Show-through Cancellation in Scanned Images Using Blind Source Separation Techniques

Boaz Ophir and David Malah

Department of Electrical Engineering, Technion - Israel Institute of Technology

Contact:

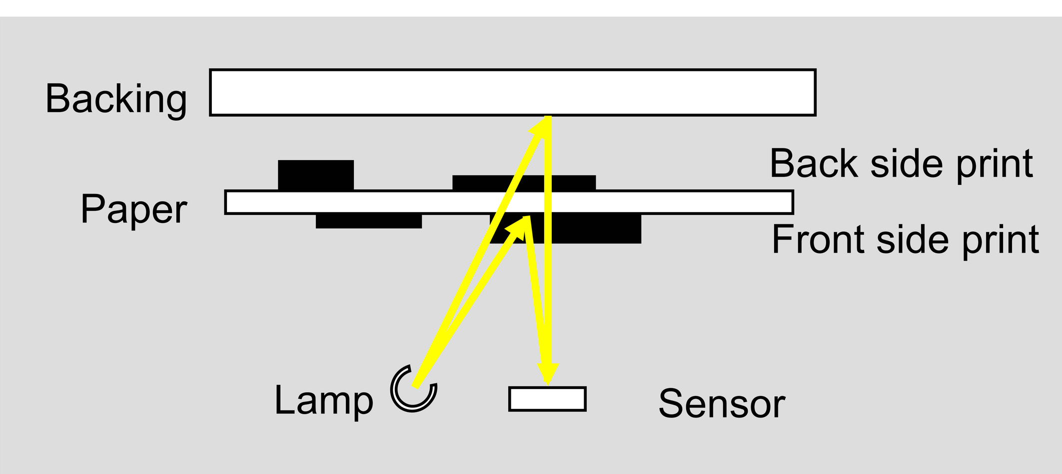
Boaz Ophir boazo@il.ibm.com

David Malah malah@ee.technion.ac.il

Abstract

- Show-through is a common occurrence when scanning duplex printed documents. The back-side printing shows through the paper, contaminating the front side image.
- Previous work modeled the problem as a *non-linear convolutive mixture* of images.
- We propose a cleaning process based on a Blind Source Separation (BSS) approach.
- We define a cost function incorporating the non-linear mixing model in a mean-squared error term, along with a regularization term based on Total-Variation.
- Location-dependent regularization tradeoff preserves image edges while removing show-through edges.
- Images and mixing parameters are estimated using an alternating minimization process, with each stage using only convex optimization methods.
- The resulting images exhibit significantly lower show-through, both visibly and in objective measures.

Physical Model



Scanned reflectance value:

$$R_f^s = R_p T_f^2 + T_f^2 T_p^2 R_{bk} T_b^2$$

Desired reflectance value:

$$R_f = R_p T_f^2 + T_f^2 T_p^2 R_{bk}$$

R	reflectance
T	transmittance
f,b	front/back images
p	paper
bk	scanner backing
s	scanned images

Definitions:

Reflectance of white paper:

$$R_p^w = R_p + T_p^2 R_{bk}$$

Normalized optical density:

$$D = -ln(R/R_p^w)$$

Normalized optical absorptance:

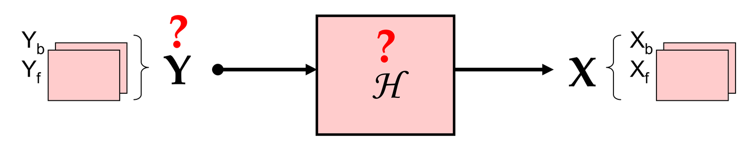
$$A = 1 - R/R_p^w$$

Linearized Creation Model:

$$D_f^s(x,y) \cong D_f(x,y) + h(x,y) \otimes A_b(x,y)$$

G. Sharma, "Show-Through Cancellation in Scans of Duplex Printed Documents", IEEE Transactions on Image Processing, 2001

Show-through as a BSS Problem



- Mixing operator $\mathcal{H}$ contains both non-linear and convolutive elements

$$\mathcal{H} = \begin{bmatrix} 1 & h_{12} * f(\cdot) \\ h_{21} * f(\cdot) & 1 \end{bmatrix}$$

- Goal – Estimate images **Y** and mixing parameters $\mathcal{H}$
- Method – Minimization of a cost function:

$$\mathbf{J} = \|\mathcal{H}(\mathbf{Y}) - \mathbf{X}\|_2^2 + \lambda Reg(\mathbf{Y})$$

Fidelity Term — Regularization

Choice of Regularization Function

- Assumption: Images Y_f and Y_b are independent

$$Reg(\mathbf{Y}) = \lambda_1 Reg(Y_f) + \lambda_2 Reg(Y_b)$$

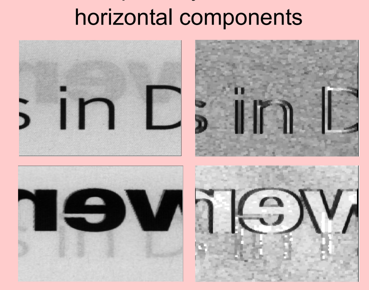
- Regularization functions:
  - Smoothness — ☺ Simple, ☹ Oversmoothing
  - Markov Random Fields — ☺ Edge-preserving, ☹ Many parameters, ☹ Non-convex
  - Total-Variation — ☺ Edge-preserving, ☺ Convex

Fidelity/Regularization Tradeoff

$$\mathbf{J} = \|\mathcal{H}(\mathbf{Y}) - \mathbf{X}\|_2^2 + \lambda_1 TV(Y_f) + \lambda_2 TV(Y_b)$$

- Weighting factor λ determines strength of TV smoothing
- In the context of image restoration, Strong *et al* (1997) suggest setting location-dependent λ inversely proportional to gradient strength to preserve image edges
- Problem: Show-through interference contains edges itself
- Solution: Weighting is set using both images
- Strong edge in one image and weak edge at the same location in the other image => probable show-through scenario
  - Set **high** λ at show through edge
    - Proportional to true (reverse side) edge strength
  - Set **low** λ at true edge
    - Inversely proportional to edge strength

λ is set separately for vertical and horizontal components



Cost Function Minimization

$$\mathbf{J} = \|\mathcal{H}(\mathbf{Y}) - \mathbf{X}\|_2^2 + \lambda_1 TV(Y_f) + \lambda_2 TV(Y_b)$$

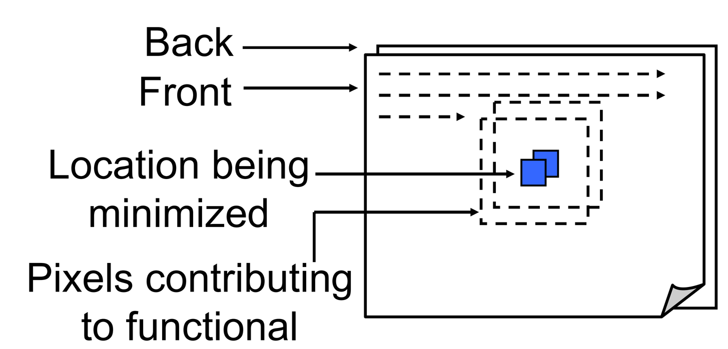
- Simultaneous minimization is very difficult:
  - **Non-linear** in joint space { $\mathcal{H}$,**Y** }
  - **Non-convex** in joint space { $\mathcal{H}$,**Y** }
  - **Large** number of variables
- Reduction method:

  Alternating minimization in image space **Y** and mixing parameters $\mathcal{H}$

  - **Linear** optimization problems
  - **Convex** cost functional

Stage 1: Image Estimation (Y)

- **Y** optimization is done on a very large number of dependent variables
- We use the **Iterated Conditional Modes** (ICM) method (Besag 86) to minimize the energy functional



- **J** is minimized one pixel pair location at a time (in both images simultaneously)
- Each minimization is done only on partial sum of elements of **J** - only the surrounding pixels take part
- Pixel values updated in place
- Algorithm typically converges after 3 – 5 passes
- To avoid directional preference, we alternate scan direction

Stage 2: Mixing Parameters Estimation ($\mathcal{H}$)

- Assumption: $\mathcal{H}$ is the same throughout the image
- Least squares problem
- Only pixels with single-sided activity are useful for the solution

=> Use activity masks – only pixels with single-sided activity contribute

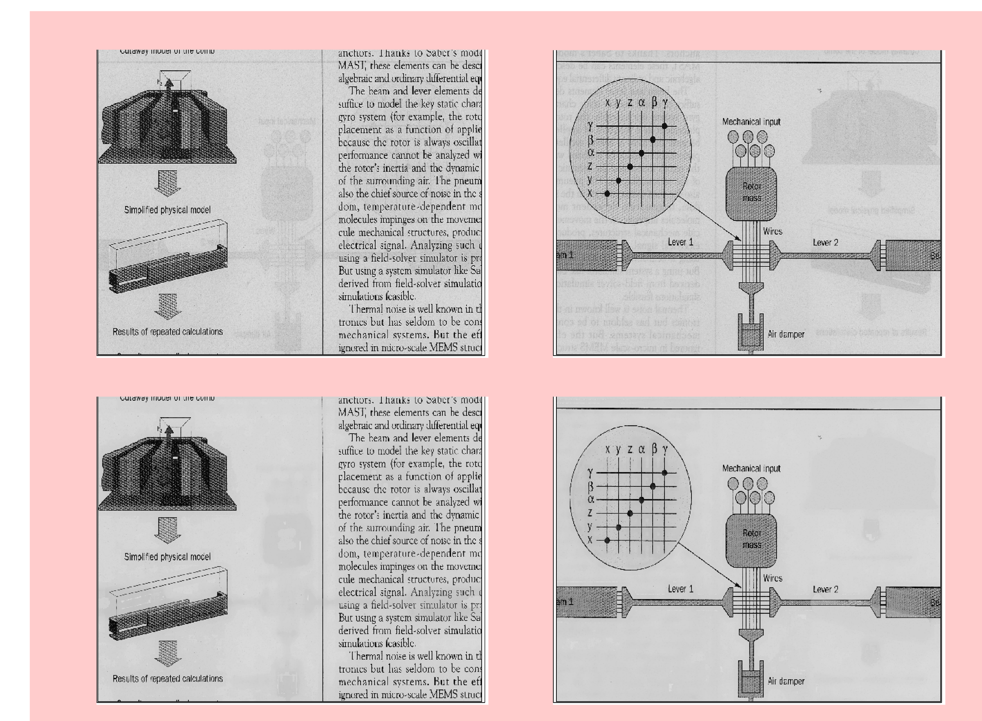
$$\mathbf{J} = \|M(\mathcal{H}(\mathbf{Y}) - \mathbf{X})\|_2^2$$

Activity mask

Minimization Algorithm

- Initialization : $\mathbf{Y} = \mathbf{X}, \quad \mathcal{H} = \begin{bmatrix} 1 & \epsilon \\ \epsilon & 1 \end{bmatrix}$
- Alternating optimization:
  - Y Optimization:
    - Compute λ maps
    - ICM iterations
      - Top-Bottom scan
      - Bottom-Top scan
  - $\mathcal{H}$ Optimization:
    - Compute activity masks
    - Least squares minimization
  - Normalize PSFs and images
- Repeat for larger filter support

Results



Conclusion

- In this work, we showed how the BSS framework can be used for separating complicated image mixtures as demonstrated in the show-through problem.
- We attempt to separate the image through the minimization of a cost function. The non-linear convolutive mixing model is incorporated in the fidelity term, while image independence is reflected in the use of separate prior terms for each image.
- We use Total-Variation regularization coupled with a novel location-dependent scheme for setting the trade-off terms. We thus achieve good edge preservation, while avoiding the use non-convex functions and optimization methods.
- We use an alternating minimization reduction of the problem, applying the ICM optimization method to further simplify the process.
- The combination of all these methods is found to give good show-through cancellation.