High Quality MRC Document Coding [1]

Guotong Feng and *Charles A. Bouman*
Electrical and Computer Engineering
Purdue University
West Lafayette IN 47907

Hui Cheng
Sarnoff Corporation
Visual Information Systems
Princeton, NJ 08543-5300

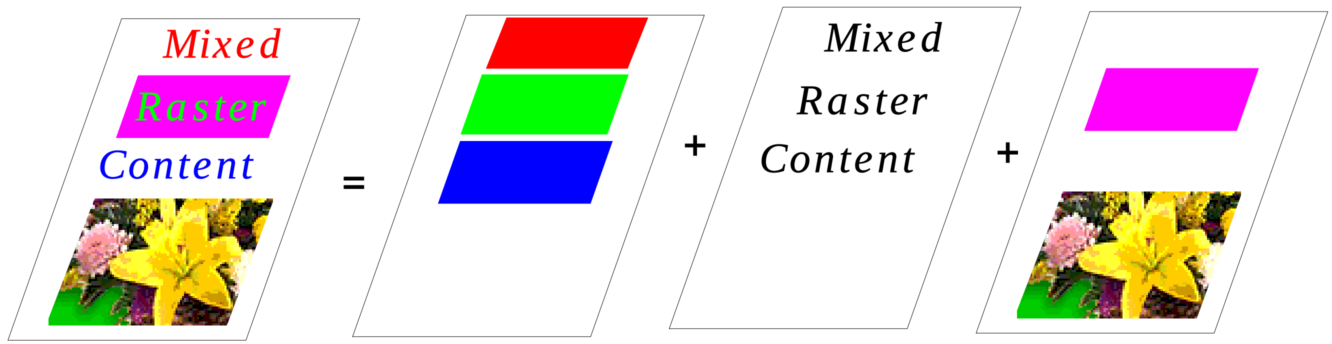
http://www.ece.purdue.edu/~bouman

[1] This project is supported by the Xerox Foundation.

Document Image Compression

- High quality raster documents
  - Require 600dpi scanning using 24 bits per pixel.
  - Result in 45-90 Megabytes per page
- Natural image compression algorithms perform very poorly on documents.
  - At moderate bit rates (20:1 compression) text and graphics are badly degraded.
  - Ringing artifacts look very bad with text.
  - Text is high resolution/low bit depth
  - Pictures are low resolution/high bit depth
- Document compression must function well at very low bit rates (80:1 compression).

Segmentation Based Document Compression



- Mixed Raster Content (MRC)
  - Document is typically separated into three layers: foreground, background, and binary mask.
  - Each layer is coded independently using conventional image coder: JPEG, wavelet, CCITT Group 4, or JBIG(2).
  - MRC part of JPEG2000 and ITU T.44 standards.
- Block-based segmentation
  - Segment non-overlapping blocks of pixels into different classes.
  - Compress each block differently according to its characteristics.
  - More efficient than MRC.

How to Compute Segmentation?

Computing Segmentation using RDOS

- $R_i(x_i|x_{i-1}) = 0$ - Number of bits required to encode block i using class x_i.
- $D(x_i) = 0$ - Distortion produced by encoding block i using class x_i.
- "Optimum" segmentation is given by

$$x^* = \arg \min_{x \in \mathcal{N}^L} \sum_{i=1}^{L-1} R_i(x_i|x_{i-1}) + \lambda D_i(x_i) \qquad (1)$$

- Can be solved using dynamic programming.
- Requires detailed formulation of rate and distortion for each class.

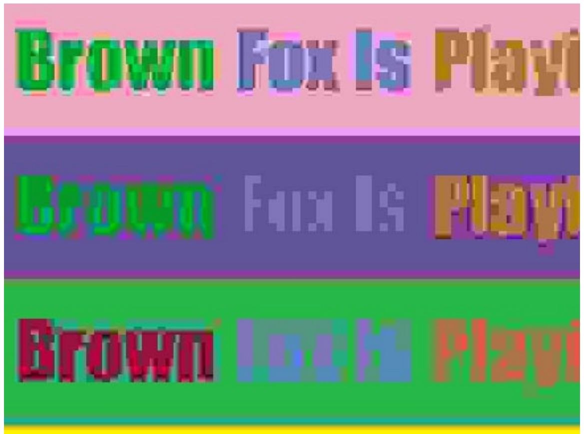
Comparison of Results I

Original

RDOS, 0.101 bpp

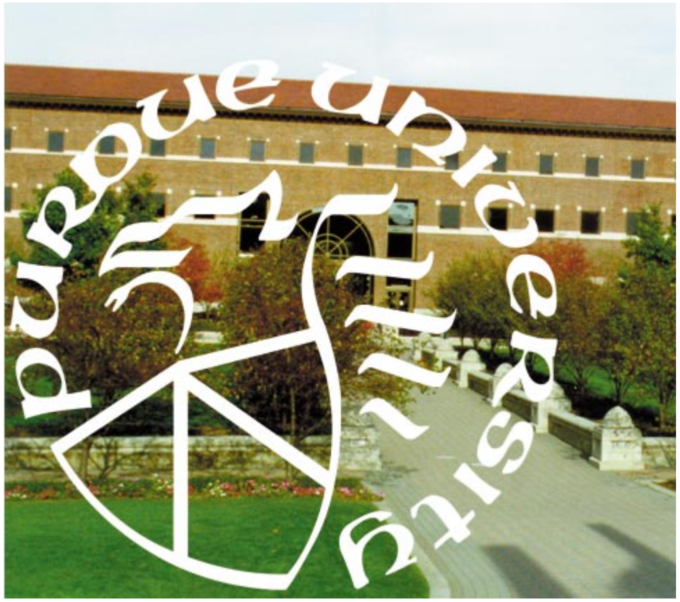
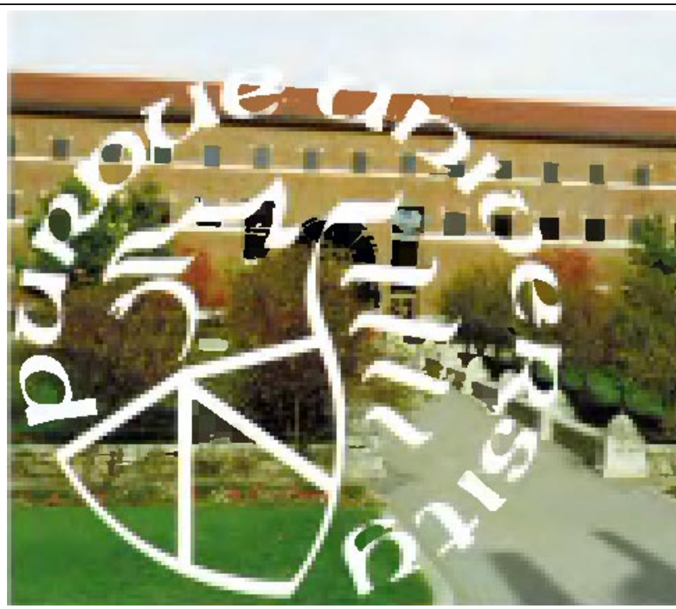
DjVu, 0.103 bpp

SPIHT, 0.103 bpp

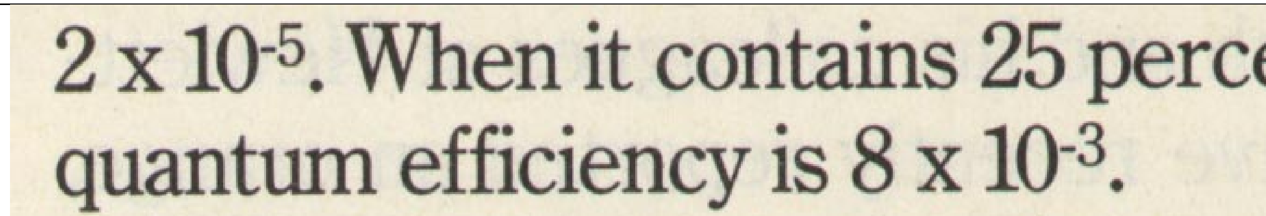
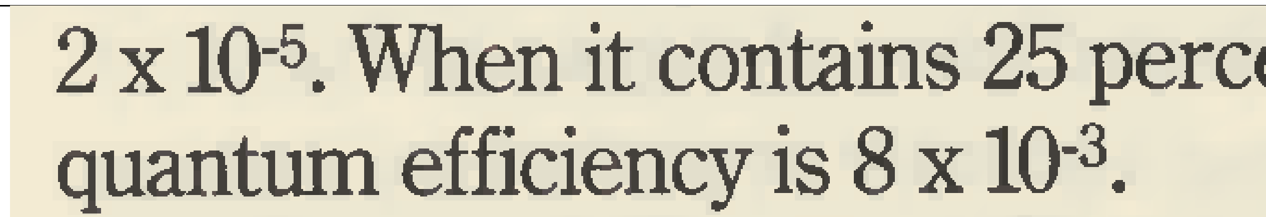
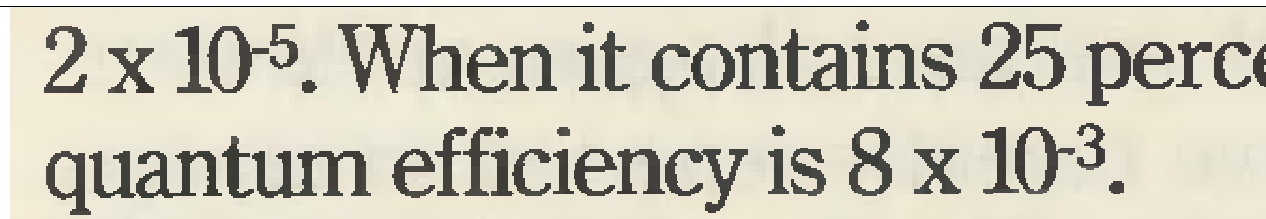
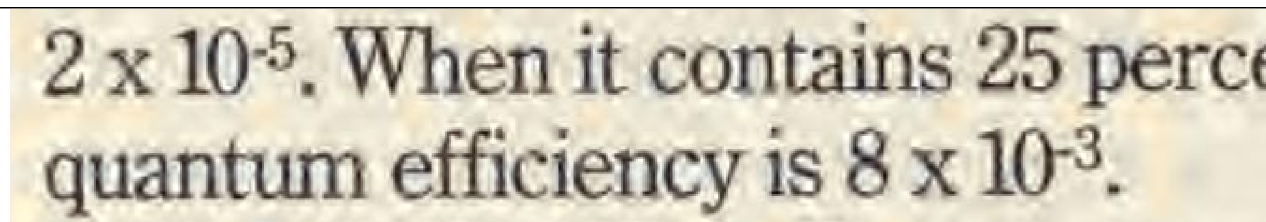


JPEG, 0.184 bpp

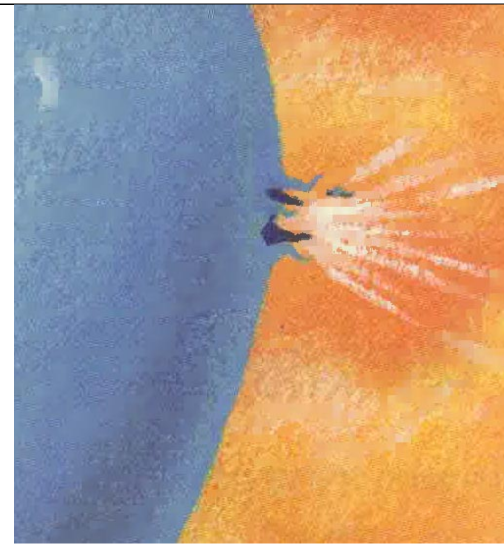
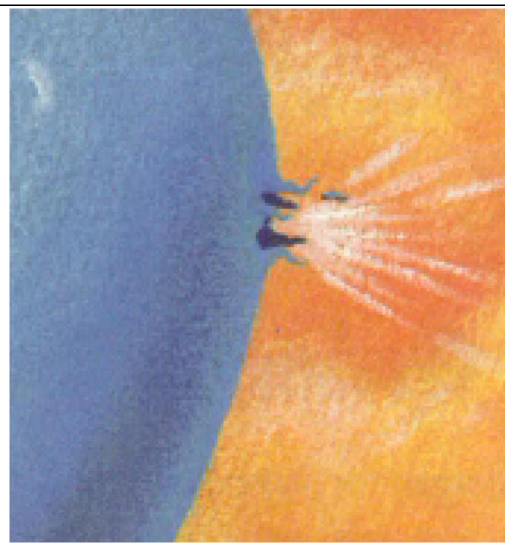
Comparison of Results II

Org			RDOS
DjVu			SPIHT

Comparison of Results III

Original	 2 x 10^{-5}. When it contains 25 perce quantum efficiency is 8 x 10^{-3}.
RDOS 0.114 bpp	 2 x 10^{-5}. When it contains 25 perce quantum efficiency is 8 x 10^{-3}.
DjVu 0.114 bpp	 2 x 10^{-5}. When it contains 25 perce quantum efficiency is 8 x 10^{-3}.
SPIHT 0.114 bpp	 2 x 10^{-5}. When it contains 25 perce quantum efficiency is 8 x 10^{-3}.

Comparison of Results IV

Original			RDOS 0.125 bpp
DjVu 0.132 bpp			SPIHT 0.125 bpp

What is the Problem?

(b)original image scanned at 400dpi

(a)compressed by Djvu at 400dpi, 473:1 compression ratio

- Binary mask makes hard selection between foreground and background pixels.
- This causes “jaggies” and other artifacts in document.

Resolution Enhanced Rendering (RER) Approach

- RER Encoder
  - Encode gray level mask using adaptive dithering.
  - Adaptive dithering is controlled by four parameters.
- RER Decoder
  - Decode gray level mask using tree-based nonlinear predictor.
  - Nonlinear predictor is controlled by parameters of decision tree.
- Joint Encoder/Decoder Optimization
  - Design decoder to recognize patterns of encoder.
  - Design encoder to produce recognizable patterns.
  - Jointly optimize encoder/decoder to minimize distortion.

Resolution Enhanced Rendering Method

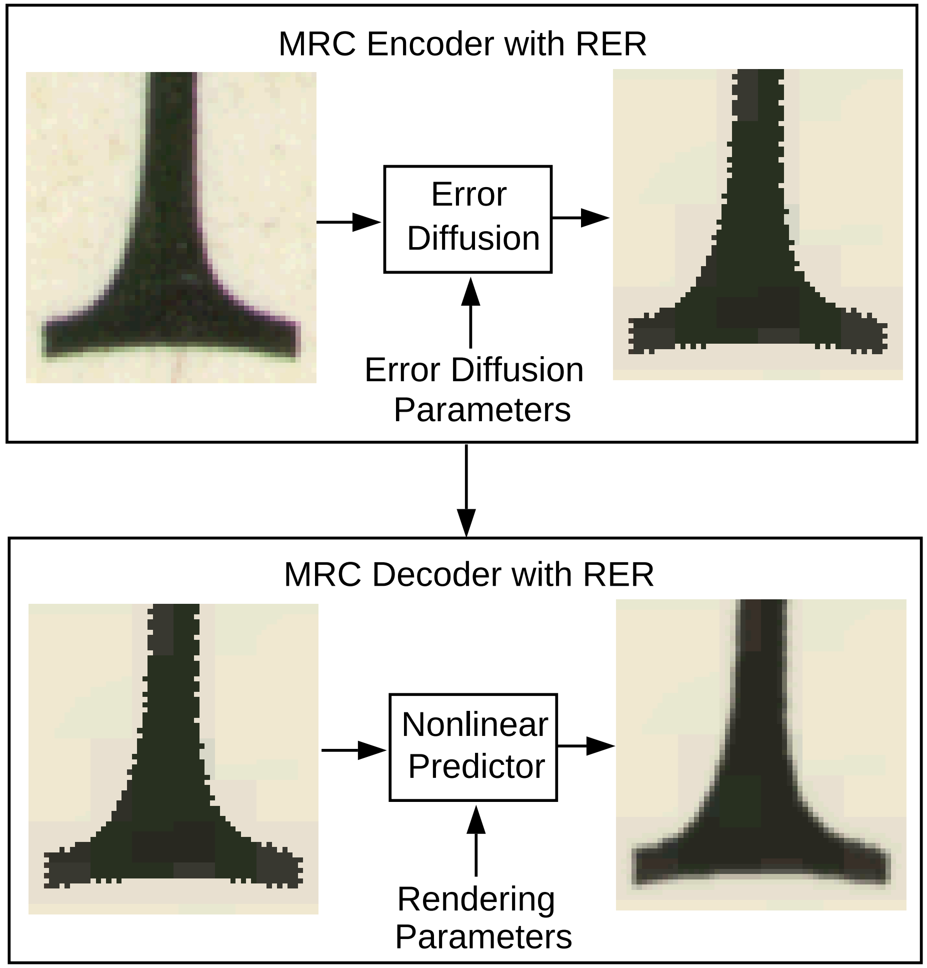
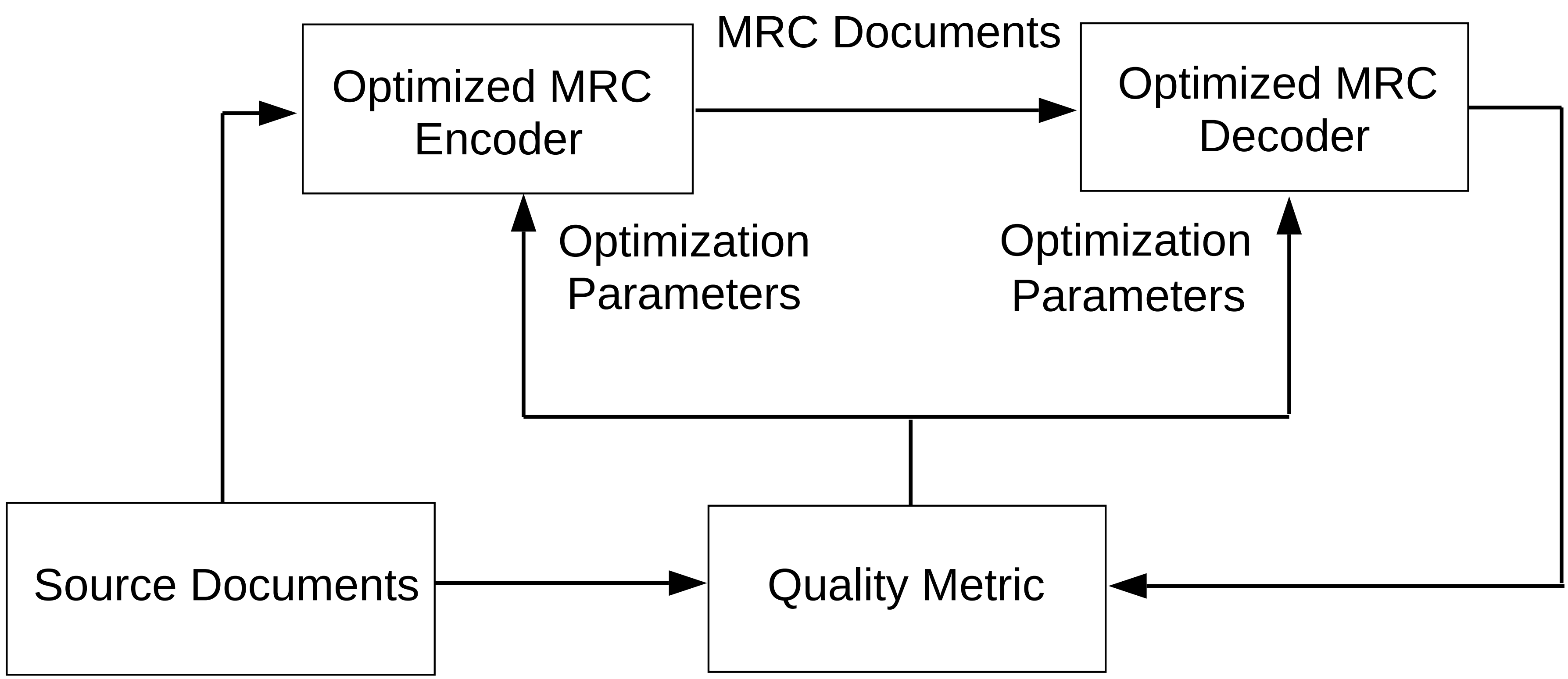


Illustration of MRC encoder and decoder with RER.

Optimization for Encoder & Decoder with RER

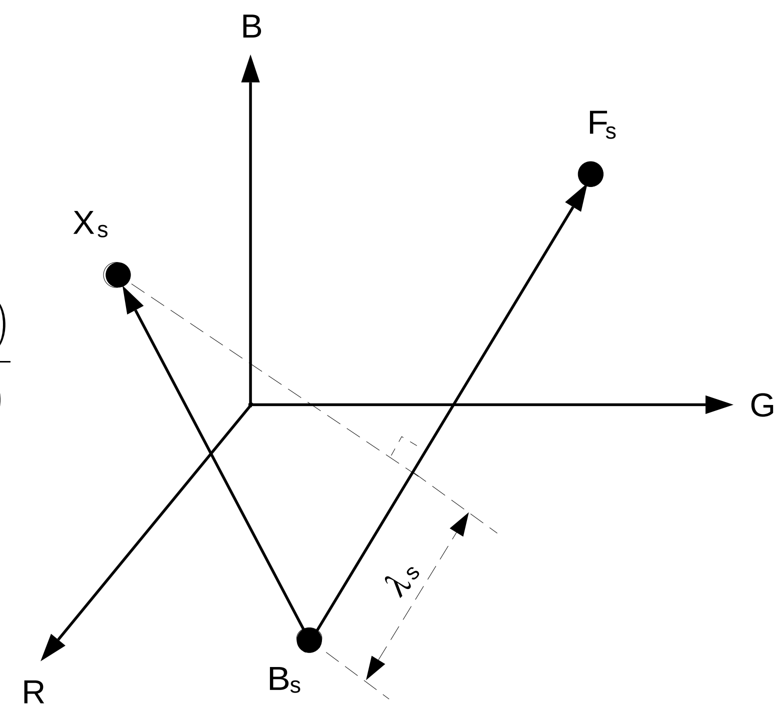


The parameters of RER encoder and decoder are jointly optimized to maximize the quality of the decoded document.

Computing "Gray" Level Mask

1. Project pixel onto line connecting foreground and background color.

$$\lambda_s = \frac{(X_s - B_s)^t (F_s - B_s)}{(F_s - B_s)^t (F_s - B_s)}$$



2. Clip the result to the range $[0, 1]$.

$$\gamma_s = \min\left\{1, \max\left\{0, \lambda_s\right\}\right\}$$

3. Compute the error of a binary mask at each pixel.

$$\Delta_s = \min\{\gamma_s, 1 - \gamma_s\}$$

The Gray Level Mask Encoder

- Binary mask is dithered using error diffusion.
- Error diffusion weights are varied according to:

$$w_{s_j} = \frac{\alpha_j \Delta_{s_j} u(\Delta_{s_j} - \tau)}{\Sigma_{j=0}^{3} \alpha_j \Delta_{s_j} + 0.001}$$

  Δ_{s_j} - binary mask error at pixel s_j

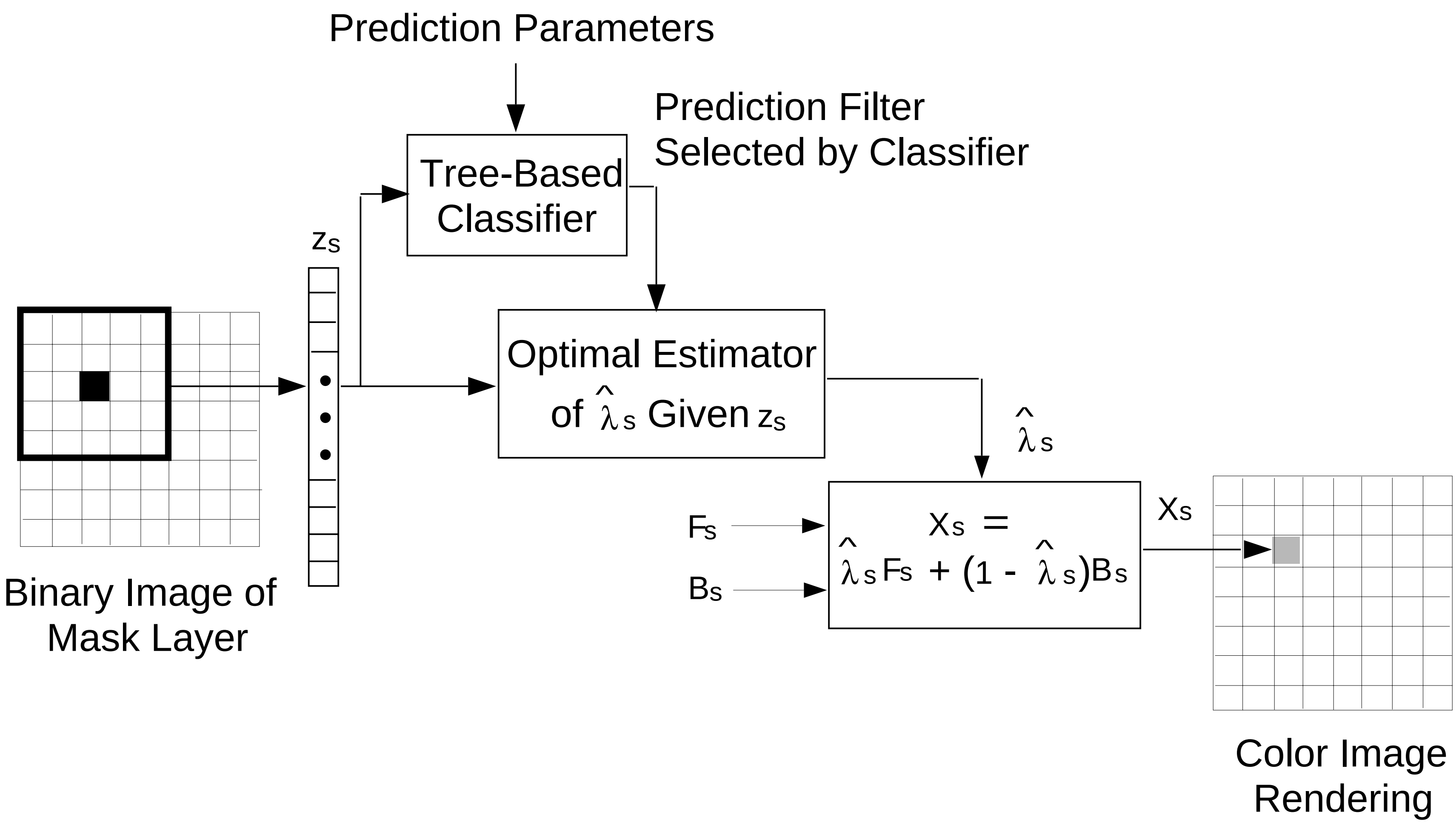
- Adaptive error diffusion:
  - Weights change at each pixel.
  - Diffuses error along boundaries of binary mask.
  - $(\tau, \alpha_0, \alpha_1, \alpha_2, \alpha_3)$ - parameter parameterize algorithm

The RER Decoder

- For each pixel in image:
  1. Extract a 5×5 window of the binary mask.
  2. Use tree structured predictor to estimate gray level mask, $\hat{\lambda}_s$. (Atkins, Bouman, and Allebach '99)
  3. Use gray level mask value to reconstruct pixel

$$\hat{X}_s = \hat{\lambda}_s F_s + (1 - \hat{\lambda}_s) B_s$$

The RER Decoder (Cont'd)

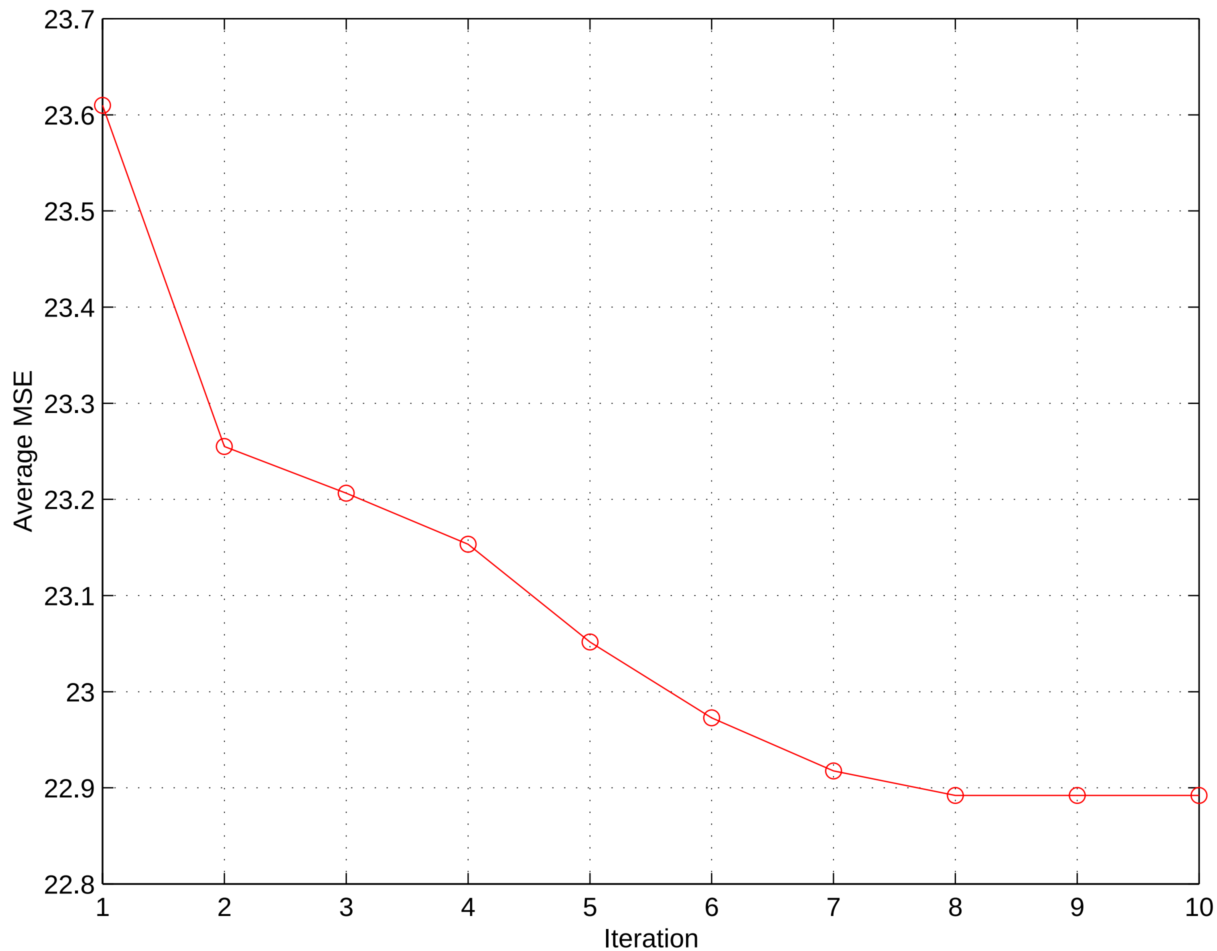


Structure of RER decoder using nonlinear predictor.

Approach to Training and Testing

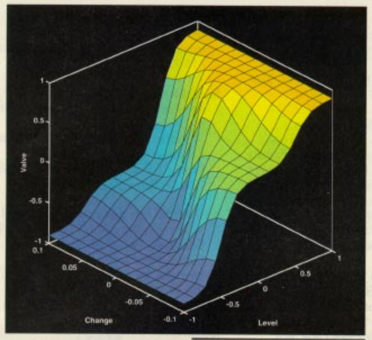
- Experimental data
  - Encoder training data - 5 documents scanned at 400dpi
  - Decoder training data - 16 documents scanned at 400dpi
  - Testing data - combination of scanned and synthetic images.
- The training process alternated between optimization of the encoder and decoder parameters.
- The distortion metric is mean squared error.
- Pixels with significant binary mask approximation errors were selected.

Convergence



Axis y - minimum average MSE of testpages used for encoder training after each loop of joint optimization.
Axis x - Iteration number of the joint optimization for encoder and decoder.

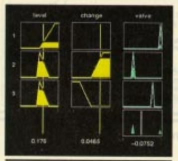
Scanned Test Image



This surface shows the response of a fuzzy controller for a distillation column intake valve. The system was implemented with three simple rules using the MATLAB Fuzzy Logic Toolbox.

MATLAB® Fuzzy Logic Toolbox, the clear path to sharper solutions.

The MATLAB Fuzzy Logic Toolbox is the first product that lets you rapidly learn, evaluate, and integrate fuzzy logic into real world engineering solutions.



The exclusive Fuzzy Inference Diagram lets you interactively explore fuzzy rules and rapidly analyze system behavior.

Fuzzy logic in a complete engineering environment

The Fuzzy Logic Toolbox is the heart of a complete environment for fuzzy system design, dynamic simulation, and implementation.

Unlike fuzzy-only packages, the Toolbox lets you apply the power of fuzzy logic to your everyday engineering work—whether you're exploring new ideas, simulating alternative designs, or generating real-time code.

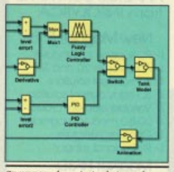
Clearly intuitive GUI

The Fuzzy Logic Toolbox's simple point-and-click user interface guides you effortlessly through the steps of fuzzy design, from setup to diagnosis. The interactive graphics let you instantly visualize and fine tune system behavior.

Focused power and flexibility

With the Fuzzy Logic Toolbox, you get built-in support for the latest methods such as fuzzy clustering and adaptive neuro-fuzzy learning. Using the high-level MATLAB language, you can easily modify any Toolbox function, and add your own.

You can even combine fuzzy logic with other MATLAB Toolboxes to create custom solutions for your applications.


SIMULINK's dynamic simulation tools simplify the design and development of real-time fuzzy systems.

Integrated simulation and code generation with SIMULINK®

The Toolbox is the only fuzzy software that is integrated with a full-function dynamic simulation system. SIMULINK lets you easily simulate fuzzy designs, compare their performance to conventional control or signal processing methods, and generate C code for embedded processors—all within a single interactive environment.

MATLAB
Picture the Power

NEW Fuzzy Logic Toolbox

- *Comprehensive fuzzy design tools*
- *Innovative GUI for design, analysis, and debugging*
- *Advanced methods for adaptive systems and pattern recognition*
  - *Fuzzy clustering*
  - *Neuro-fuzzy learning*
  - *Sugeno fuzzy inferencing*
- *Integrated with SIMULINK and MATLAB Toolboxes, such as:*
  - *Control design*
  - *Signal and image processing*
  - *Neural networks*
- *Code generation for standalone and embedded real-time systems*

The Ultimate Technical Computing Environment™

MATLAB is an extensible, interactive Technical Computing Environment that seamlessly integrates computation, visualization, and modeling on PCs and workstations. Now, it's the ideal way to bring the power of fuzzy logic into focus.

Call today for your free fuzzy logic information package.

508-653-1415

The MATH WORKS Inc.

24 Prime Park Way / Natick, MA 01760
Tel: 508/653-1415 Fax: 508/653-6284
Email: info@mathworks.com

The MathWorks is represented in the following countries:
Australia: +61-2-922-6311 • Brazil: +55-11-816-3144
France: +33-1-45-34-23-91 • Germany: +49-241-26041
India: +91-80-2-260-260 • Israel: +972-3-561-5151
Italy: +39-11-24-85-332 • Japan: +81-3-5978-5410
Korea: +82-2-517-1257 • Portugal: +34-3-415-49-04
Scandinavia: +46-8-15-30-22 • Spain: +34-3-415-49-04
Switzerland: +41-31-998-44-11 • Taiwan: +886-2-501-8787
For Belgium, Luxembourg, The Netherlands, United Kingdom and Republic of Ireland call Cambridge Control, Ltd: +44-1223-423-200 or Rapid Data, Ltd: +44-1903-821-266

217v0

Circle No. 9

Synthetic Test Image

Example 1

Original Image

Example 1

Standard RDOS

Example 1

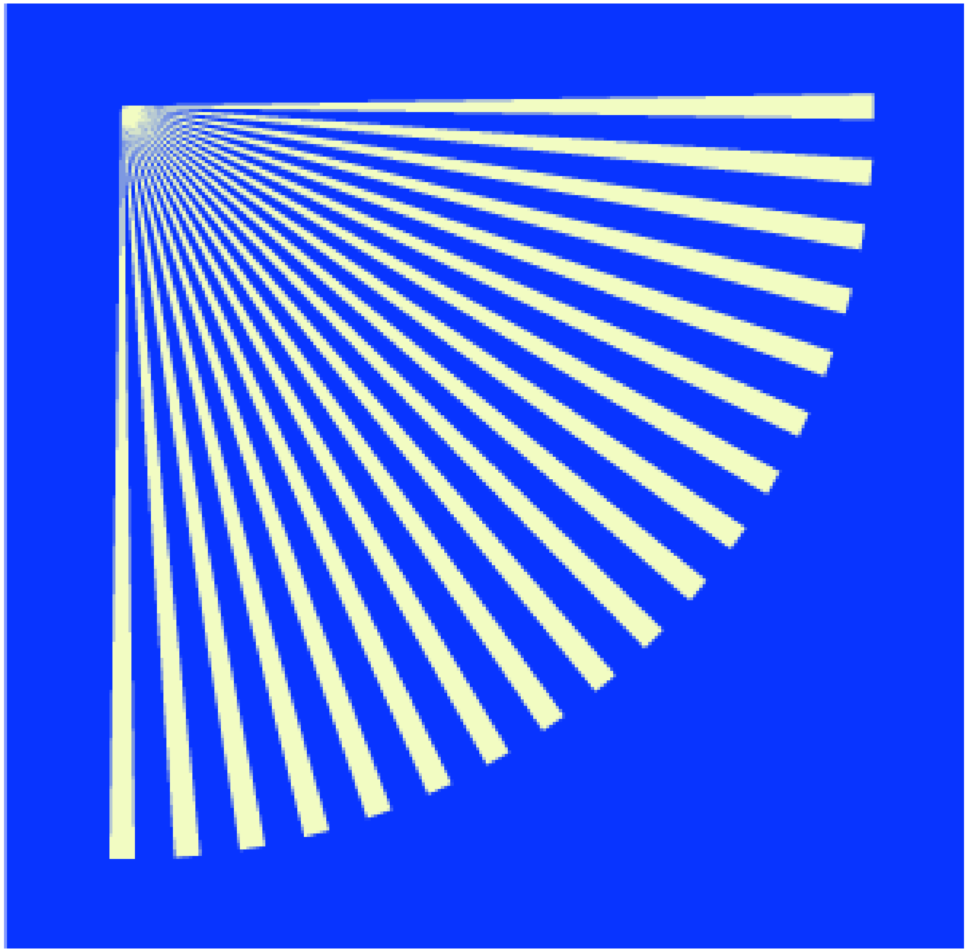
Resolution Enhanced RDOS

Example 1

Original image

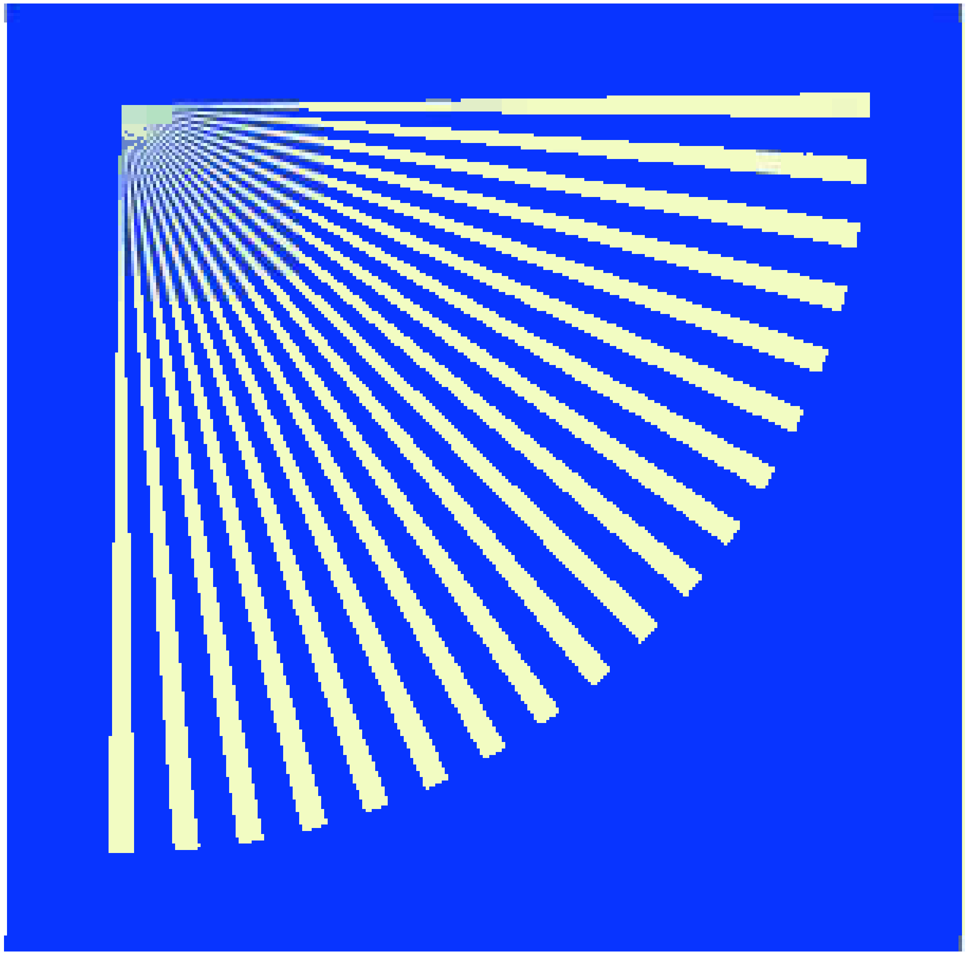
Resolution Enhanced RDOS

Example 2



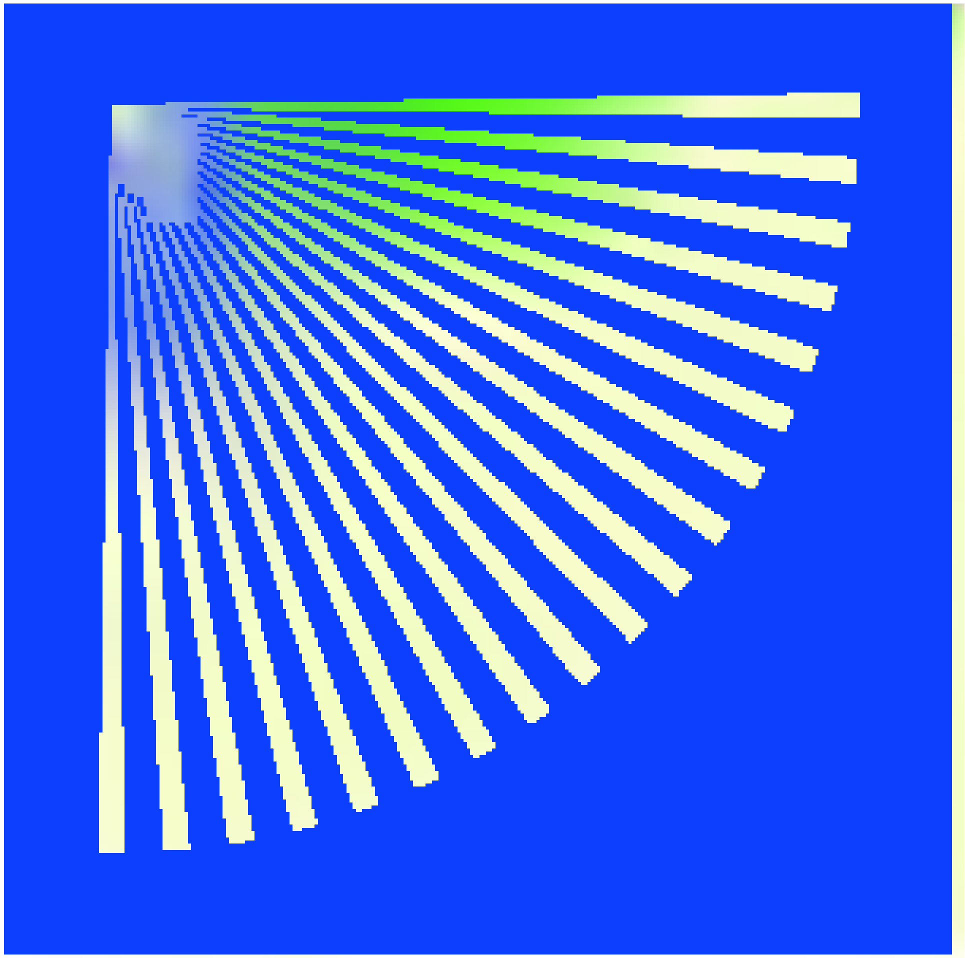
Original

Example 2



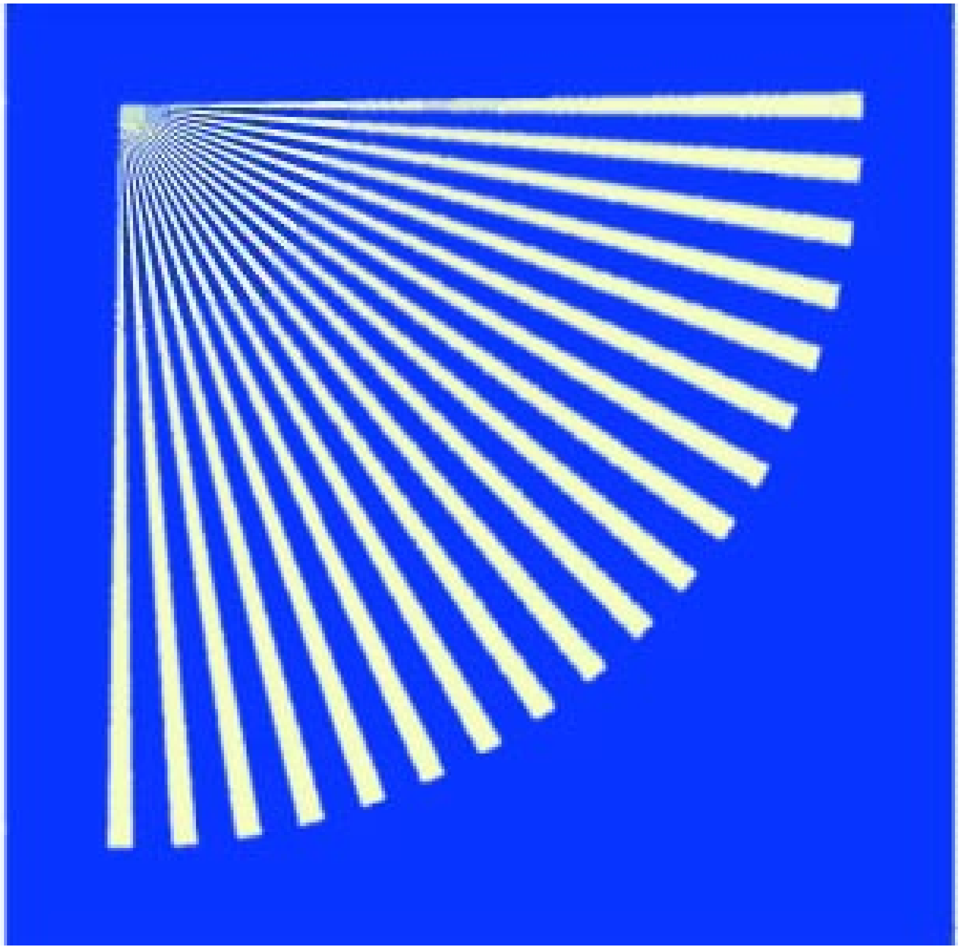
RDOS

Example 2



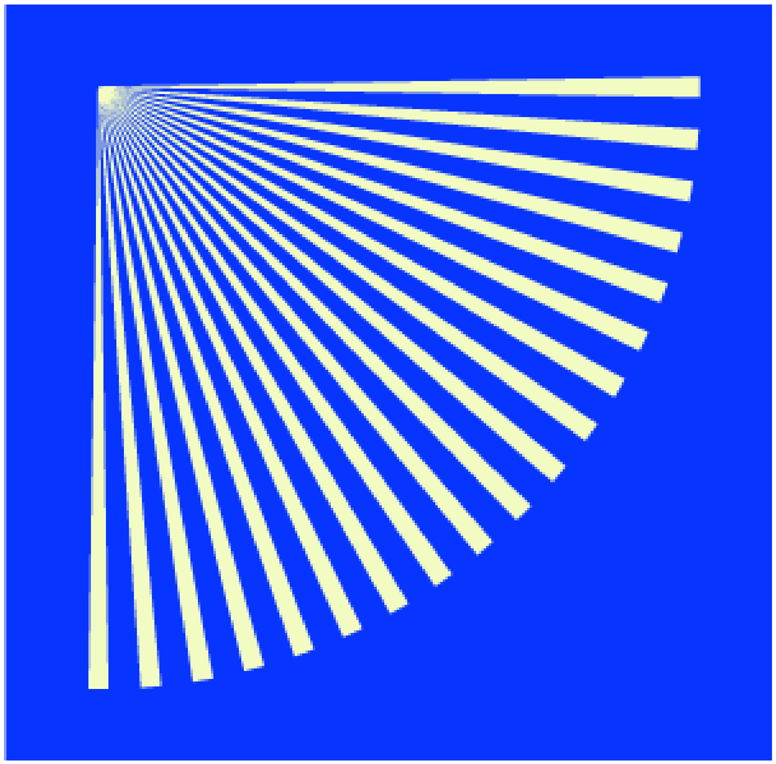
DjVu

Example 2

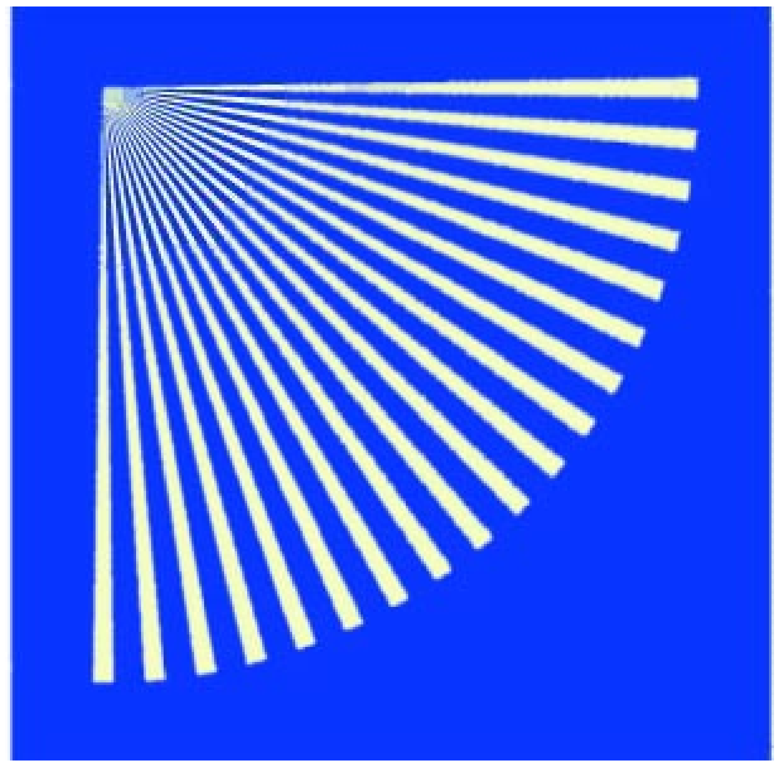


Resolution Enhanced RDOS

Example 2

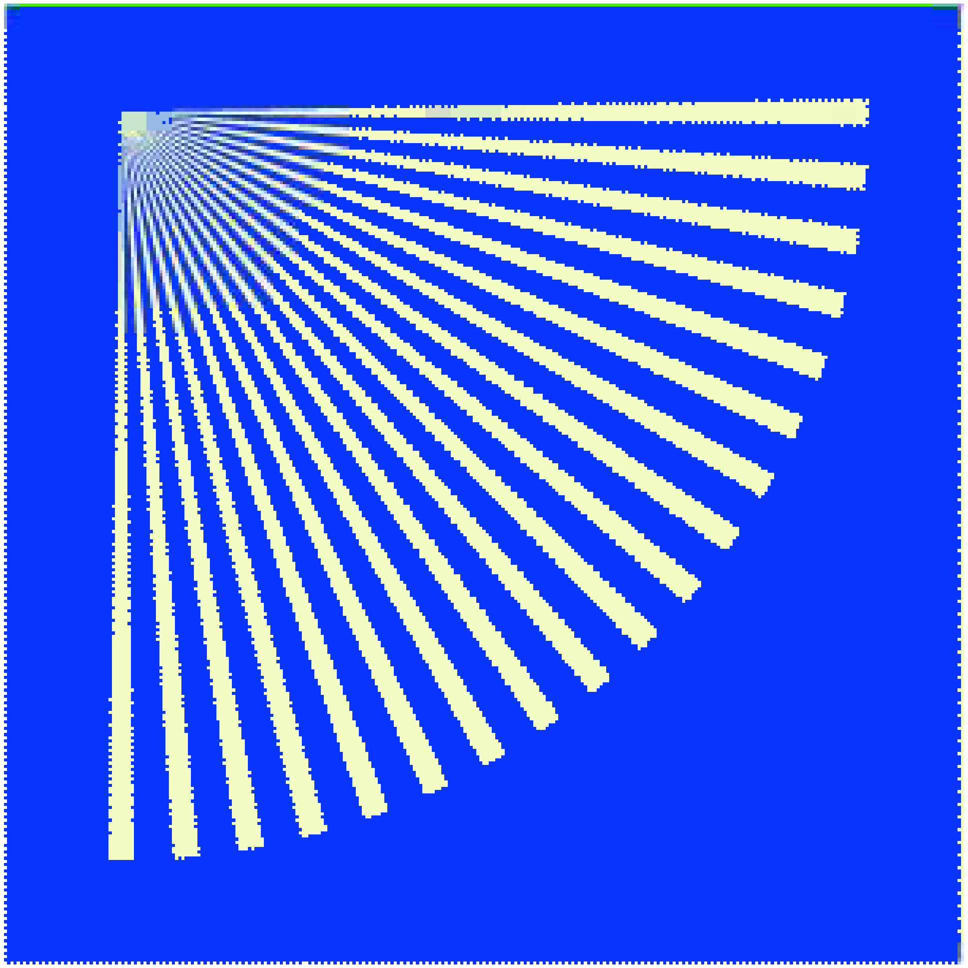


Original



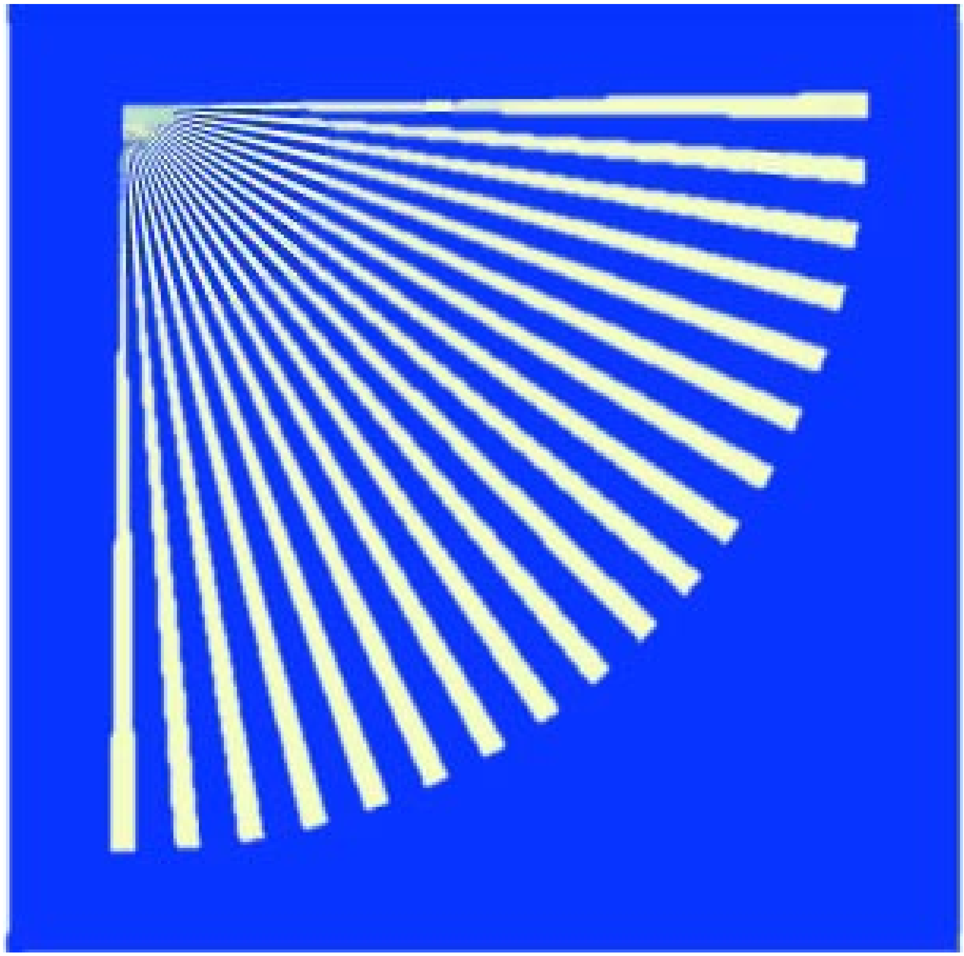
Resolution Enhanced RDOS

Example 2



RER Encoder Only

Example 2

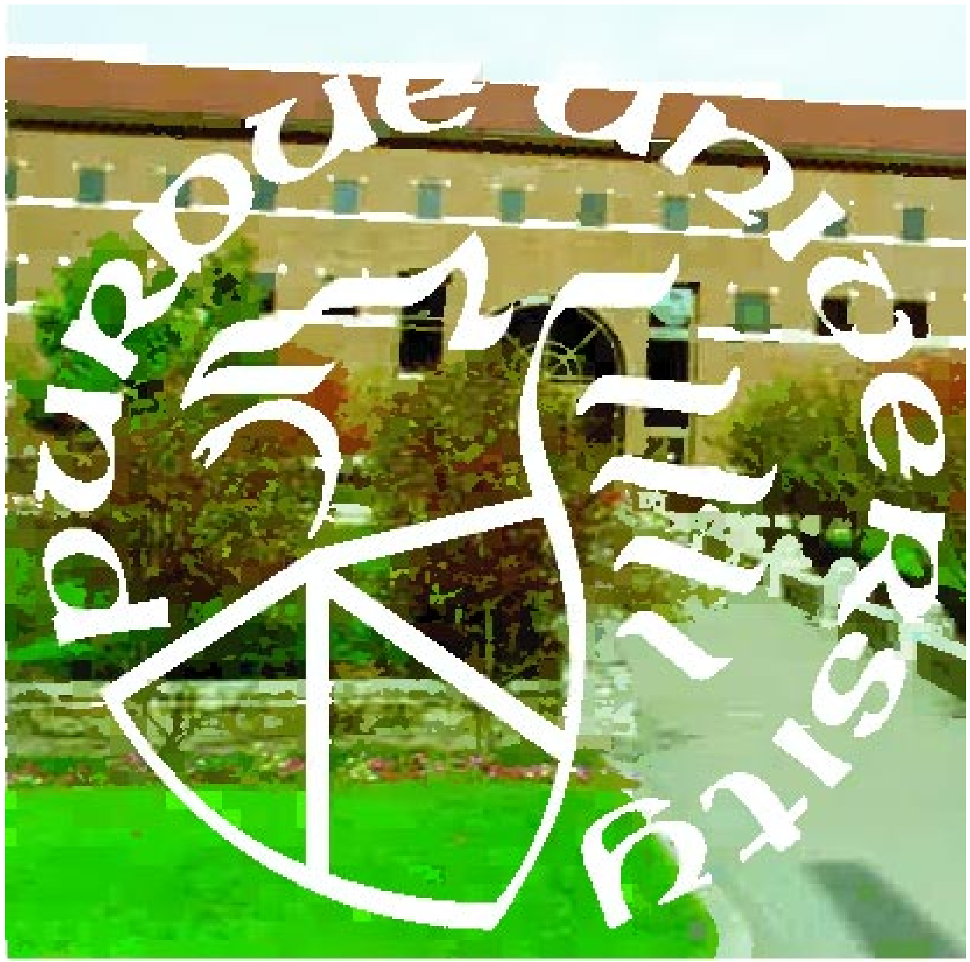


RER Decoder Only

Example 3

Original

Example 3

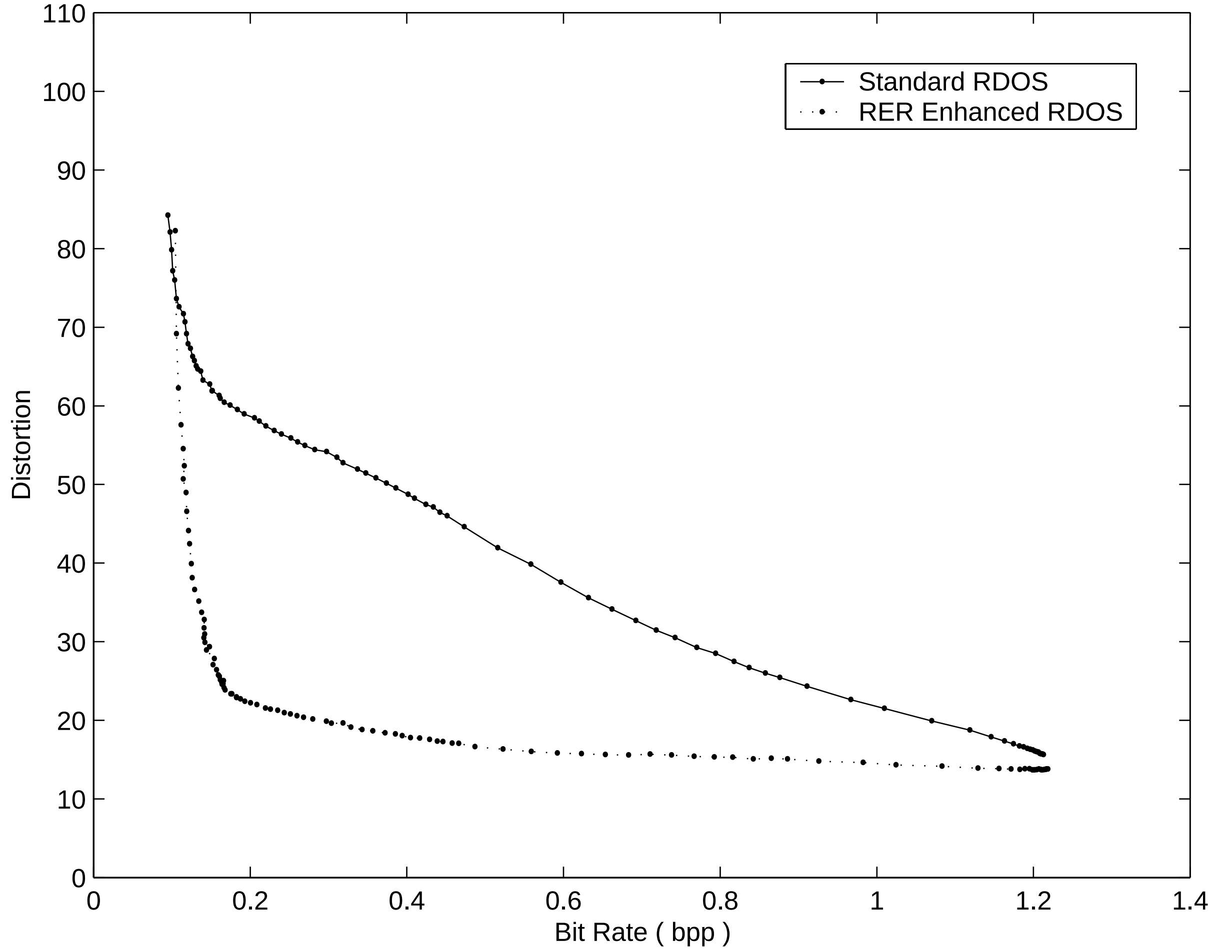


RDOS

Example 3

Resolution Enhanced RDOS

Experimental Results



R-D performance of RER enhanced RDOS and standard RDOS.

Conclusion

- Binary masks can cause significant artifacts in coded raster documents.
- The RER method:
  - Can improve document quality by efficiently encoding mask transitions.
  - Is backward compatible with standard MRC coders.
  - Is most effective after joint optimization of both encoder/decoder.