Error Diffusion and Delta-Sigma Modulation for Digital Image Halftoning

Thomas D. Kite, Brian L. Evans, and Alan C. Bovik
Department of Electrical and Computer Engineering
The University of Texas at Austin

Presented at Hewlett-Packard Research Labs, Palo Alto, CA
March 17, 1997

Outline

Introduction to Digital Halftoning

– Applications
– Common halftoning methods
– 1-D delta-sigma modulation
– Halftoning with the 1-D DSM

Analysis of Error Diffusion

– Extension of delta-sigma modulation to 2-D
– Linear analysis of the 2-D DSM
– Improved modelling
– Image quality metrics

Design and Implementation

– Filter optimization
– Implementation issues

Conclusions

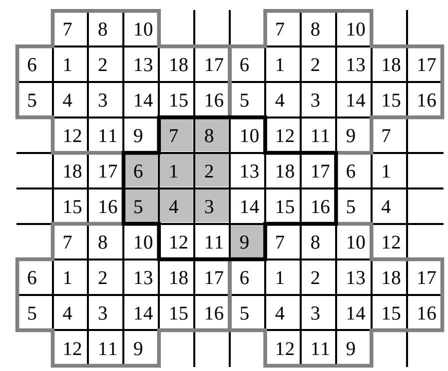
– Future research

The Need for Digital Halftoning

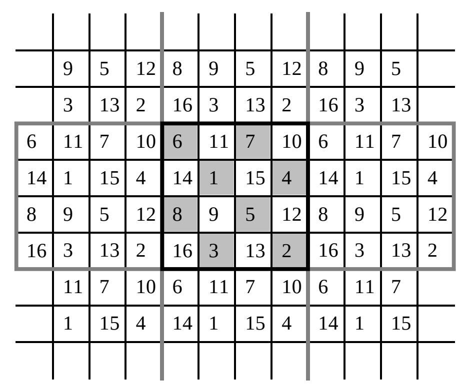
- Grayscale and color imagery now ubiquitous
- Many devices are incapable of reproducing grayscale
  - Laser printers
  - Inkjet printers
  - Facsimile machines
  - Low-cost liquid crystal displays
- Grayscale imagery must be binarized for these devices
- Halftoning attempts to reproduce the full range of gray while preserving image quality and spatial resolution
- Screening techniques are fast and simple
- Error diffusion gives better results on some media

Halftoning by Screening

- Tessellate image with threshold screen (two screens are shown)
- Threshold pixel relative to corresponding screen value; activate output pixel if image pixel ≥ screen pixel
- Clustered-dot screen clumps output pixels together
- Dispersed-dot screen keeps output pixels apart
- Point operation; very simple computationally (parallel implementation possible)



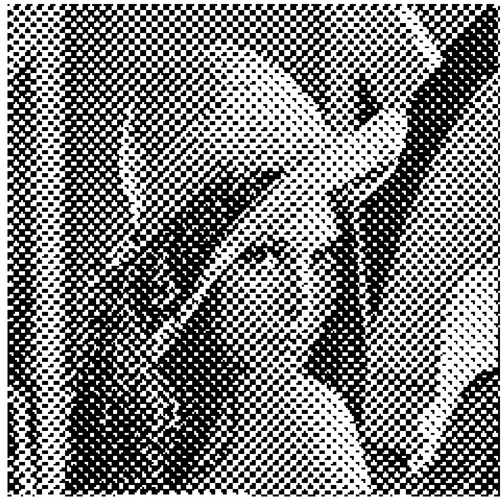
19-level clustered-dot screen



16-level dispersed-dot screen

Typical Screening Results

Original image



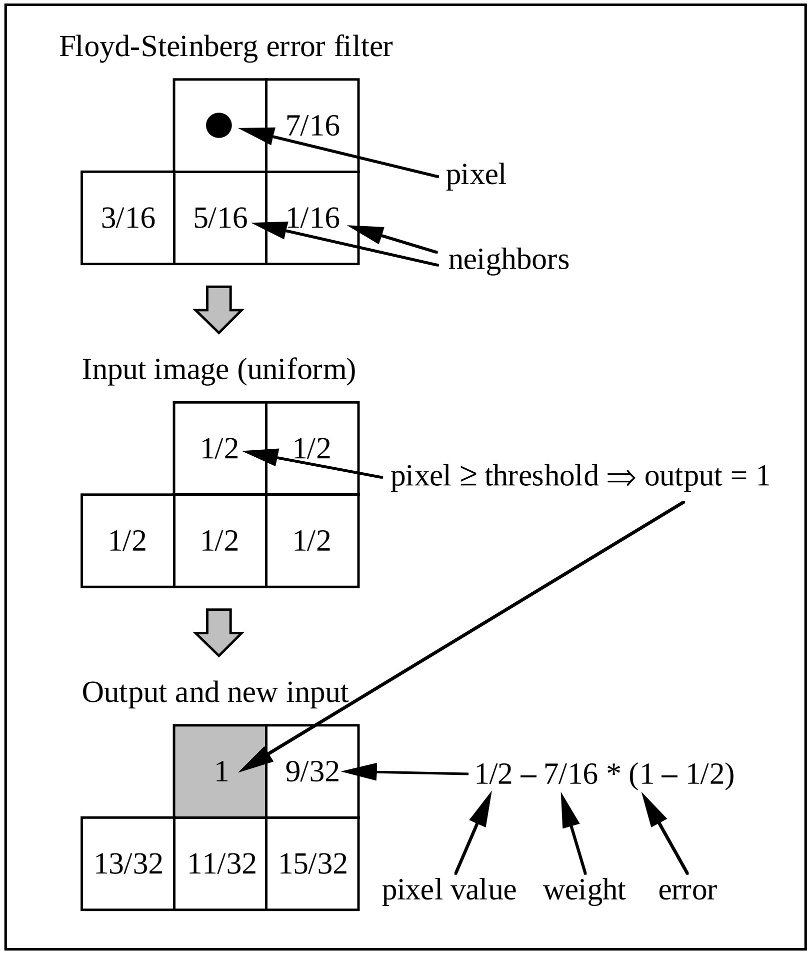
19-level clustered dot

16-level dispersed dot

- Clustered dot screening produces a coarse image that is more resistant to defects such as ink spread
- Dispersed dot screening has higher spatial resolution
- Both have equal computational complexity and noticeable artifacts

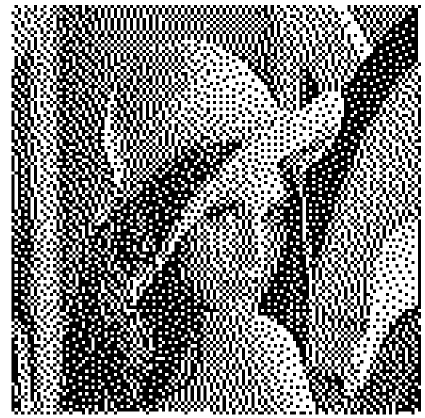
Halftoning by Error Diffusion

- Introduced by Floyd and Steinberg, 1975
- Not a screening technique—instead, quantizes each pixel and distributes quantization error among neighbors
- Error filter weights chosen by trial and error for good visual results
- Artifacts due to scan order are visible—can be ameliorated somewhat with non-raster scans (Fan, 1994; Knox, 1994)



Typical Error Diffusion Results

Original Image



Floyd-Steinberg

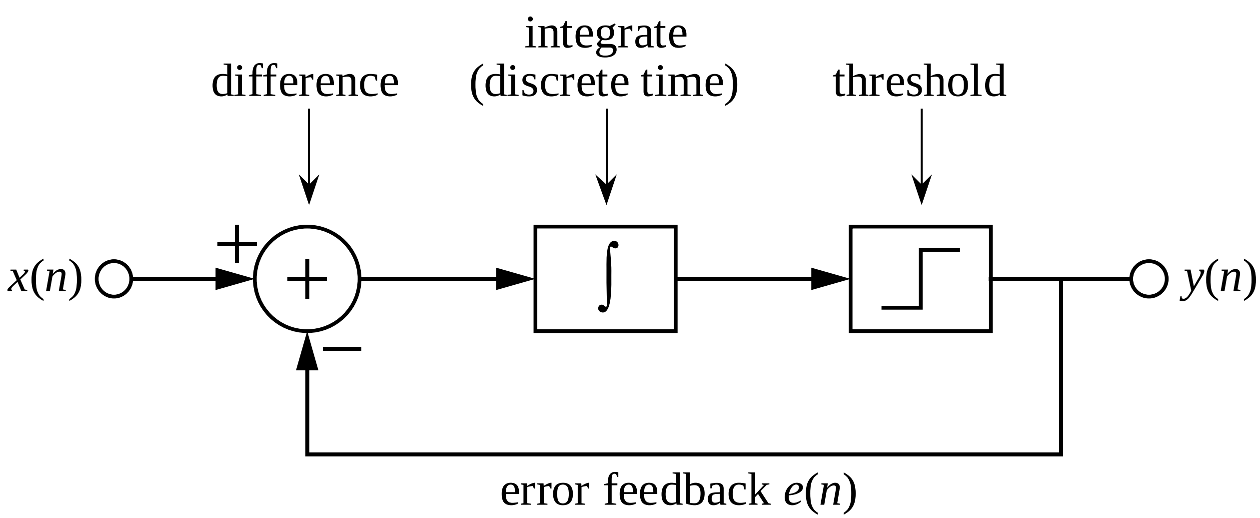
Jarvis *et al.*

Stucki

Developments in Error Diffusion

- Floyd and Steinberg's error filter modified (Jarvis, Judice & Ninke, 1976) to produce sharper images
- Further modification (Stucki, 1980) gave similar results
- General framework for error diffusion and 'blue noise' concept developed (Ulichney, 1987)
- Connection between error diffusion and delta-sigma converters identified (Bernard, 1991)
- Blue noise screening technique demonstrated (Mitsa & Parker, 1991)
- Error diffusion partially analyzed and improved (Knox, 1992; Eschbach, 1993; Wong, 1995; others)

1-D Delta-Sigma Modulation



- First-order modulator shown; usually run at high oversampling ratio
- First stage computes difference between input and previous output
- Error is integrated in second stage and thresholded for a one-bit output
- Quantizer can be substituted for thresholder to give *N*-bit output with shaped noise (wordlength reduction)
- Higher order modulators use multiple integrators

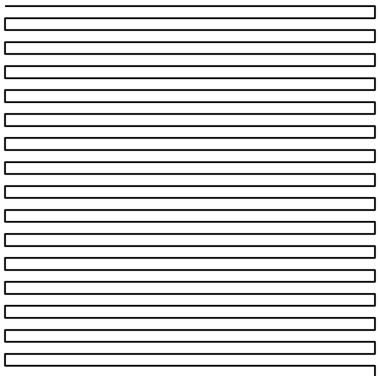
Linear Analysis of 1-D Modulator

- Assume quantizer simply adds noise (linear assumption)
- Analysis then predicts

$$Y(z) = z^{-1}X(z) + [1 - z^{-1}]N(z)$$

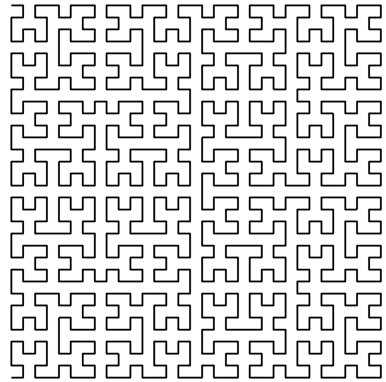
- Signal is passed with a delay; noise is high-pass filtered
- Higher order loops can achieve shaping as $[1 - z^{-1}]^k$, where k is the order of the loop
- Since output is always ±1 and signal transfer function is flat, noise power does not change with loop order; noise is merely redistributed spectrally
- Linear analysis does not predict idle tones, distortion

Halftoning with a 1-D Modulator



Serpentine path

Halftoned



Peano path

Halftoned

- To operate on a 2-D image with a 1-D process, we scan the image
- Choice of scan greatly affects the results, which are visually noisy
- Quantization error is distributed along the path of the scan:
  - Raster, serpentine scans distribute error mostly horizontally
  - Peano scan (Witten & Neal, 1984) distributes error haphazardly
- Genuine 2-D extension needed for high-quality visual results

Outline

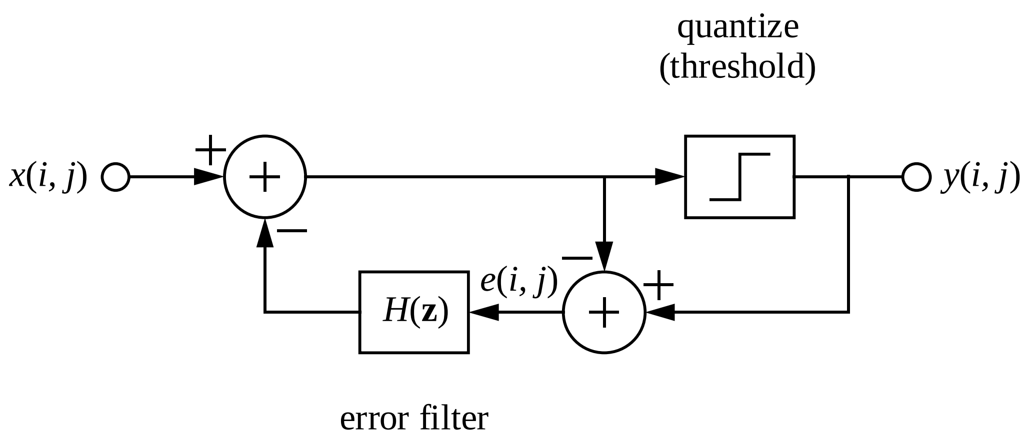
Introduction to Digital Halftoning
- Applications
- Common halftoning methods
- 1-D delta-sigma modulation
- Halftoning with the 1-D DSM

Analysis of Error Diffusion
- Extension of delta-sigma modulation to 2-D
- Linear analysis of the 2-D DSM
- Improved modelling
- Image quality metrics

Design and Implementation
- Filter optimization
- Implementation issues

Conclusions
- Future research

Noise Shaping Feedback Coder



- Used for wordlength reduction (e.g., 8 bits to 1 bit for images)
- Quantization error is shaped spectrally by the error filter, increasing effective resolution in part of the passband
- Shaping designed to achieve high resolution where noise would be most objectionable psychophysically
- Equivalent to conventional delta-sigma modulator; offers simplicity of form (which is identical to error diffusion)

Linear Analysis of the NSFC

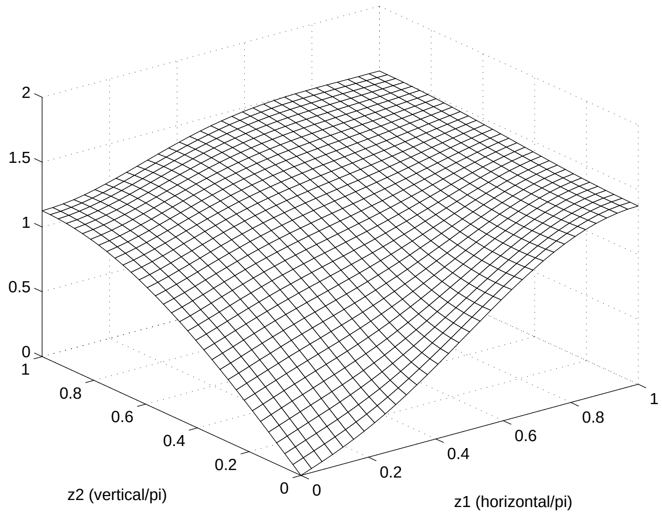
- Again, assume quantizer is an adder of white noise
- Analysis then predicts:

$$Y(z) = X(z) + [1 - H(z)]N(z)$$

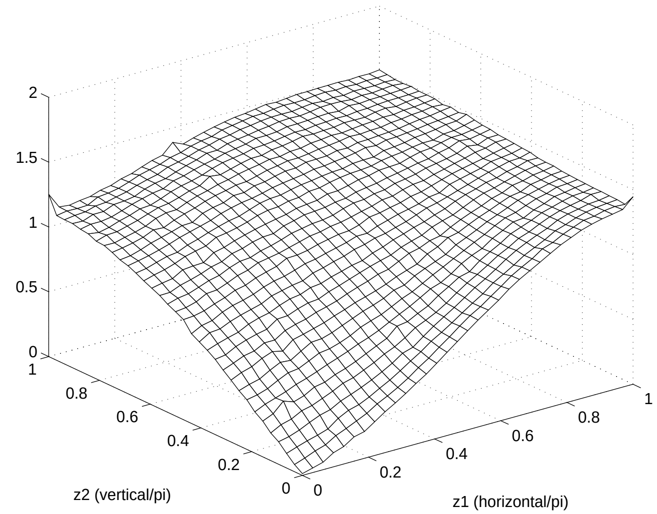
- Signal $X(z)$ passes unchanged
- Noise is filtered by $1 - H(z)$
- For error diffusion, z is now a vector (z_1, z_2)
- All error diffusion schemes have $H(0, 0) = 1$, i.e. the noise transfer function (NTF) has a zero at DC
- High-pass NTF shapes noise to frequencies where the human visual system is less sensitive

Success of Linear Analysis

- Noise images halftoned with Floyd-Steinberg algorithm
- Measured averaged NTF compares well with NTF predicted from linear analysis



Predicted NTF



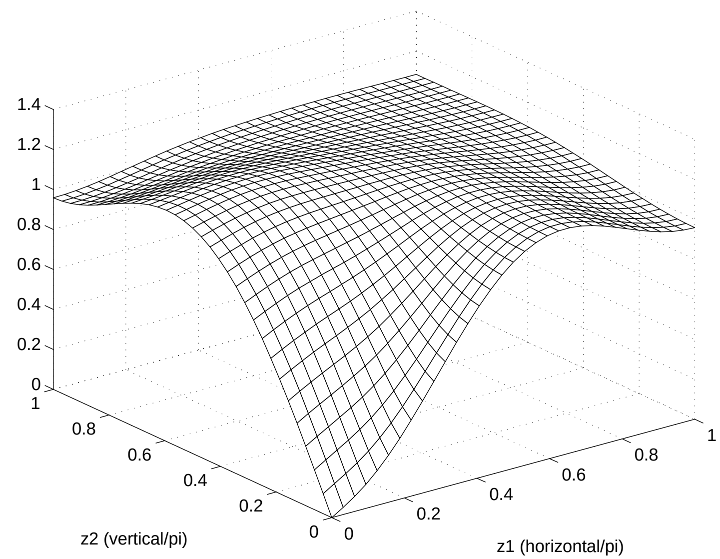
Measured NTF

Failure of Linear Analysis

- Predicts an output that is the sum of the input and shaped noise; this is visibly false
- Predicts a noise image uncorrelated with the input; this is demonstrably false (Knox, 1992)
- Predicts a flat signal transfer function (STF), yet the larger filters (Jarvis *et al.*, Stucki) perform edge sharpening
- Fails to predict ‘idle tones’ noticeable in smoothly-changing areas (Fan & Eschbach, 1994)
- Similar failures noted in the audio delta-sigma literature (Gray, 1997; many others)
- Some other modelling of the quantizer is needed

Jarvis Filter Edge Sharpening

- Jarvis, Stucki filters produce noticeable edge sharpening
- NTF is more exaggerated than Floyd-Steinberg
- Linear analysis fails to predict edge sharpening
- Sharpening must therefore be due to failure of the analysis
- Quantization error is correlated with the input, leading to a non-flat effective STF
- Can we model the quantizer to take account of this correlation?



Jarvis filter NTF

Nature of the Quantization Error

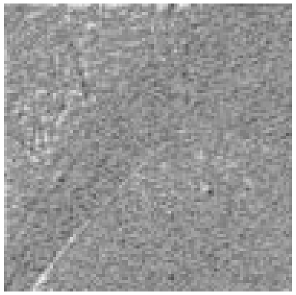
- Quantization error is highly correlated with the output (and therefore with the input)
- Error has a good linear fit with the output; for this image

$$QERR(x, y) \approx 0.85Y(x, y)$$

- Difference between output and linear fit of error is almost completely noise-like
- Large signal component in error explains edge sharpening
- Linear fit of error suggests a gain model for the quantizer

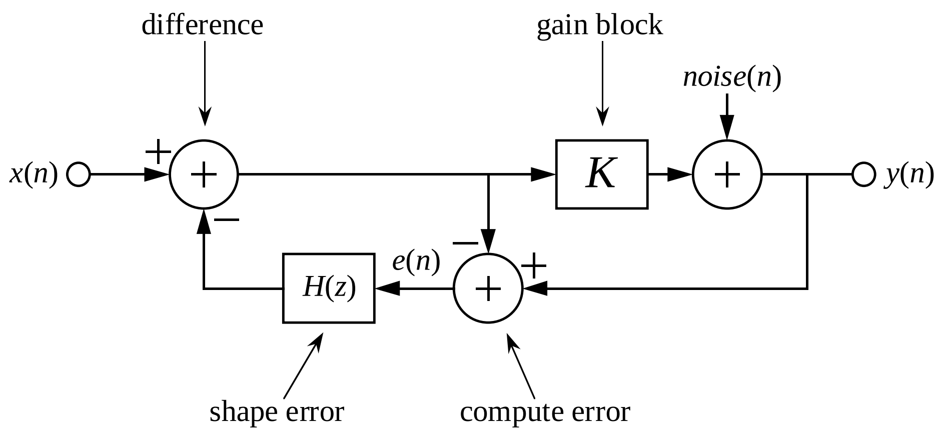
Jarvis filter output

Quantization error



Difference between error and linear fit to output

Gain Model for the Quantizer



- Replace the quantizer by a gain block (applied to an audio converter by Ardalan & Paulos, 1987). We measured the following:
  - Signal: $K \approx 2$ for Floyd-Steinberg scheme, $K \approx 5$ for Jarvis scheme
  - Noise: $K = 1$ for all schemes
- Quantization error now contains a large input signal component, which is filtered by $H(z)$ and added to the input
- Since $1 - H(z)$ is usually high-pass, input signal is boosted in the high frequencies, producing edge sharpening

Results of the Gain Model

Original image

Gain model, $K = 5.1$

Jarvis halftoned

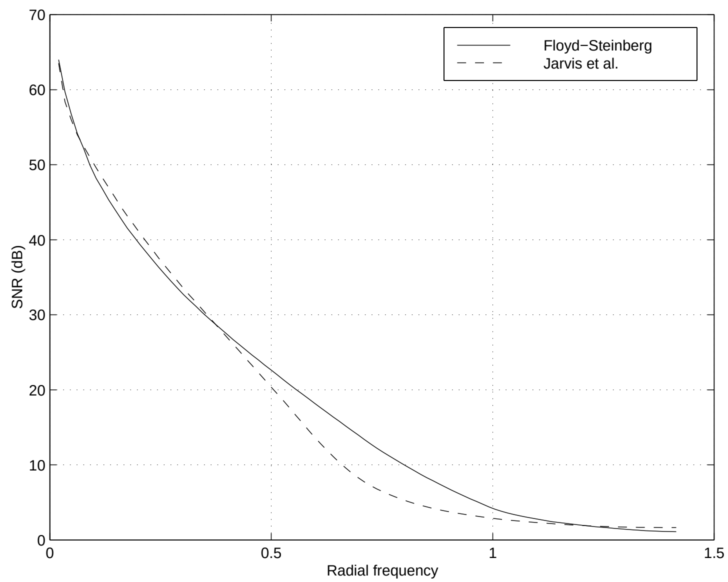
- Edge sharpening accurately modelled by assuming quantizer has gain
- Signal gain K is approximately constant for a given halftoning scheme (similar phenomenon also noted by Knox, 1992)
- Quantization noise gain K is unity for all schemes, i.e. NTF = $1 - H(z)$, as predicted by linear model

Image Quality Metrics

- We model halftoning as a linear system in which a gain block and noise adder are substituted for the quantizer
- All halftoning schemes perform spectral shaping (edge sharpening) on the signal, and add quantization noise
- We measure image quality in the following way:
  1. Sharpen the original image using the noiseless gain block model
  2. Filter the sharpened original image and the halftoned image with identical sharp cutoff low-pass filters
  3. Compute the signal-to-noise ratio (SNR) between the two filtered images
  4. Increase the cutoff frequency of the low-pass filter and repeat
- SNR vs. frequency of halftoned image relative to edge-sharpened original correlates well with visual quality

SNR vs. Frequency Results

- SNR of halftoned image is high at low frequencies (near the zero of the NTF)
- We radially low-pass filter the edge-sharpened original image and halftoned image and measure the SNR
- Cutoff frequency is increased; SNR computed at each point
- Jarvis shows improvement over Floyd-Steinberg at LF because of aggressive low-frequency noise-shaping; loses out around mid-band noise hump



SNR vs. frequency

Outline

Introduction to Digital Halftoning

- Applications
- Common halftoning methods
- 1-D delta-sigma modulation
- Halftoning with the 1-D DSM

Analysis of Error Diffusion

- Extension of delta-sigma modulation to 2-D
- Linear analysis of the 2-D DSM
- Improved modelling
- Image quality metrics

Design and Implementation

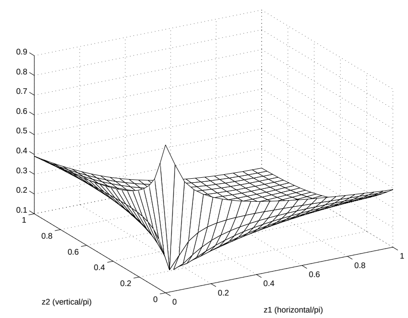
- Filter optimization
- Implementation issues

Conclusions

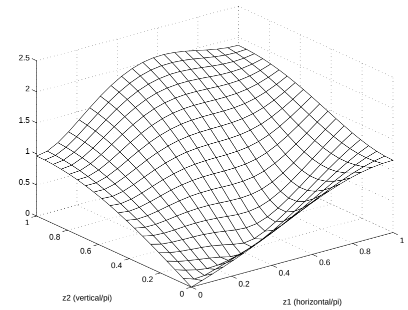
- Future research

Design of the Error Filter

- Error filter weights determine both noise shaping and edge sharpening effects
- The ability to adjust the two independently is desirable (Ulichney, 1987)
- We are using optimization techniques to achieve a target NTF; the gain model then predicts the STF
- Optimize filter according to a human visual criterion
- Choice of target response greatly affects filter results

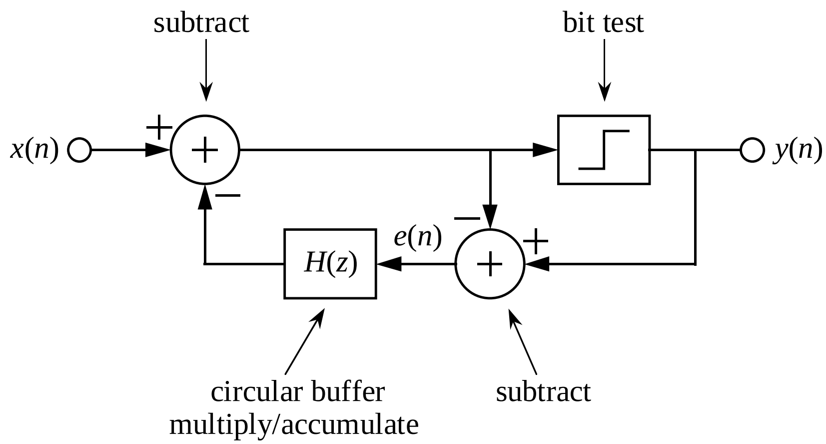


Possible weighting scheme



Resulting NTF

Implementation



- Floyd-Steinberg requires approximately four multiplies, six additions and a bit test per pixel, plus some circular addressing (approximately 7 cycles per pixel on a pipelined general purpose DSP)
- Screening requires a single comparison (subtraction - approximately 1 cycle per pixel on an equivalent processor)
- For small error filters, error diffusion is manageable in real-time using a low-cost digital signal processor
- We are also investigating parallel hardware implementations

Summary

- Error diffusion well-established but not fully understood
- Results from 1-D delta-sigma modulation are now being applied to halftoning; many analogies apparent
- Modelling of the quantizer is necessary
- Gain block model gives good results and predictions
- SNR vs. frequency correlates with visual performance
- Response of HVS should guide error filter design
- Edge enhancement is predictable and should be adjustable independent of noise shaping
- Implementation issues must be addressed to make error diffusion viable in commercial products

Future Work

- Error filter family
  - A set of error filters for different requirements
  - Varying supports, wordlengths and edge sharpening effects
- Extension to the oversampling case
  - 2-D interpolation and halftoning should be combined
  - Requirements on error filter are different from non-oversampled case
  - Smoothing of simple interpolation schemes can be counteracted by edge-sharpening halftoning algorithms, allowing fast implementation
- Extension to video sequences
  - Human spatio-temporal contrast sensitivity function used to optimize error filter in three dimensions (Hilgenberg *et al.*, 1994)
  - Applications in real-time video and low-cost displays