional matrix. The number of possible 2x2x2 seed matrices for the three dimensional case is much larger than for the two-dimensional case. If (and this is a big if) we ignore rotations in space-time (6), as well as reflections (2) and purely spatial rotations (4), then the number of distinct seed matrices is 8!/(6*2*8)=840. We can apply the following heuristics to reduce the search space: first, by analogy with the two-dimensional case, we might require that for a requested density of 0.5 that we obtain a spatial checkerboard pattern in each temporal frame; secondly, we might require that the checkerboards in frame 1 and frame 2 have opposite spatial phase. Another desireable constraint is that for any input gray level, the number of thresholds exceeded in each the two frames differ by at most 1.

A matrix that satisfies all of these properties may be constructed from the 2x2 seed matrix shown in figure 1 as follows: the first frame is made by multiplying the matrix in figure 1 by 2. The second temporal frame is formed by rotating this matrix by 90 degrees and adding 1 to each of the entries. This seed cube can then be expanded to any desired size (which is a power of 2) by repeated application of the algorithm described above.

3.2. Temporal Error Diffusion

The method known as error diffusion was introduced by Floyd and Steinberg[13] , and is generally agreed to produce output images of higher quality than those produced using ordered dither. It is a serial process which proceeds as follows: at each pixel, the desired level is rounded to the nearest quantization level, which is output. The error is computed by subtracting the desired value from the quantized value. This error is "diffused" by subtracting fractions of it from the desired values of nearby unquantized pixels. The precise pattern of how the error is distributed determines the resulting patterns.

The most obvious way to generalize this algorithm to three dimensions is simply to diffuse the error with a three dimensional spread function. The proportion of error which is diffused in the current frame, as opposed to spread into the following frame, can be used to shape the spatiotemporal spectrum of the resulting quantization noise.

The principle of temporal error diffusion can also be combined with any other two-dimensional spatial dithering algorithm as follows: we first obtain a quantized image using any method we choose. We then compute the error image by subtracting the desired image from the quantized image. The desired image for the second frame is then modified by subtracting the error image from the previous frame. This algorithm has the advantage that no extra filtering is required, and the individual frames can be processed using a parallelizable algorithm such as ordered dither. There is little opportunity, however, to control the spatiotemporal parameters of the resulting quantization noise.

3.3. Visually optimal methods

Because of the inherently serial way in which pixels are processed in the error diffusion algorithm, it is impossible for a pixel to share its error equally with *all* of its neighbors. This difficulty has been overcome by a number of stochastic algorithms which have been proposed recently[3,4,14-17] all of which do basically the same thing. First, a filter is specified which is applied to the error image. This filter is usually designed to capture the contrast sensitivity of the human visual system. The filtered error image is then condensed to a single number, such as the sum of the squared error, to give an overall quality measure. Pixels are visited sequentially and their states are changed so as to minimize the error measure. To increase the chances of finding the global optimum, the technique of simulated annealing may be used, in which pixels are occasionally set to states which increase the overall error measure, in order to escape from local minima. The probability that these (usually unfavorable) transitions occur is controlled by a parameter known as the "temperature." The image is "annealed" by slowly reducing the temperature. Eventually, the image will "freeze" into a final state. The probability that this state is the global optimum increases as the rate of cooling decreases.

Although this class of algorithm is the most computationally intensive of those which have been discussed, it has been recommended as a benchmark because of the straightforward way in which models of visual quality can be integrated with the halftoning process. Unlike error diffusion, where the effective error filter must be obtained by recursive application of the diffusion weights, here the finite impulse response filter that is applied to the error *is* the filter with respect to which the error is minimized. Thus for the three dimensional case it is a fairly straightforward matter to design a filter using data on spatiotemporal contrast sensitivity.

To date most applications of this algorithm have used a single two-dimensional low pass filter to model visibility. This type of filter has the disadvantage that it is not particularly sensitive to oriented patterns for which the visual system seems to have specialized detectors. This defect becomes more pronounced if we apply this simple approach to three dimensions. The spatiotemporal analog of two-dimensional orientation is *motion;* a three dimensional lowpass filter tends to blur patterns which move coherently through it, although such patterns may be quite visually salient. For results which are truly visually optimal, we need to filter the error in a manner more consistent with how we believe the visual system processes the input imagery, namely with not one but an array of filters tuned for different spatiotemporal orientations[18] , Visual models such as this will also be valuable for obtaining objective measures of the performance of the simpler algorithms described in the preceding sections.

4. Conclusions

Spatiotemporal dithering is a potentially useful way of extending the grayscale resolution of a display device while at the same time reducing the visibility of static artifacts. All of the two dimensional spatial dithering algorithms currently in use can be generalized to three dimensions in a fairly straightforward manner.

5. Acknowledgements

This work was supported by NASA RTOP's 505-64-53 and 506-59-65, and DARPA's High Resolution Systems Program.

6. References

1. Ulichney, R., *Digital Halftoning*, MIT Press, Cambridge, Mass., 1987.
2. Mulligan, J. B., ''Digital halftoning methods for selectively partitioning error into achromatic and chromatic channels,'' in *Human Vision and Electronic Imaging: Models, Methods, and Applications*, ed. B. E. Rogowitz and J. P. Allebach, pp. 261-270, Proc. SPIE 1249, paper #21, 1990.
3. Balasubramanian, R., Carrara, D., and Allebach, J., ''Color image quantization using direct binary search,'' in *Human Vision, Visual Processing and Digital Display IV*, ed. B. E. Rogowitz and J. P. Allebach, Proc. SPIE 1913, paper #20, 1993.
4. Mulligan, J. B. and Ahumada, A. J. Jr, ''Principled methods for color dithering based on models of the human visual system,'' *SID Int. Symp. Dig. Tech. Papers*, pp. 194-197, 1992.
5. Robson, J. G., ''Spatial and temporal contrast sensitivity functions of the visual system,'' *J. Opt. Soc. Am.*, vol. 56, pp. 1141-1142, 1966.
6. van Nes, F. L., Koenderink, J. J., Nas, H., and Bouman, M. A., ''Spatiotemporal modulation transfer in the human eye,'' *J. Opt. Soc. Am.*, vol. 57, pp. 1082-1088, 1967.
7. Kelly, D. H., ''Spatio-temporal frequency characteristics of color-vision mechanisms,'' *J. Opt. Soc. Am.*, vol. 64, pp. 983-990, 1974.
8. Koenderink, J. J. and van Doorn, A. J., ''Spatiotemporal contrast detection threshold surface is bimodal,'' *Optics Letters*, vol. 4, pp. 32-34, 1979.
9. Noorlander, C., Heuts, M. J. G., and Koenderink, J. J., ''Sensitivity to spatiotemporal combined luminance and chromaticity contrast,'' *J. Opt. Soc. Am.*, vol. 71, pp. 453-459, 1981.
10. Noorlander, C., Koenderink, J. J., den Ouden, R. J., and Edens, B. W., ''Sensitivity to spatiotemporal colour contrast in the peripheral visual field,'' *Vision Res.*, vol. 23, pp. 1-11, 1983.
11. Watson, A. B., ''Temporal sensitivity,'' in *Handbook of perception and human performance, vol. I Sensory processes and perception*, ed. K. R. Boff, L. Kaufman, and J. P. Thomas, John Wiley and Sons, New York, 1986.
12. Bayer, B. E., ''An optimum method for two-level rendition of continuous-tone pictures,'' *Conference Record of the Intl. Conf. on Communications*, pp. 26-11 - 26-15, 1973.
13. Floyd, R. and Steinberg, L., ''An adaptive algorithm for spatial gray scale,'' *SID 1975 Symp. Dig. Tech. Papers*, p. 36, 1975.
14. Carnevali, P., Coletti, L., and Patarnello, S., ''Image Processing by Simulated Annealing,'' *IBM J. Res. Develop.*, vol. 29, pp. 569-579, 1985.
15. Anastassiou, D., ''Neural net based digital halftoning of images,'' *Proc. IEEE Symp. Circuits & Systems*, pp. 507-510, 1988.
16. Analoui, M. and Allebach, J. P., ''Model based halftoning using direct binary search,'' in *Human Vision, Visual Processing and Digital Display III*, ed. B. E. Rogowitz, J. P. Allebach, and S. A. Klein, pp. 96-108, Proc. SPIE 1666, paper #9, 1992.
17. Mulligan, J. B. and Ahumada, A. J. Jr., ''Principled halftoning based on human vision models,'' in *Human Vision, Visual Processing and Digital Display III*, ed. B. E. Rogowitz, J. P. Allebach, and S. A. Klein, pp. 109-121, Proc. SPIE 1666, paper #10, 1992.
18. Watson, A. B., ''Perceptual-components architecture for digital video,'' *J. Opt. Soc. Am. A*, vol. 7, pp. 1943-1954, 1990.