

A novel model for show-through in scan of duplex printed documents

M. Ryyan Khan · Md. Kamrul Hasan

Received: 24 February 2010 / Revised: 20 October 2010 / Accepted: 22 October 2010 / Published online: 21 November 2010
© Springer-Verlag London Limited 2010

Abstract In scan of duplex printed documents, show-through is a phenomenon where the printing of the opposite side “shows through” and contaminates the data of the scanning side. Despite the efforts of prior works to model and correct the show-through effect, an accurate mathematical model based on rigorous analysis of the show-through phenomenon is yet to be reported in the literature. In this paper, we present a detailed analysis of the phenomenon and derive a physics-based novel mathematical model for the show-through effect. We follow a probabilistic approach to explain the scattering of photons in the scanning setup of document-backing system. The show-through model, i.e., the relationship between the reflectances of the two sides of the printed document is then derived using the various components of the photon fluxes emerging from those sides. The model is validated using practical show-through data. Finally, the effectiveness of our model is tested by applying it to show-through correction. The results of experimental tests demonstrate superior performance of our model and algorithm when compared to a well-cited work reported in the literature.

Keywords Scanned image · Show-through · Photon scattering in paper · Show-through model · Joint histogram

1 Introduction

In modern times, digitization is a preferred option for the preservation and archiving of documents. The digitization process includes the capture of images through scanning the document of interest. One of the phenomena degrading the image acquired from scanning is known as the *show-through effect*. In the event of scanning duplex printed documents, the scanned side is often contaminated by an image from the printing of the opposite side. This occurrence, known as show-through, does not hinder processing in the case of low quality scans. But in the application of high quality scans, or in the case of digitization method like optical character recognition (OCR), the show-through data can hamper processing. Besides, an image with show-through is also irritating to the human eye. Higher machine readability and improved quality of digitized documentation for electronic database is the major inspiration for the development of show-through correction techniques.

The show-through effect has been analyzed, and a mathematical model representing the phenomenon was derived in [1]. Based on the model, an adaptive filtering technique was used in [1] and [2] for the estimation of a point spread function (PSF). In this process, the back side image was used as an estimation of noise signal for cleaning the front side. Similarly, the back side was then cleaned.

The show-through phenomenon was viewed as a Blind Source Separation (BSS) problem in [3] and [4]. In [3], the correction was applied through simultaneous estimation of the separated images of both sides of the document and the mixing operator. The mean squared error (MSE) fidelity and total-variation (TV) terms were used yielding a non-linear mixing model for the images. The optimization of the images and mixing operator was simplified using the Iterated Conditional Modes (ICM) optimization method [5]. A new

M. R. Khan · Md. K. Hasan
Department of Electrical and Electronic Engineering, Bangladesh University of Engineering and Technology, Dhaka, Bangladesh
e-mail: ryyan.khan.eee@gmail.com

Md. K. Hasan (✉)
Department of Biomedical Engineering, Kyung Hee University,
Kyungki 446-701, South Korea
e-mail: khasan@eee.buet.ac.bd

show-through effect model was assumed in [4] where the nonlinear BSS algorithm proposed in [6] was used for the correction process.

Recently, a wavelet-based show-through correction algorithm was proposed in [7]. This correction process assumed that, for most images, the wavelets with high-frequency coefficients are likely to be sparsely distributed. The second assumption was that, the front side mixture component of the image will be dominant compared to the component of the opposite side. For documents containing mixture of shaded images and texts on both sides, the above-mentioned assumptions may not hold.

In our previous work [8], we presented an empirical model for show-through taking into account the limitations of the model presented in [1]. The cleaning processes was implemented based on the proposed model by calculating the necessary model parameters from the statistical information extracted using the joint histogram.

In this paper, we present a novel mathematical model for show-through. We analyze the scattering of photons in the scanning system consisting of ink in the form of printing, paper substrate and the scanner backing. The scattering of light in printed paper and its effect on the observed reflectance has been studied in [9–13]. We consider the probabilistic approach for light scattering followed in [13] and extend its concept for the modeling of the show-through effect. Then, the derived model is verified using practical data for show-through. Finally, we present a show-through correction algorithm based on this model. The model parameters are estimated from the joint histogram.

The rest of this paper is organized as follows. Section 2 introduces the show-through effect in mathematical form giving insight into how we intend to express the model. Section 3 provides a detailed analysis of photon scattering in the case of show-through and presents the derivation of the mathematical model for a specific region. Section 4 extends the model presented in Sect. 3 for the space varying data of the scanned document. Section 5 verifies the derived model for practically obtained show-through data. The show-through correction algorithm based on the mathematical model is presented in Sect. 6. The results obtained by the application of our algorithm is then presented in Sect. 7. Finally, we give our concluding remarks in Sect. 8. All the necessary derivations needed for our calculations are presented in the appendices.

2 Problem statement

In scanning of one side of a duplex printed document, often the printing of the opposite side shows through. This is known as the *show-through* effect. Let us consider the two sides of a duplex printed document as *side A* and *side B* as shown in Fig. 1. The reflectance detected by the scanner head at

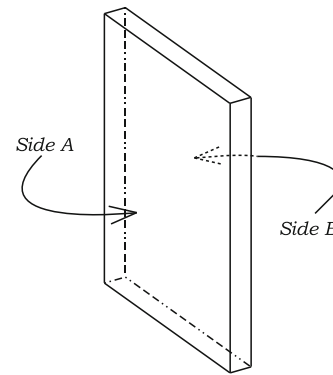


Fig. 1 A sample document

side A is R_A at a *specific region*. The *region* is such that the corresponding reflectances of both the sides are space invariant. The original reflectance (R_A^0) of *side A* is contaminated by the reflectance of *side B* (R_B^0) and thus forms the show-through affected reflectance R_A . Here, R_A^0 is the reflectance found at the scanner head of *side A* when there is no printing on *side B* and R_B^0 is that of *side B*.

Mathematically, the show-through corrupted reflectance observed at the scanner head at *side A* (R_A) would be dependent on R_A^0 and R_B^0 , i.e.,

$$R_A = f_{st}^A(R_A^0, R_B^0) \quad (1)$$

where, $f_{st}^A(\cdot)$ is the function incorporating the effect of show-through for *side A*. Here, the subscript ‘*st*’ denotes ‘show-through’. Similarly, the reflectance observed at the scanner head at *side B* (R_B) would be,

$$R_B = f_{st}^B(R_B^0, R_A^0) \quad (2)$$

where, $f_{st}^B(\cdot)$ is the function associating the show-through effect for *side B*.

In this paper, we explain the effect of show-through and derive expression relating R_A to R_A^0 and R_B^0 . A similar expression for R_B is also given. These relationships are fitted in practical show-through corrupted data to validate our model. Finally, this model is applied to correct show-through and thus strengthen the validity of the model.

3 Model derivation

Figure 2 demonstrates the basic geometry of the surfaces of the two sides for a *specific region* of a printed paper substrate. In Fig. 2, the surface of the *region* selected for *side A* is partitioned into two subsets:

1. Σ_{1A} is the set of localities under the ink dots, showed as the shaded portions in the diagram.

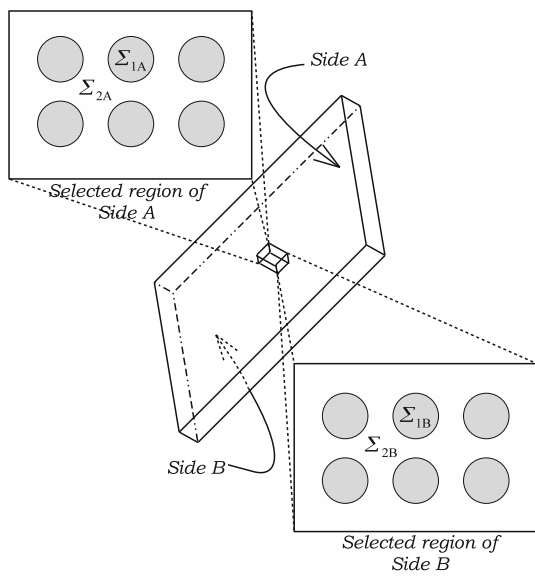


Fig. 2 The halftone image of the *specific region* of the two sides

2. Σ_{2A} is the portions between the ink dots, i.e., Σ_{2A} refers to the localities of bare paper in the specified *region*.

Similarly, we get two subsets in the corresponding selected *region* of *side B*, marked as Σ_{1B} and Σ_{2B} in Fig. 2. These subsets of the specified region can be related as,

$$\frac{\text{Area of } (\Sigma_{1m})}{\text{Area of } (\Sigma_{2m})} = \frac{f_m}{1 - f_m}, \quad m = A, B \tag{3}$$

Here, f_A and f_B are the dot percentages of *side A* and *side B*, respectively.

For the derivation of a show-through model, we extend the concept and method used in [13]. We intend to find the reflectance (R_A) of *side A* as seen by the scanner. For calculation purposes, we consider a region of the page for which the reflectance is constant in both the sides, i.e., the dot percentages of the two sides do not have space variation in the *specific region*. Finally, after deriving a relationship among the reflectances for the *specific region*, we will assume the reflectances to be space variant (i.e., varying in different regions of the paper) and express a general relation for the whole page.

Reflectance of an object is defined as the ratio of the light intensity reflected to the light intensity incident on the surface of the object. In terms of photons, the reflectance is the ratio of the reflected photon flux to the incident photon flux. Consider the setup for scanning *side A* of a document as shown in Fig. 3. The photon flux incident on the surface of *side A* from the source is I_0 . The average photon flux emitting from the region, conforming to our criteria, of this surface is J_A . Therefore, the reflectance of the region of *side A* is,

$$R_A = \frac{J_A}{I_0} \tag{4}$$

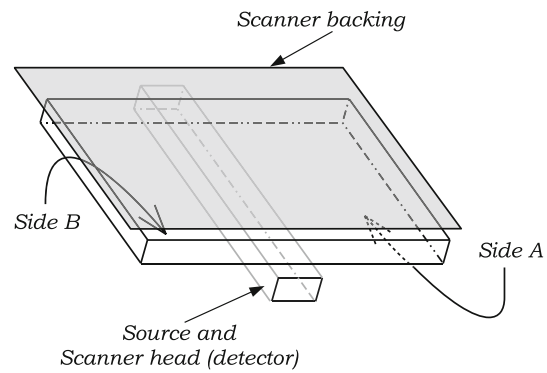


Fig. 3 A setup for scanning *side A*

Now, we can divide J_A into two components: J_{AA} and J_{AB} . J_{AA} is the component of photon flux J_A without taking into account any interaction of *side B*. J_{AB} is the component of photon flux J_A incorporating only the effect of *side B*. Therefore,

$$J_A = J_{AA} + J_{AB} \tag{5}$$

Using (4) and (5), we have,

$$\begin{aligned} R_A &= \frac{J_{AA}}{I_0} + \frac{J_{AB}}{I_0} \\ &= R_{AA} + R_{AB} \end{aligned} \tag{6}$$

Here, R_{AA} is the *self-reflectance* of *side A*, defined as the component of R_A without taking into account the effect of *side B*. Again, R_{AB} is the *cross-reflectance* of *side A*, which is the component of R_A incorporating only the effect of *side B*.

The calculation and derivation for the model is organized as follows:

1. In the first step of our calculations, we determine the different photon flux components relating to the show-through phenomenon. As shown in Figs. 4, 5 and 6, this is done in three phases:
 - (a) In phase 1 of this step, we calculate J_{AA} and the photon flux J_B transmitted through to *side B*.
 - (b) In phase 2 of this step, we calculate the net photon flux I_B^c incident on *side B*.
 - (c) In phase 3 of this step, we calculate the photon flux J_{AB} that is reflected back to the scanner head from the backing through *side B*.
2. In the second step, we determine all the required reflectances using the derived expressions of the photon fluxes in the first step. Finally, we establish relationship among the reflectances and present the mathematical model for show-through.

Before, we move onto the calculations of the photon flux, let us define all the probability terms that we would be using in the expressions:

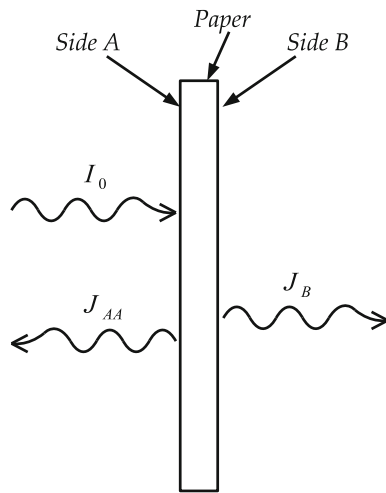


Fig. 4 The photon fluxes: J_{AA} and J_B

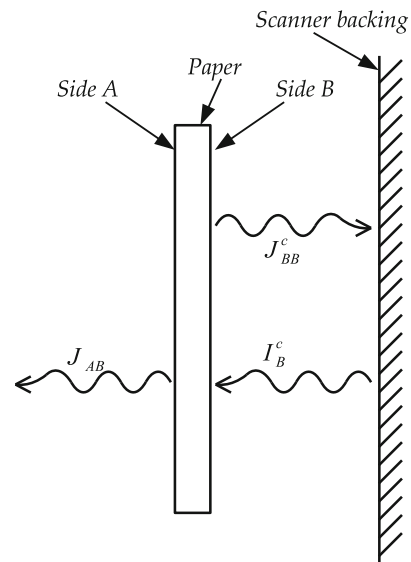


Fig. 6 Transmission of photon flux from side B

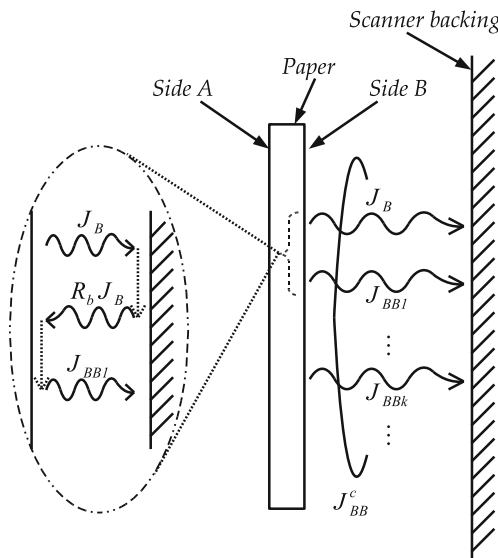


Fig. 5 The cumulated photon flux emitting from side B

1. The Point Spread Function (PSF), $p(\bar{r}_m, \bar{r}_n)$ is the joint probability of photons striking any arbitrary point at position \bar{r}_m of Σ_m and exiting from the paper substrate at position \bar{r}_n of Σ_n . Here, $m, n = 1A, 2A, 1B$ or $2B$.
2. $P(m)$ is the probability of photon striking the substrate in Σ_m .
3. $P(m, n)$ is the joint probability of photon striking the substrate in Σ_m and emerging from Σ_n .
4. ${}_m P_n$ is the conditional probability of photons emerging from Σ_n given that the photons struck the substrate in Σ_m .

3.1 Calculation of the photon fluxes J_{AA} and J_B

For the derivation of the show-through effect model, we follow a probabilistic approach in representing the scattering of

photons in the substrate paper. Therefore, we would obtain expressions of the different photon fluxes and their constituent components having corresponding probability terms. In this Section, we relate J_{AA} and J_B to the probability functions and other characteristic parameters of the printed document.

The photon flux incident on side A is I_0 . As shown in Fig. 4, a fraction of this flux is reflected as J_{AA} , another fraction is transmitted through to side B as J_B and the rest is absorbed in the paper and ink. The photon flux J_{AA} can be split up into two components as,

$$J_{AA} = J_{AA}^P + J_{AA}^i \tag{7}$$

Here, J_{AA}^P is the component of J_{AA} emerging from the paper substrate, i.e., from Σ_{2A} and J_{AA}^i is the component of photon flux emerging from the localities of the ink dots marked as Σ_{1A} . The photon flux components J_{AA}^P and J_{AA}^i are the fluxes observed from a microscopic point of view because, for these cases, we consider the microscopic localities Σ_{1A} and Σ_{2A} in our selected region of the document. The macroscopic photon flux J_{AA} is the photon flux yielding from the average of the microscopic flux components.

The microscopic photon fluxes can be further subdivided into elementary components. In the sequel, these elementary photon flux components will be denoted by ${}_m J_n$, where $m, n = 1A, 2A, 1B$ or $2B$. Here, ${}_m J_n$ is the flux component entering through Σ_m and exiting from Σ_n .

Now, J_{AA}^P consists of two photon fluxes: ${}_{1A} J_{2A}$ and ${}_{2A} J_{2A}$. We can express these elementary photon fluxes as,

$$\left. \begin{aligned} {}_{1A} J_{2A} &= I_0 f_A T_{iA} ({}_{1A} P_{2A}) \\ {}_{2A} J_{2A} &= I_0 (1 - f_A) ({}_{2A} P_{2A}) \end{aligned} \right\} \tag{8}$$

where, T_{iA} is the transmittance of ink on *side A*. Similarly, the elementary components ${}_{1A}J_{1A}$ and ${}_{2A}J_{1A}$ of J_{AA}^i are given as follows,

$$\left. \begin{aligned} {}_{1A}J_{1A} &= I_0 f_A T_{iA} ({}_{1A}P_{1A}) T_{iA} \\ {}_{2A}J_{1A} &= I_0 (1 - f_A) ({}_{2A}P_{1A}) T_{iA} \end{aligned} \right\} \quad (9)$$

Therefore, the *microscopic* components of J_{AA} in (7) can be obtained as,

$$\begin{aligned} J_{AA}^p &= {}_{1A}J_{2A} + {}_{2A}J_{2A} \\ &= I_0 [{}_{1A}P_{2A} f_A T_{iA} + {}_{2A}P_{2A} (1 - f_A)] \end{aligned} \quad (10)$$

and,

$$\begin{aligned} J_{AA}^i &= {}_{1A}J_{1A} + {}_{2A}J_{1A} \\ &= I_0 [{}_{1A}P_{1A} f_A T_{iA}^2 + {}_{2A}P_{1A} (1 - f_A) T_{iA}] \end{aligned} \quad (11)$$

In the rest of this section, the photon flux J_B transmitted from *side A* through to *side B* is calculated. In this case, the photon flux is affected by both *side A* and *side B*.

Now, the macroscopical value of the photon flux emerging from *side B* can be expressed as,

$$J_B = I_0 \bar{\chi}_{AB}(f_A, f_B) \quad (12)$$

where, $\bar{\chi}_{AB}(f_A, f_B)$ is the transmission ratio for the combined paper and ink system for our concerned region.

3.2 Calculation of the total photon flux incident on *side B* (I_B^c)

The photon flux J_B transmitted through to *side B* will undergo multiple reflections between the surfaces of *side B* and scanner backing as shown in Fig. 5. Therefore, there would be a *cumulated* photon flux incident on *side B*. In this section, we account for the multiple reflections and determine the net photon flux I_B^c incident on *side B*.

For a backing of reflectance R_b , the photon flux J_B will be reflected back to the paper substrate with the value of $R_b J_B$. This photon flux will be reflected from *side B* giving a photon flux of J_{BB1} . This photon flux will have two microscopic components: J_{BB1}^p and J_{BB1}^i . Now, with the light source of flux $R_b J_B$, using the similar concept used in the preceding section, we have,

$$J_{BB1}^p = R_b J_B \underbrace{[f_B T_{iB} ({}_{1B}P_{2B}) + (1 - f_B) ({}_{2B}P_{2B})]}_{\beta_p} \quad (13)$$

$$J_{BB1}^i = R_b J_B \underbrace{[f_B T_{iB}^2 ({}_{1B}P_{1B}) + (1 - f_B) T_{iB} ({}_{2B}P_{1B})]}_{\beta_i} \quad (14)$$

Therefore, the first reflected photon flux J_{BB1} of J_B from *side B* is,

$$\begin{aligned} J_{BB1} &= J_{BB1}^p + J_{BB1}^i \\ &= R_b J_B (\beta_p + \beta_i) \\ &= R_b J_B R_{BB} \end{aligned} \quad (15)$$

Here, R_{BB} is called the *self-reflectance* of *side B*. The *self-reflectance* of *side B*, defined as the component of the observed reflectance of *side B* (R_B) without taking into account the effect of *side A*, is given by,

$$R_{BB} = \frac{J_{BB1}}{R_b J_B} = \beta_p + \beta_i \quad (16)$$

Similarly, J_{BB1} will be reflected again from the backing as $R_b J_{BB1}$ and it will be reflected from *side B* as J_{BB2} . Then,

$$\begin{aligned} J_{BB2} &= (R_b J_{BB1}) R_{BB} \\ &= J_B (R_b R_{BB})^2 \end{aligned} \quad (17)$$

The k th reflected photon component toward the scanner backing is,

$$J_{BBk} = J_B (R_b R_{BB})^k \quad (18)$$

Now, taking into account all the photon flux emerging from the surface of *side B*, the *cumulated photon flux* emerging from *side B* is given by,

$$\begin{aligned} J_{BB}^c &= J_B + \sum_{k=1}^{\infty} J_{BBk} \\ &= J_B + J_B \sum_{k=1}^{\infty} (R_{BB} R_b)^k \\ &= J_B \left[1 + \sum_{k=1}^{\infty} (R_{BB} R_b)^k \right] \\ &= J_B R_{BB}^c \end{aligned} \quad (19)$$

where,

$$\begin{aligned} R_{BB}^c &= \left[1 + \sum_{k=1}^{\infty} (R_{BB} R_b)^k \right] \\ &= \frac{1}{1 - R_{BB} R_b} \end{aligned} \quad (20)$$

is the *cumulated reflectance component* of *side B*, taking into account only the effect of *side B*.

Incorporating all photon fluxes incident on *side B*, the *cumulated photon flux* incident on *side B* is given by,

$$I_B^c = R_b J_{BB}^c \quad (21)$$

Using (12), (19) and (21), the cumulated photon flux incident on side B (I_B^c) can be expressed as,

$$I_B^c = R_b J_{BB}^c = R_b J_B R_{BB}^c \tag{22}$$

$$= R_b R_{BB}^c I_0 \bar{\chi}_{AB}(f_A, f_B) \tag{23}$$

3.3 Calculating J_{AB}

J_{AB} is the photon flux emerging from side A due to only the photons reflected from the scanner backing through the paper. The net flux incident on side B, I_B^c , acts as a source for the transmitted photon flux, J_{AB} , from side B to side A. This is shown in Fig. 6. Using a similar concept of (12), J_{AB} can be expressed as,

$$J_{AB} = I_B^c \bar{\chi}_{AB}(f_A, f_B) \tag{24}$$

$$= R_b R_{BB}^c I_0 \bar{\chi}_{AB}^2(f_A, f_B) \quad [\text{using (23)}] \tag{25}$$

Now that we have found the expressions for the necessary photon fluxes, we can begin the calculations for determining the reflectances. Before moving onto the next step, we find the relationships between PSFs and the probability functions, which will assist in expressing the reflectances in compact forms. The following subsection is intended for this purpose.

3.4 Relationships between PSFs and the probability functions

In this Subsection, we relate the average PSFs to the conditional probability functions. Then, the relationships among the conditional probability functions are presented. These relationships would reduce the number of probability functions in our expressions and enable us to present the expressions for photon fluxes in a more compact form.

The area of both the sides corresponding to the specific region is equal, i.e.,

$$\text{area of } \Sigma_A = \text{area of } \Sigma_B \tag{26}$$

$$\therefore (\text{area of } \Sigma_{1A}) + (\text{area of } \Sigma_{2A}) = (\text{area of } \Sigma_{1B}) + (\text{area of } \Sigma_{2B}) \tag{27}$$

Therefore, we have, the average PSFs,

$$\bar{P}_A = \frac{1}{f_A(1-f_A)} \int_{\Sigma_{1A}} \int_{\Sigma_{2A}} p(\bar{r}_{1A}, \bar{r}_{2A}) d\sigma_1 d\sigma_2 \tag{28}$$

$$\bar{P}_B = \frac{1}{f_B(1-f_B)} \int_{\Sigma_{1B}} \int_{\Sigma_{2B}} p(\bar{r}_{1B}, \bar{r}_{2B}) d\sigma_1 d\sigma_2 \tag{29}$$

$$\bar{P}_{11} = \frac{1}{f_A f_B} \int_{\Sigma_{1A}} \int_{\Sigma_{1B}} p(\bar{r}_{1A}, \bar{r}_{1B}) d\sigma_1 d\sigma_2 \tag{30}$$

$$\bar{P}_{12} = \frac{1}{f_A(1-f_B)} \int_{\Sigma_{1A}} \int_{\Sigma_{2B}} p(\bar{r}_{1A}, \bar{r}_{2B}) d\sigma_1 d\sigma_2 \tag{31}$$

$$\bar{P}_{21} = \frac{1}{(1-f_A)f_B} \int_{\Sigma_{2A}} \int_{\Sigma_{1B}} p(\bar{r}_{2A}, \bar{r}_{1B}) d\sigma_1 d\sigma_2 \tag{32}$$

$$\bar{P}_{22} = \frac{1}{(1-f_A)(1-f_B)} \int_{\Sigma_{2A}} \int_{\Sigma_{2B}} p(\bar{r}_{2A}, \bar{r}_{2B}) d\sigma_1 d\sigma_2 \tag{33}$$

The average PSFs are assumed to be constants independent of the dot percentages of side A and side B (i.e., f_A and f_B). The elementary photon flux entering through Σ_{1A} and exiting from Σ_{2A} is,

$${}_{1A}J_{2A} = I_0 T_{iA} \int_{\Sigma_{1A}} \int_{\Sigma_{2A}} p(\bar{r}_{1A}, \bar{r}_{2A}) d\sigma_1 d\sigma_2 = I_0 T_{iA} f_A (1-f_A) \bar{P}_A \tag{34}$$

Similarly,

$${}_{2A}J_{1A} = I_0 T_{iA} \int_{\Sigma_{2A}} \int_{\Sigma_{1A}} p(\bar{r}_{2A}, \bar{r}_{1A}) d\sigma_1 d\sigma_2 = I_0 T_{iA} f_A (1-f_A) \bar{P}_A \tag{35}$$

Using (8) and (34), (9) and (35), we get,

$$\left. \begin{aligned} {}_{1A}P_{2A} &= \bar{P}_A(1-f_A) \\ {}_{2A}P_{1A} &= \bar{P}_A f_A \end{aligned} \right\} \tag{36}$$

and, similarly, for side B,

$$\left. \begin{aligned} {}_{1B}P_{2B} &= \bar{P}_B(1-f_B) \\ {}_{2B}P_{1B} &= \bar{P}_B f_B \end{aligned} \right\} \tag{37}$$

Thus far, we have established the relationships among the corresponding conditional probability functions through the dot percentages and the average PSFs.

3.5 Calculating the reflectances

From (6), (7), (10) and (11), we have,

$$R_{AA} = \frac{J_{AA}}{I_0} = ({}_{1A}P_{1A})T_{iA}^2 f_A + 2f_A(1-f_A)\bar{P}_A T_{iA} + ({}_{2A}P_{2A})(1-f_A) \tag{38}$$

Using (6) and (25), we obtain the compact form of R_{AB} as,

$$R_{AB} = \frac{J_{AB}}{I_0} = R_b R_{BB}^c \bar{\chi}_{AB}^2(f_A, f_B) \tag{39}$$

As shown in Appendix A, the transmission ratio of the combined paper and ink system of our concerned region can be

expressed as,

$$\bar{\chi}_{AB}(f_A, f_B) = T_w[1 - f_A(1 - T_{iA})][1 - f_B(1 - T_{iB})] \tag{40}$$

For unprinted (white) *side B*, $f_B = 0$. In this particular condition, there is no show-through data from *side B* to *side A*. Now, we define R_{AB}^0 as the value of the reflectance R_{AB} when there is no printing on *side B*, i.e., when $f_B = 0$. Therefore, using (39) and (40), we have,

$$\begin{aligned} R_{AB}^0 &= (R_{AB}|f_B = 0) \\ &= R_b R_{BB}^{c0} \bar{\chi}_{AB}(f_A, f_B = 0) \\ &= R_b R_{BB}^{c0} T_w^2 [1 - (1 - T_{iA})f_A]^2 \\ &= R_{AB}^w [1 - (1 - T_{iA})f_A]^2 \end{aligned} \tag{41}$$

Here, the term $R_{AB}^w = R_b R_{BB}^{c0} T_w^2$ is a constant and depends on material properties of the paper and scanner backing. T_w is the transmittance of white paper and $R_{BB}^{c0} = (R_{BB}^c|f_B = 0)$.

Now, from the calculations done in Appendix B, the conditional probabilities relating to the photon scattering on *side A* can be expressed as functions of f_A as given below,

$$\left. \begin{aligned} {}_{1A}P_{1A} &= \frac{P(1A,1A)}{f_A} = \bar{p}_A f_A + (R_{AA}^w - \bar{p}_A) \\ {}_{2A}P_{2A} &= \frac{P(2A,2A)}{1-f_A} = -\bar{p}_A f_A + R_{AA}^w \\ {}_{1A}P_{2A} &= \frac{P(1A,2A)}{f_A} = -\bar{p}_A f_A + \bar{p}_A \end{aligned} \right\} \tag{42}$$

We can now replace the conditional probabilities ${}_{1A}P_{1A}$, ${}_{2A}P_{2A}$ and ${}_{1A}P_{2A}$ with expressions containing f_A and the constants \bar{p}_A and R_{AA}^w . Here R_{AA}^w is the value of the reflectance R_{AA} when there is no printing on *side A*, i.e., $R_{AA}^w = (R_{AA}|f_A = 0)$. From (38) and (42), we get,

$$\begin{aligned} R_{AA} &= ({}_{1A}P_{1A})T_{iA}^2 f_A + 2f_A(1 - f_A)\bar{p}_A T_{iA} \\ &\quad + ({}_{2A}P_{2A})(1 - f_A) \\ &= [\bar{p}_A f_A + (R_{AA}^w - \bar{p}_A)]T_{iA}^2 f_A \\ &\quad + 2f_A(1 - f_A)\bar{p}_A T_{iA} \\ &\quad + (-\bar{p}_A f_A + R_{AA}^w)(1 - f_A) \\ &= -(1 - f_A)f_A \bar{p}_A T_{iA}^2 + f_A R_{AA}^w T_{iA}^2 \\ &\quad + 2f_A(1 - f_A)\bar{p}_A T_{iA} - (1 - f_A)f_A \bar{p}_A \\ &\quad + (1 - f_A)R_{AA}^w \\ &= f_A R_{AA}^w T_{iA}^2 + (1 - f_A)R_{AA}^w \\ &\quad - f_A(1 - f_A)\bar{p}_A(1 - T_{iA})^2 \end{aligned} \tag{43}$$

We define R_A^0 as the value of the reflectance R_A when there is no printing on *side B*, i.e., when $f_B = 0$. Therefore, using (6), (41) and (43),

$$\begin{aligned} R_A^0 &= (R_A|f_B = 0) \\ &= ((R_{AA} + R_{AB})|f_B = 0) \\ &= R_{AA} + R_{AB}^0 \\ &= R_A^w [f_A T_{iA}^2 + (1 - f_A)] \\ &\quad - (1 - f_A)f_A(1 - T_{iA})^2(\bar{p}_A + R_{AB}^w) \\ &= f_A R_A^w T_{iA}^2 + (1 - f_A)R_A^w \\ &\quad - (1 - f_A)f_A(1 - T_{iA})^2 k_{pA} \\ &= R_A^w - R_A^w(1 - T_{iA}^2)f_A - f_A(1 - T_{iA})^2 k_{pA} \\ &\quad + f_A^2 k_{pA}(1 - T_{iA})^2 \end{aligned} \tag{44}$$

where, $k_{pA} = (\bar{p}_A + R_{AB}^w)$ is a constant that depends on the properties of the page. Here, R_A^w is the *white paper reflectance* of *side A*. The reflectance observed at the scanner head at *side A* would be R_A^w if there were no printing on either sides of the document, i.e.,

$$\begin{aligned} R_A^w &= (R_A|f_A = 0, f_B = 0) \\ &= ((R_{AA} + R_{AB})|f_A = 0, f_B = 0) \\ &= R_{AA}^w + R_{AB}^w \end{aligned} \tag{46}$$

Similarly, we can define *white paper reflectance* of *side B*, R_B^w .

As shown in Appendix C, the solution for (45) with respect to f_A is given by,

$$f_A = \frac{b_1 - \sqrt{b_1^2 - 4a_1c_1}}{2a_1} \tag{47}$$

where,

$$\left. \begin{aligned} a_1 &= k_{pA}(1 - T_{iA})^2 \\ b_1 &= R_A^w(1 - T_{iA}^2) + (1 - T_{iA})^2 k_{pA} \\ c_1 &= R_A^w - R_A^0 \end{aligned} \right\}$$

Similarly, for *side B*,

$$f_B = \frac{b_2 - \sqrt{b_2^2 - 4a_2c_2}}{2a_2} \tag{48}$$

where,

$$\left. \begin{aligned} a_2 &= k_{pB}(1 - T_{iB})^2 \\ b_2 &= R_B^w(1 - T_{iB}^2) + (1 - T_{iB})^2 k_{pB} \\ c_2 &= R_B^w - R_B^0 \end{aligned} \right\}$$

where, $k_{pB} = (\bar{p}_B + R_{BA}^w)$ is a constant similar to k_{pA} .

Substituting the value of $\bar{\chi}_{AB}(f_A, f_B)$ from (40) into (39), R_{AB} can be expressed as,

$$\begin{aligned}
 R_{AB} &= R_b R_{BB}^c \bar{\chi}_{AB}^2(f_A, f_B) \\
 &= R_b R_{BB}^{c0} T_w^2 \frac{R_{BB}^c}{R_{BB}^{c0}} \\
 &\quad \times [1 - f_A(1 - T_{iA})]^2 [1 - f_B(1 - T_{iB})]^2 \\
 &= R_{AB}^w \frac{R_{BB}^c}{R_{BB}^{c0}} [1 - f_A(1 - T_{iA})]^2 [1 - f_B(1 - T_{iB})]^2
 \end{aligned} \tag{49}$$

where, $R_{AB}^w = R_b R_{BB}^{c0} T_w^2$. The term $[1 - f_A(1 - T_{iA})]^2$ can be expressed as a linear function of R_A^0 and f_A using (45). Rearranging (45), we have,

$$\begin{aligned}
 [1 - f_A(1 - T_{iA})]^2 &= \frac{1}{k_{pA}} [-(R_A^w - R_A^0) \\
 &\quad + f_A(1 - T_{iA}^2)(-k_{pA} + R_A^w) + k_{pA}] \\
 &= \frac{1}{k_{pA}} R_A^0 - (1 - T_{iA}^2) \left(1 - \frac{R_A^w}{k_{pA}}\right) f_A \\
 &\quad + \left(1 - \frac{R_A^w}{k_{pA}}\right) \\
 &= \mu_{1A} R_A^0 + \mu_{2A} f_A + \mu_{3A}
 \end{aligned} \tag{50}$$

Here, μ_{1A}, μ_{2A} and μ_{3A} are constants that depend on the page and ink properties of *side A*. They are given by,

$$\left. \begin{aligned}
 \mu_{1A} &= \frac{1}{k_{pA}} \\
 \mu_{2A} &= -(1 - T_{iA}^2) \left(1 - \frac{R_A^w}{k_{pA}}\right) \\
 \mu_{3A} &= \left(1 - \frac{R_A^w}{k_{pA}}\right)
 \end{aligned} \right\} \tag{51}$$

Using (44) in (6), we have,

$$\begin{aligned}
 R_A &= R_{AA} + R_{AB} \\
 &= R_A^0 - R_{AB}^0 + R_{AB}
 \end{aligned} \tag{52}$$

In (52), we intend to express R_{AB}^0 and R_{AB} in terms of R_A^0, R_B^0, f_A and f_B .

Therefore, using (41) and (49), we obtain the expression of R_{AB} given by,

$$R_{AB} = \frac{R_{AB}^0}{R_{BB}^{c0}} R_{BB}^c [1 - f_B(1 - T_{iB})]^2 \tag{53}$$

Using (41) and (50), R_{AB}^0 can be expressed as,

$$\begin{aligned}
 R_{AB}^0 &= R_{AB}^w [1 - f_A(1 - T_{iA})]^2 \\
 &= R_{AB}^w (\mu_{1A} R_A^0 + \mu_{2A} f_A + \mu_{3A})
 \end{aligned} \tag{54}$$

Before substituting R_{AB} and R_{AB}^0 into (52), we present a relation similar to (50) for *side B* as,

$$[1 - f_B(1 - T_{iB})]^2 = \mu_{1B} R_B^0 + \mu_{2B} f_B + \mu_{3B} \tag{55}$$

where,

$$\left. \begin{aligned}
 \mu_{1B} &= \frac{1}{k_{pB}} \\
 \mu_{2B} &= -(1 - T_{iB}^2) \left(1 - \frac{R_B^w}{k_{pB}}\right) \\
 \mu_{3B} &= \left(1 - \frac{R_B^w}{k_{pB}}\right)
 \end{aligned} \right\} \tag{56}$$

Using (52) and (53), the reflectance observed at the scanner head at *side A* can be expressed as,

$$\begin{aligned}
 R_A &= R_{AA} + R_{AB} \\
 &= R_A^0 - R_{AB}^0 + \frac{R_{AB}^0}{R_{BB}^{c0}} R_{BB}^c [1 - f_B(1 - T_{iB})]^2 \\
 &= R_A^0 - R_{AB}^0 \left[1 - \frac{R_{BB}^c}{R_{BB}^{c0}} [1 - f_B(1 - T_{iB})]^2\right]
 \end{aligned} \tag{57}$$

This equation can be further processed by replacing the terms R_{AB}^0 and $[1 - f_B(1 - T_{iB})]^2$ by the corresponding expressions containing the terms R_A^0, f_A and R_B^0, f_B . Finally, using (54) and (55) in (57), we get,

$$\begin{aligned}
 R_A &= R_A^0 - R_{AB}^w (\mu_{1A} R_A^0 + \mu_{2A} f_A + \mu_{3A}) \\
 &\quad \times \left[1 - \frac{R_{BB}^c}{R_{BB}^{c0}} (\mu_{1B} R_B^0 + \mu_{2B} f_B + \mu_{3B})\right]
 \end{aligned} \tag{58}$$

Here, R_A can be expressed in terms of R_A^0 and R_B^0 , by substituting the expressions of f_A and f_B given in (47) and (48).

It is possible to simplify (58) if we assume that $\bar{p}_A \approx R_{AA}^w$ and $\bar{p}_B \approx R_{BB}^w$ which results, $\mu_{2A} = \mu_{3A} = \mu_{2B} = \mu_{3B} = 0$. Now, \bar{p}_A is the fraction of photon flux that would be emerging from Σ_{1A} and Σ_{2A} taking into account only the effect of *side A* and with $T_{iA} = 1$ (i.e., no ink dot on Σ_{1A}). This should be a measure of the *self-reflectance* of unprinted *side A* R_{AA}^w , providing the logic behind our assumption. The same logic can be stated for *side B*.

Therefore, applying this assumption, we can simplify (58) to the following form,

$$R_A = R_A^0 - R_{AB}^w \left(\frac{R_A^0}{R_A^w}\right) \left[1 - \frac{R_{BB}^c}{R_{BB}^{c0}} \left(\frac{R_B^0}{R_B^w}\right)\right] \tag{59}$$

where, R_{BB}^c is dependent on *side B* reflectance R_B^0 as,

$$R_{BB}^c = \frac{1}{1 - R_B^0 \left(1 - \frac{R_{AB}^w}{R_B^w}\right) R_b} \tag{60}$$

From (59) and (60) we can see that, as the *side B* reflectance R_B^0 increases, R_{BB}^c increases and the *side A* reflectance R_A also increases. Similarly, R_A decreases with the decrease in R_B^0 .

The reflectance observed at the scanner head at *side B* can similarly be expressed as,

$$R_B = R_B^0 - R_{BA}^w \left(\frac{R_B^0}{R_B^w} \right) \left[1 - \frac{R_{AA}^c}{R_{AA}^0} \left(\frac{R_A^0}{R_A^w} \right) \right] \tag{61}$$

where, R_{AA}^c is dependent on *side A* reflectance R_A^0 as,

$$R_{AA}^c = \frac{1}{1 - R_A^0 \left(1 - \frac{R_{BA}^w}{R_A^w} \right) R_b} \tag{62}$$

It is to be noted that R_{AB}^w , R_{BB}^c and R_{BB}^{c0} in (59) depend on the reflectance of the scanner backing (R_b). In fact, R_{AB}^w is proportional to R_b and thus the presented formulation ideally shall not predict any show-through in the case when the scanner backing is black ($R_b = 0$). This is due to the fact that the scattering of photons from the inner surface of *side B* has been neglected for the derivation of the expression for R_A . But, as for most practical cases, when R_b is high, the reflectance component due to the scattering of photons from the inner surface of *side B* will be negligible compared to other components. Thus, the expressions presented here are valid for most of the practical cases with bright scanner backing ($R_b \gtrsim 0.8$).

4 Model for space varying data

In Sect. 3, we have derived a model for mathematical representation of the show-through effect. The expressions hold for the *specific region* where the reflectances of both the sides corresponding to that region are space invariant. Therefore, by segmenting a whole document into such *regions*, we can establish the show-through effect in mathematical form for the complete document.

Now, let us consider the case when we scan any side of an arbitrary duplex printed document to obtain a digital form. The acquired digital form of the page will be divided into pixels. In most of the practical applications, the scanning resolution is lower than the printing resolution. Therefore, the *microscopic* geometric structure will not be visible in the digital version of this page. The pixels in this page give a *macroscopic* value of the reflectance and, as these pixels are small in size, we can assume that the pixel region conforms to the conditions applied to the *specific region* selected for our model. Thus, for the (m, n) th pixel of *side A*, we can write (59) as,

$$R_A(m, n) = R_A^0(m, n) - R_{AB}^w \left(\frac{R_A^0(m, n)}{R_A^w} \right) \times \left[1 - \frac{R_{BB}^c(m, n)}{R_{BB}^{c0}} \left(\frac{R_B^0(m, n)}{R_B^w} \right) \right] \tag{63}$$

where, $R_{BB}^c(m, n)$ varies with *side B* reflectance $R_B^0(m, n)$ as,

$$R_{BB}^c(m, n) = \frac{1}{1 - R_B^0(m, n) \left(1 - \frac{R_{AB}^w}{R_B^w} \right) R_b} \tag{64}$$

Here, all the reflectance values are in the range 0 to 1. The reflectances R_A^0 and R_B^0 are the show-through unaffected or show-through corrected reflectances. We would get a similar set of expressions for *side B*.

Another phenomenon associated with the show-through effect is *blurring*. In some cases, image or text edges are blurred when they show-through from the opposite side. To incorporate this effect, we need to consider the scattering of photons from one *specific region* (that conforms to our conditions) to another. For simplicity, we ignore blurring in this paper.

5 Verification of the derived model

In this section, we validate the derived model by fitting it to the data of a practical show-through affected document. Consider an A4 sized page with printing on only *side A*. The printed side of the page consists of blocks with shades of gray of different intensities. Figure 7 shows the scanned images of the two sides of the test document. The *side B* has been vertically reflected. The blocks printed on *side A* shows through on *side B* as seen in Fig. 7b.

Now, each of the blocked regions of the two sides has been marked as r_k , where $k = 1, 2, \dots, 10$. Here, the region r_k is darker compared to r_{k+1} . Due to nonuniformity in printing, the reflectance of a certain region, r_k , on *side A* will not be constant. Similarly, the blocks will not show through uniformly on *side B*. Therefore, we take the average reflectances for each region:

$$R_A^{av}(k) = \text{average}\{R_A(m, n)\} \text{ for region } r_k \tag{65}$$

$$R_B^{av}(k) = \text{average}\{R_B(m, n)\} \text{ for region } r_k \tag{66}$$

The white paper reflectance of both the sides (e.g., R_A^w and R_B^w) are calculated by averaging the reflectances of a region with no printing on either side. Now, as there is no printing on *side B*, the actual or corrected reflectance (R_B^0) would be equal to the white paper reflectance. Therefore, we set $R_B^{0av}(k) = R_B^w$, for all k .

Now, as there is no printing on *side B*, the reflectance scanned at *side A* (R_A) would be the show-through free reflectance (R_A^0). Therefore, we can write, $R_A^{0av}(k) = R_A(k)$.

For this test document, we have thus far determined the following values:

1. The reflectances for *side A*: $R_A^{0av}(k)$.

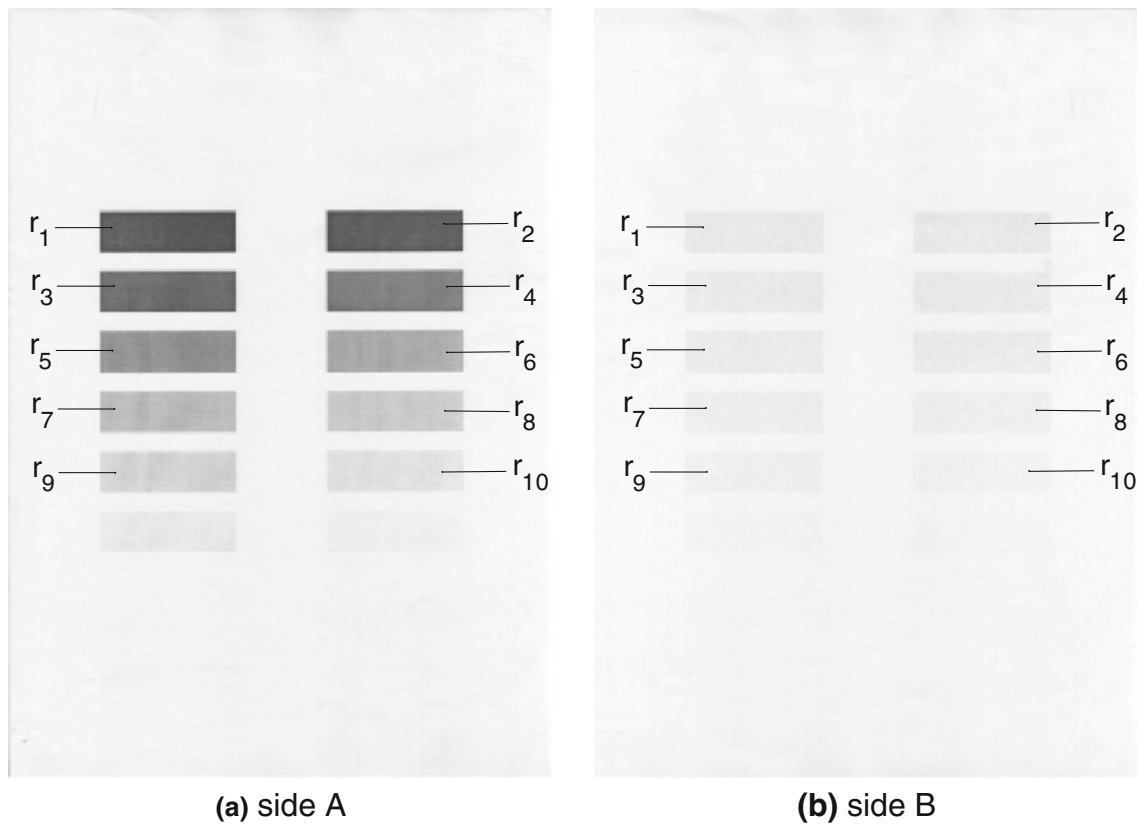


Fig. 7 Test Document: Blocks of varying intensity shades on one side of the page

2. The reflectances for *side B*: $R_B^{av}(k)$ and $R_B^{0av}(k)$.
3. The white paper reflectances of the two sides: R_A^w and R_B^w .

Now, in the model expressions for *side B*, (61) and (62), we observe that we have two unknown parameters: R_b and R_{BA}^w . These two parameters are assumed to be constants for the whole *side B*. We, therefore, fit our model, (61) and (62), in the calculated data ($R_A^{av}(k)$, $R_B^{av}(k)$, $R_B^{0av}(k)$, R_A^w and R_B^w) of the document and determine the unknowns (R_b and R_{BA}^w).

Figure 8 presents the plot of $R_B^{av}(k)$ versus $R_A^{0av}(k)$ for the current data set. As we can see, the model generated curve using the estimated parameters follows the scanned data. The curve fitting process yielded the following values of the two unknown parameters:

1. $R_b = 0.9638$
2. $R_{BA}^w = 0.0724$

which are practically acceptable. Here, the reflectance of the scanner backing R_b is very high (close to unity), which is obvious. And, R_{BA}^w has a small value, contributing very little to the white paper reflectance of *side B* (R_B^w). The major contribution to R_B^w is from $R_{BB}^w (= R_B^w - R_{BA}^w)$.

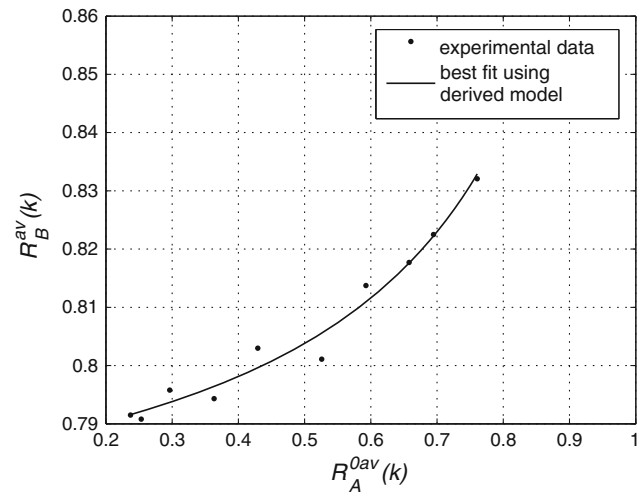


Fig. 8 Curve fitting using the derived model in practical data

6 Show-through correction algorithm

In this section, we present a show-through correction algorithm based on the derived model and by using the information extracted from the joint histogram.

For show-through correction using the proposed method, the parameters required to be calculated are the white paper

reflectances R_A^w and R_B^w of *side A* and *side B*, respectively, the reflectance of the scanner backing (R_b) and, the *cross-reflectance* of white paper for *side A* and *side B* (i.e., R_{AB}^w and R_{BA}^w , respectively). These are calculated by using the data extracted from logically defined points (referred to as *tracks*) on the joint histogram as described below.

6.1 The joint histogram

The joint histogram formed from the scans of the two sides of an arbitrary image is shown in Fig. 9. Here, the amplitude, z , of the point (x, y) indicates that there are z number of pixels of reflectance y (of *side B*) exactly opposite to the pixels of reflectance x (of *side A*). The points on the joint histogram with $z = 0$ indicate that the images do not jointly contain the corresponding reflectances (x, y) . Then, the joint histogram is smoothed as shown in Fig. 10 using the wavelet coefficient thresholding technique for further processing.

6.2 Smoothing the joint histogram

The joint histogram needs smoothing for further processing. Only the relative amplitudes of the peaks are required to be the same. But the major peaks on the joint histogram should retain their original positions after smoothing. Any appropriate smoothing technique following the above-mentioned criterion may be used. We choose wavelet coefficient thresholding technique for this purpose.

If the thresholding process is done on the surface shown in Fig. 9, there will be some unwanted edge effects, which will damage some of the information of the joint histogram. Therefore, to avoid the contamination of the edge artifacts, zero is padded around the joint histogram. We perform 3-level 2-D discrete stationary wavelet transform. All the

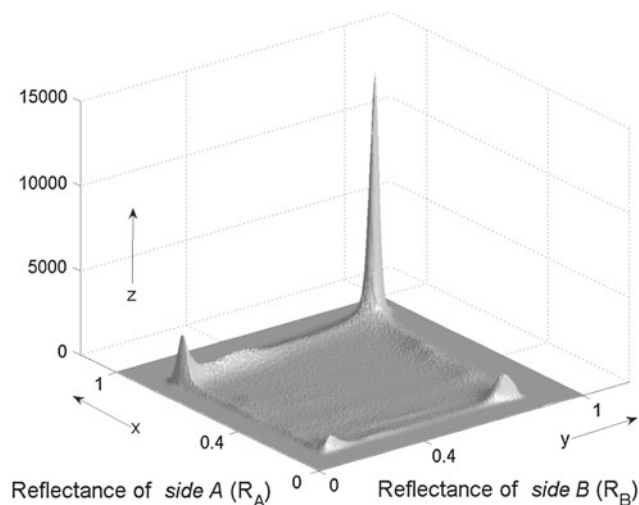


Fig. 9 The joint histogram of an arbitrary image

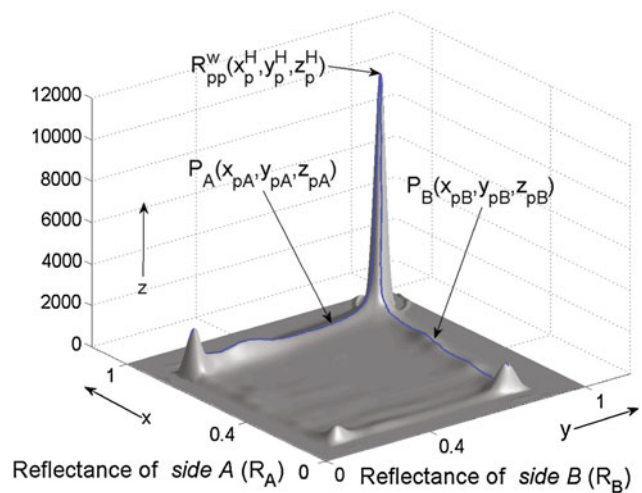


Fig. 10 Tracks for calculating transfer functions

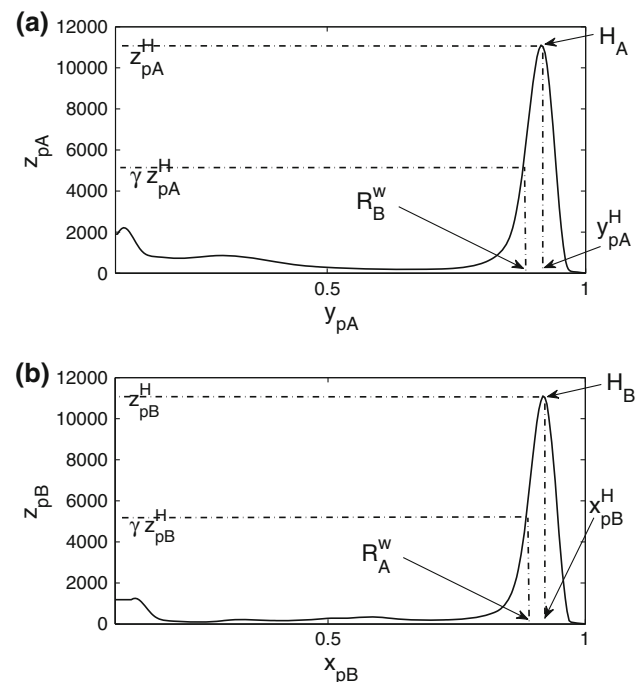


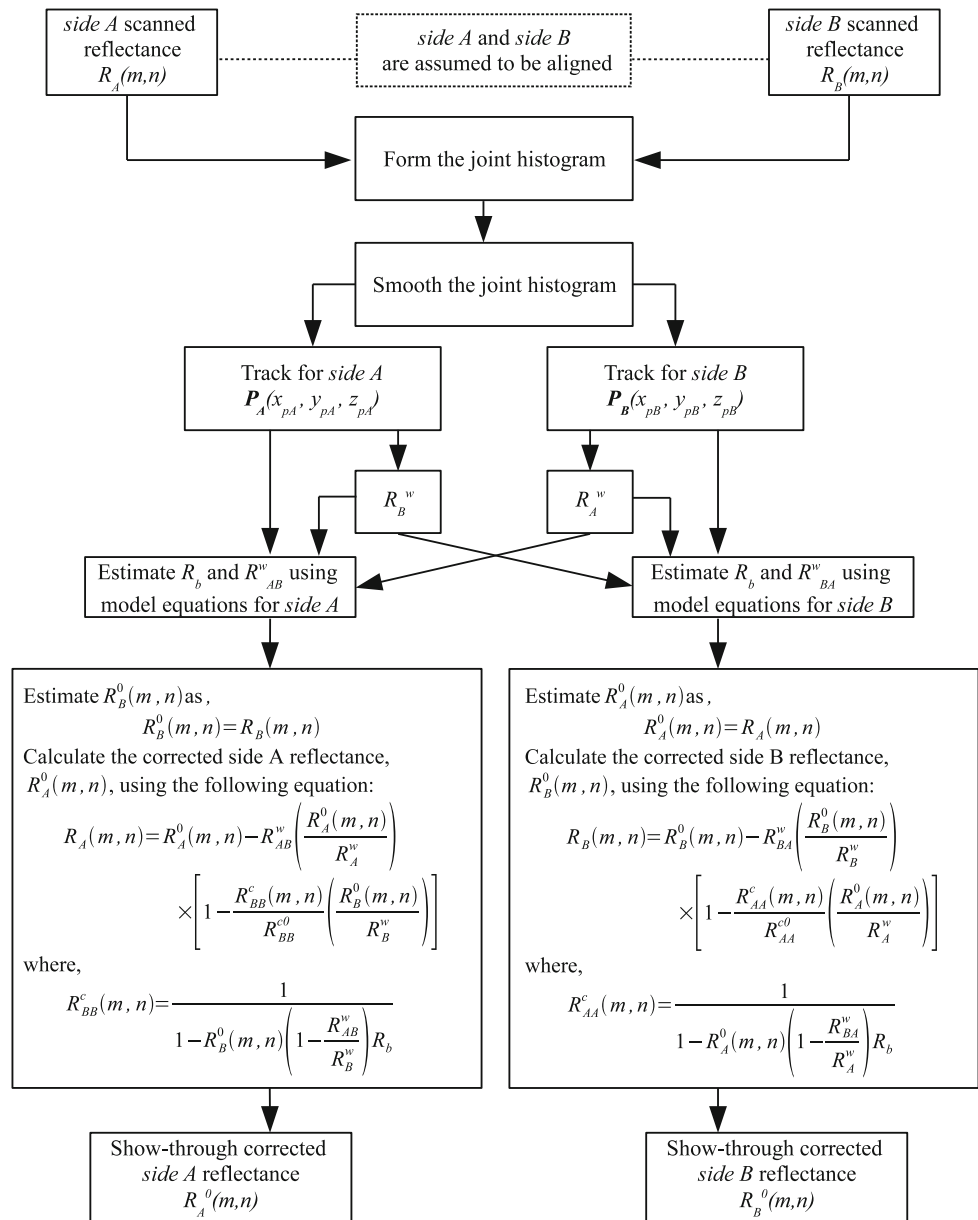
Fig. 11 Simplified diagrams of the tracks: a Simplified form of P_A , b simplified form of P_B

details of the 3 levels are set to zero and then the surface is reconstructed. The padded zeros are removed. The resulting joint histogram surface is shown in Fig. 10.

6.3 Determining the tracks

In the joint histogram of Fig. 10, the *nonzero amplitudes* (i.e., $z > 0$) of the brightest value (high reflectance) on *side A*, which are associated with the dark (low reflectance) pixels of *side B* (say, $x > x_H$ and $y < y_L$), will correspond only

Fig. 12 Proposed show-through correction algorithm



to the show-through reflectance from *side B* to *side A*. That is, there is actually printing (text or image) on *side B* but no printing on *side A*, at the pixels fulfilling the above-mentioned criterion. It means, these pixels are supposed to have reflectance of R_A^w (of *side A*), as if they had no show-through effect. As there is no printing on *side A* for these pixel, the reflectance of *side B* would be free of show-through. Therefore, for these pixels, we can set,

$$\left. \begin{aligned} R_A^0(m_p, n_p) &= R_A^w \\ R_B^0(m_p, n_p) &= R_B(m_p, n_p) \end{aligned} \right\} \quad (67)$$

where, (m_p, n_p) are the pixels having the above specified reflectance.

Due to the nonuniformity of the paper, all the dark pixels of *side B* will not show through to give the same brightness on the unprinted region of *side A*. Statistically speaking, instead of taking the highest reflectance value from *side A*, we should take the values corresponding to the peaks in the joint histogram from the brightest region (high reflectances) of *side A* associated with a dark (low reflectance) value of *side B* for the above calculation. The value chosen from the peaks (high frequency of pixels) will correspond to the expected values of our desired reflectances.

Now, we can set similar conditions, as shown in (67), for all *side B* reflectances. Using these conditions in (59) and (60), we obtain the following set of expressions $\forall(R_B^0 = y_{pA})$, given by,

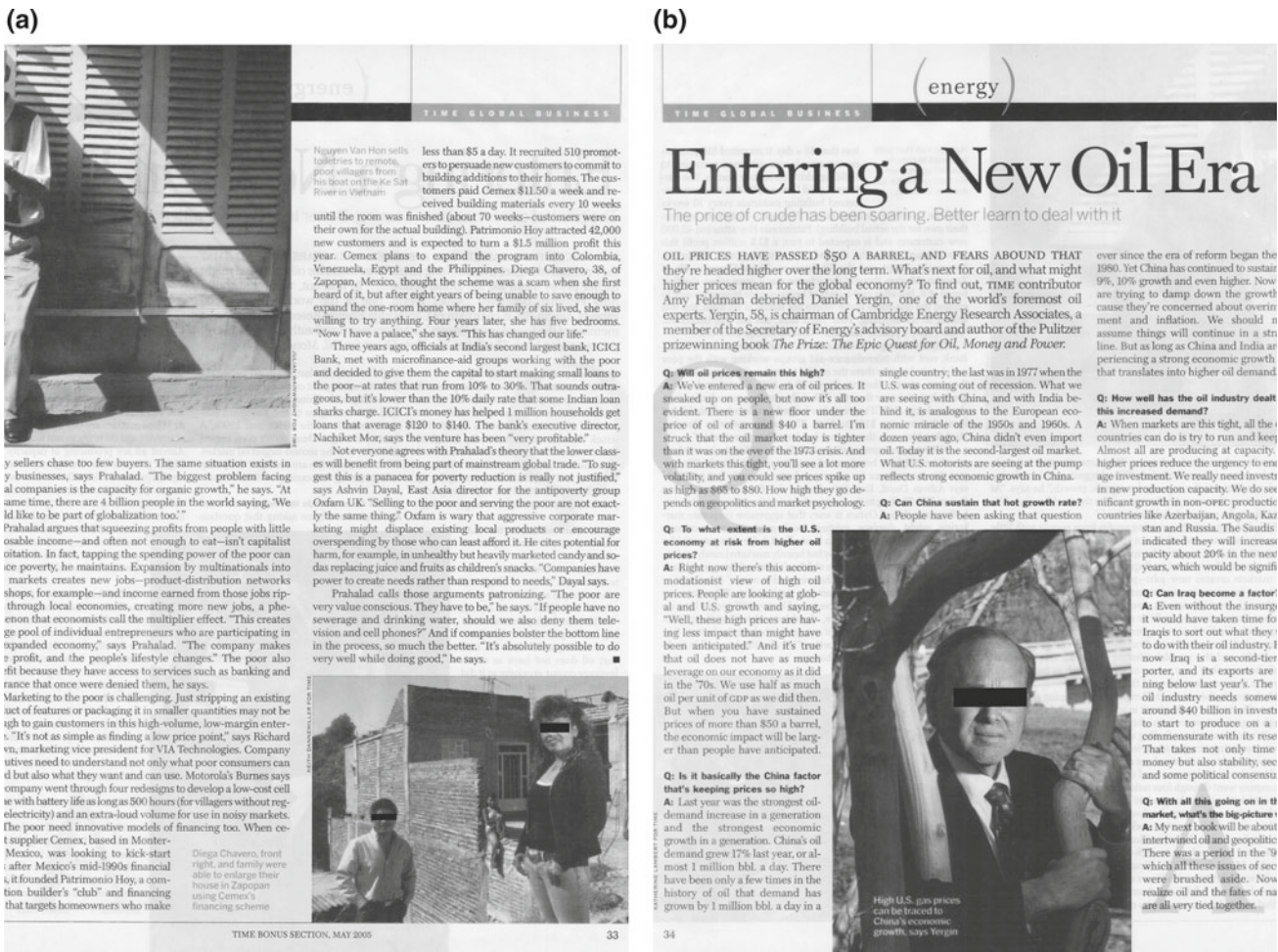


Fig. 13 Show-through corrupted document (1st example): a side A, b side B

$$R_A = R_A^w - R_{AB}^w \left[1 - \frac{R_{BB}^c}{R_{BB}^0} \left(\frac{R_B^0}{R_B^w} \right) \right] \quad (68)$$

and,

$$R_{BB}^c = \frac{1}{1 - R_B^0 \left(1 - \frac{R_{AB}^w}{R_B^w} \right) R_b} \quad (69)$$

For the above calculations, the points chosen are given by the locus of $\mathbf{P}_A(x_{pA}, y_{pA}, z_{pA})$ in the smoothed joint histogram as shown in Fig. 10. In (68) and (69), the unknown parameters for side A are:

1. The white paper reflectances of the two sides: R_A^w and R_B^w .
2. The reflectance of the scanner backing (R_b).
3. The cross-reflectance of white paper for side A (R_{AB}^w).

R_A^w and R_B^w are determined from the joint histogram and then R_b and R_{AB}^w are calculated using the set of expressions $\forall (R_B^0 = y_{pA})$, given in (68) and (69).

Similarly, the locus of $\mathbf{P}_B(x_{pB}, y_{pB}, z_{pB})$ is determined for estimating parameters for side B. We refer the loci, \mathbf{P}_A and \mathbf{P}_B as the tracks on the joint histogram. These tracks will be used to estimate the white paper reflectance (to be explained in Sect. 6.4) and the reflectances R_b , R_{AB}^w and R_{BA}^w (to be explained in Sect. 6.5).

6.4 Estimating white paper reflectances: R_A^w and R_B^w

The peak $R_{pp}^w(x_p^H, y_p^H, z_p^H)$, as shown in the joint histogram of Fig. 10, corresponding to the brightest regions of both the sides, can be used to calculate the white paper reflectance. But practically, the white paper is not uniform. Therefore, it would have a lot of pixels in its scanned image having reflectances below and over x_p^H (or y_p^H).

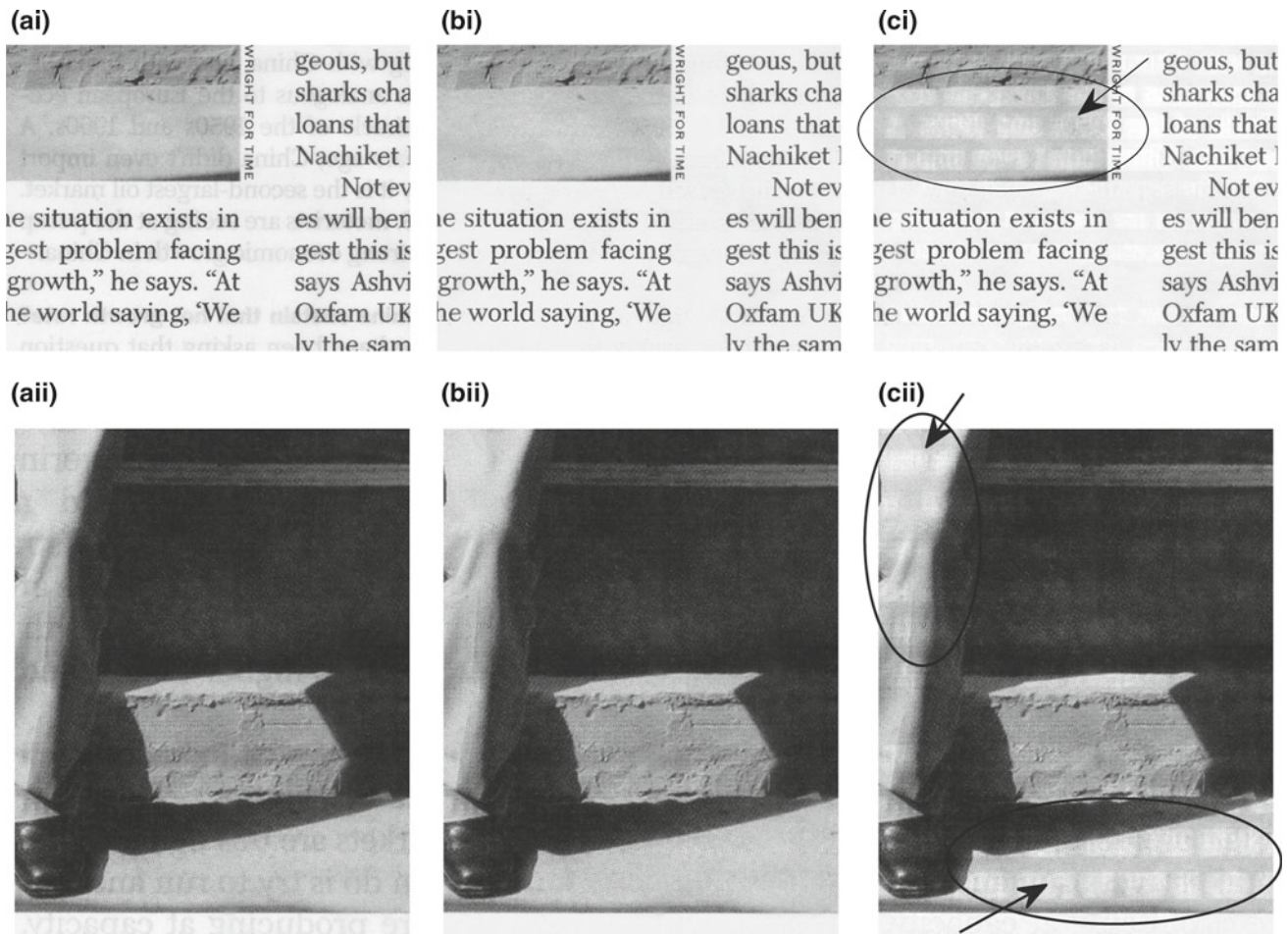
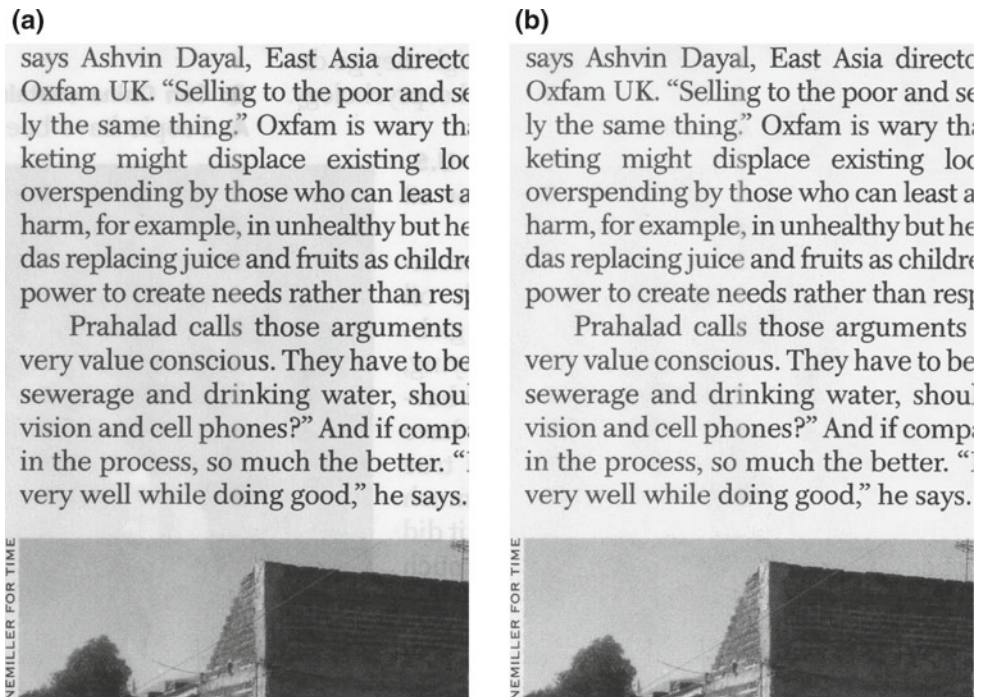


Fig. 14 Comparative results on show-through correction (from 1st example)

Fig. 15 a Image distorted by show-through, b results obtained using the proposed method



Pixels having higher reflectances than the estimated white paper reflectance are assumed to be white (unprinted). Hence, to consider larger number of pixels with reflectance values near x_p^H (or y_p^H), as white, we should take the white paper reflectance to be less than x_p^H (or y_p^H). Besides, human eye is more sensitive to contrast of light with brighter intensities [14]. An overestimated white paper reflectance would result in a visibly poor show-through corrected image with bright patches in the corrected regions.

For the calculation of the *frontside* white paper reflectance (R_A^w), we follow the track for *side B*, i.e., $\mathbf{P}_B(x_{pB}, y_{pB}, z_{pB})$ mentioned in Sect. 6.3. Let the highest peak in the brightest region of the track be given by the point $H_B(x_{pB}^H, y_{pB}^H, z_{pB}^H)$.

Simplified plots of these tracks, \mathbf{P}_A and \mathbf{P}_B , are shown in Fig. 11. Figure 11b shows the frequency of pixels (z_{pB}) versus the *side A* reflectances (x_{pB}) on the locus \mathbf{P}_B of the joint histogram. Now, in Fig. 11b we locate the frequency of the pixels given by $z_{pB}^H \times \gamma$ going from H_B toward the darker region, i.e., $x_{pB} < x_{pB}^H$. At this point, we get R_A^w . Here, $\gamma < 1$. R_A^w is chosen in this manner so that the brightest pixels which occur in high frequencies are incorporated in the calculation of R_A^w .

In a similar manner, white paper reflectance for *side B* (R_B^w) is calculated using the plot in Fig. 11a.

6.5 Calculating the reflectance of scanner backing and the cross-reflectances

The white paper reflectances R_A^w and R_B^w has been determined as explained in the previous subsection. Therefore, in (68) and (69), the reflectance of the scanner backing (R_b) and the *cross-reflectance* of white paper of *side A* (R_{AB}^w), are to be obtained. We find the values of R_b and R_{AB}^w such that the error between (x_{pA}, y_{pA}) and $(R_A, R_B^0) \forall (R_B^0 = y_{pA})$ is minimized. The set of points $(R_A, R_B^0) \forall (R_B^0 = y_{pA})$ is found from (68) and (69).

Similarly, the *cross-reflectance* of white paper of *side B* (R_{BA}^w) is calculated using the *track* of *side B*, i.e., $\mathbf{P}_B(x_{pB}, y_{pB}, z_{pB})$, and the expressions for *side B* similar to (68) and (69). For better results, R_b is calculated again for the correction process of *side B*.

6.6 Cleaning the image

The necessary parameters R_A^w , R_B^w , R_{AB}^w and R_b in the model expressions of (63) and (64) have been determined in the previous sections. We use $R_B(m, n)$ as an estimate of $R_B^0(m, n)$ and find the show-through corrected reflectance of *side A* ($R_A^0(m, n)$) using (63) and (64). The calculated reflectances having values higher than the white paper reflectance (R_A^w) are then thresholded down to R_A^w . The show-through correction from *side B* is done in a similar manner.

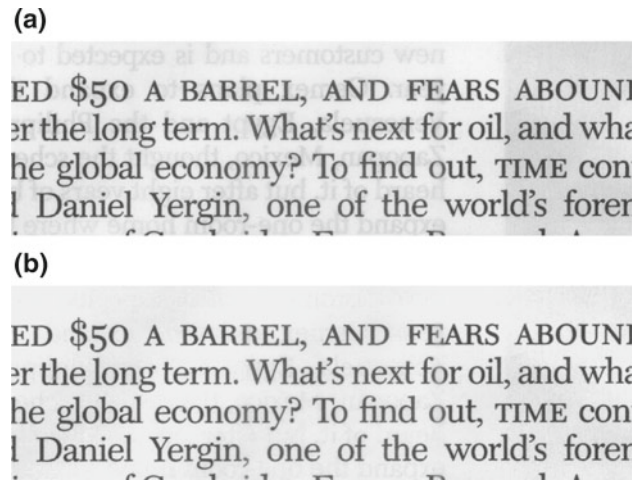


Fig. 16 **a** Image distorted by show-through, **b** results obtained using the proposed method

The rationale behind our assumption of $R_B(m, n)$ being an estimate of $R_B^0(m, n)$ is explained through the following cases:

1. **Both $\mathbf{R}_A(m, n)$ and $\mathbf{R}_B(m, n)$ have low values, i.e., the (m, n) th pixel of both the sides are dark:** For this condition, as $R_B^0(m, n)$ is already low (dark pixel), it is obvious that the reflectance of this pixel will not be significantly lowered due to show-through. Therefore, we can use $R_B(m, n)$ as an estimate of $R_B^0(m, n)$ with a very little error. Besides, the weighting of the error is further reduced due to the fact that $R_A^0(m, n)$ is also low. As seen in (63) and (64), the terms containing $R_B^0(m, n)$ is multiplied by $R_A^0(m, n)$ which is also low in this case.
2. **$\mathbf{R}_A(m, n) > \mathbf{R}_B(m, n)$, i.e., the *side B* reflectance of the (m, n) th pixel is darker compared to the corresponding *side A* reflectance:** In this case, the bright pixel of *side A* is expected to have negligible contribution in lowering the already-low $R_B^0(m, n)$, i.e., darkening the dark *side B* pixel. Therefore, we can assume $R_B(m, n)$ to be a good estimate of $R_B^0(m, n)$ in this case.
3. **$\mathbf{R}_A(m, n) < \mathbf{R}_B(m, n)$, i.e., the *side B* reflectance of the (m, n) th pixel is brighter compared to the corresponding *side A* reflectance:** In this case, $R_B^0(m, n)$ has been significantly lowered to the value $R_B(m, n)$ as there is a strong provision for show-through to bright *side B* pixel from the darker *side A* pixel. But, as explained in the first case, the effect of the erroneous estimation of $R_B^0(m, n)$ will be suppressed by the low value of $R_A^0(m, n)$.
4. **Both $\mathbf{R}_A(m, n)$ and $\mathbf{R}_B(m, n)$ have high values, i.e., the (m, n) th pixel of both the sides are bright:** For this condition, the show-through occurs to a very low extent for either of the sides. Therefore, in this case,

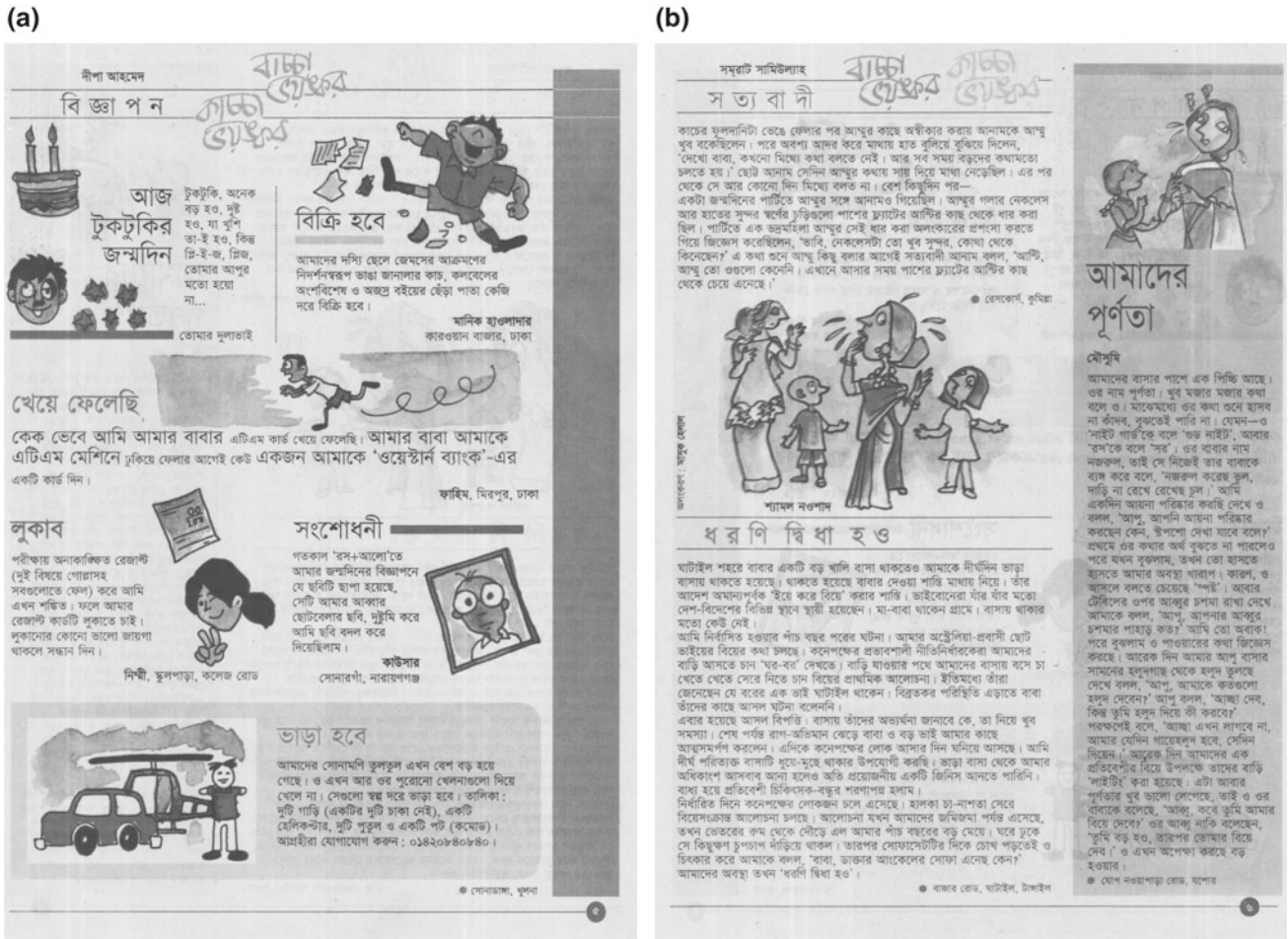


Fig. 17 Show-through corrupted document (2nd example): a side A, b side B

we can estimate $R_B(m, n)$ to be an approximation of $R_B^0(m, n)$.

Therefore, as explained above, our use of $R_B(m, n)$ as an estimate of $R_B^0(m, n)$ will not hinder the show-through correction process.

The complete show-through correction algorithm is summarized in a flow diagram given in Fig. 12.

7 Experimental results


In this section, we present show-through corrected results to demonstrate the effectiveness of the proposed mathematical model and compare its performance with that of [1]. The methods were tested for 8 pair of scanned data of duplex printed documents. A total of 3 types of papers were there all together. The documents were composite of both texts and images.

In actual applications, if an automatic document feeder is utilized for scanning the page in duplex mode, the relative alignment between the images on the two sides can be obtained from the feeder’s geometry and the detected paper edges. Alternately, techniques from image registration [15], [16] may be adopted for aligning the images. A two-stage hierarchical alignment technique, which can efficiently and accurately aligns the two sides of a document, has been presented in [17]. For our test purposes, side A and the side B are assumed to be perfectly aligned in all cases.

We present here some cropped results from only two pair of the data sets to avoid redundancy of similar results. Other test results, could be found on the website: <http://sites.google.com/site/showthroughmodel/>.

Figures 13a,b show the show-through contaminated version or the original states of the scanned documents. The images were scanned from 8" × 10.5" size pages, with a scanning resolution of 300 dpi. These images were processed using our proposed algorithm and [1]. For the simulation of our proposed algorithm, we chose $x_H = 170, y_L = 70$ and

(a)




আমাদের পূর্ণতা

মৌসুমি

আমাদের বাসার পাশে এক পিচ্চি আছে। ওর নাম পূর্ণতা। খুব মজার মজার কথা বলে ও। মাঝেমধ্যে ওর কথা শুনে হাসব না কাঁদব, বুঝতেই পারি না। যেমন—ও 'নাইট গার্ড কে বলে 'গুড নাইট', আবার 'রস'কে বলে 'সর'। ওর বাবার নাম নজরুল, তাই সে নিজেই তার বাবাকে ব্যঙ্গ করে বলে, 'নজরুল করেছে ভুল, দাড়ি না রেখে রেখেছ চুল।' আমি একদিন আয়না পরিষ্কার করছি দেখে ও বলল, 'আপু, আপনি আয়না পরিষ্কার করছেন কেন, স্টপশো দেখা যাবে বলে?' প্রথমে ওর কথার অর্থ বুঝতে না পারলেও পরে যখন বুঝলাম, তখন তো হাসতে হাসতে আমার অবস্থা খারাপ। কারণ, ও আসলে বলতে চেয়েছে 'স্পষ্ট'। আবার টেবিলের ওপর আকবুর চশমা রাখা দেখে আমাকে বলল, 'আপু, আপনার আকবুর চশমার পাহাড় কত?' আমি তো অবাক! পরে বুঝলাম ও পাওয়ারের কথা জিজ্ঞেস করছে। আরেক দিন আমার আপু বাসার সামনের হলুদগাছ থেকে হলুদ তুলছে দেখে বলল, 'আপু, আমাকে কতগুলো হলুদ দেবেন?' আপু বলল, 'আচ্ছা দেব, কিন্তু তুমি হলুদ দিয়ে কী করবে?' পরক্ষণেই বলে, 'আচ্ছা এখন লাগবে না, আমার যেদিন গায়েহলুদ হবে, সেদিন দিইন।' আরেক দিন আমাদের এক প্রতিবেশীর বিয়ে উপলক্ষে তাদের বাড়ি 'লাইটিং' করা হয়েছে। এটা আবার পূর্ণতার খুব ভালো লেগেছে, তাই ও ওর বাবাকে বলেছে, 'আবু, কবে তুমি আমার বিয়ে দেবে?' ওর আবু নাকি বলেছেন, 'তুমি বড় হও, তারপর তোমার বিয়ে দেব।' ও এখন অপেক্ষা করছে বড় হওয়ার।

● ঘোপ নওয়াপাড়া রোড, যশোর

(b)




আমাদের পূর্ণতা

মৌসুমি

আমাদের বাসার পাশে এক পিচ্চি আছে। ওর নাম পূর্ণতা। খুব মজার মজার কথা বলে ও। মাঝেমধ্যে ওর কথা শুনে হাসব না কাঁদব, বুঝতেই পারি না। যেমন—ও 'নাইট গার্ড কে বলে 'গুড নাইট', আবার 'রস'কে বলে 'সর'। ওর বাবার নাম নজরুল, তাই সে নিজেই তার বাবাকে ব্যঙ্গ করে বলে, 'নজরুল করেছে ভুল, দাড়ি না রেখে রেখেছ চুল।' আমি একদিন আয়না পরিষ্কার করছি দেখে ও বলল, 'আপু, আপনি আয়না পরিষ্কার করছেন কেন, স্টপশো দেখা যাবে বলে?' প্রথমে ওর কথার অর্থ বুঝতে না পারলেও পরে যখন বুঝলাম, তখন তো হাসতে হাসতে আমার অবস্থা খারাপ। কারণ, ও আসলে বলতে চেয়েছে 'স্পষ্ট'। আবার টেবিলের ওপর আকবুর চশমা রাখা দেখে আমাকে বলল, 'আপু, আপনার আকবুর চশমার পাহাড় কত?' আমি তো অবাক! পরে বুঝলাম ও পাওয়ারের কথা জিজ্ঞেস করছে। আরেক দিন আমার আপু বাসার সামনের হলুদগাছ থেকে হলুদ তুলছে দেখে বলল, 'আপু, আমাকে কতগুলো হলুদ দেবেন?' আপু বলল, 'আচ্ছা দেব, কিন্তু তুমি হলুদ দিয়ে কী করবে?' পরক্ষণেই বলে, 'আচ্ছা এখন লাগবে না, আমার যেদিন গায়েহলুদ হবে, সেদিন দিইন।' আরেক দিন আমাদের এক প্রতিবেশীর বিয়ে উপলক্ষে তাদের বাড়ি 'লাইটিং' করা হয়েছে। এটা আবার পূর্ণতার খুব ভালো লেগেছে, তাই ও ওর বাবাকে বলেছে, 'আবু, কবে তুমি আমার বিয়ে দেবে?' ওর আবু নাকি বলেছেন, 'তুমি বড় হও, তারপর তোমার বিয়ে দেব।' ও এখন অপেক্ষা করছে বড় হওয়ার।

● ঘোপ নওয়াপাড়া রোড, যশোর

(c)



আমাদের পূর্ণতা

মৌসুমি

আমাদের বাসার পাশে এক পিচ্চি আছে। ওর নাম পূর্ণতা। খুব মজার মজার কথা বলে ও। মাঝেমধ্যে ওর কথা শুনে হাসব না কাঁদব, বুঝতেই পারি না। যেমন—ও 'নাইট গার্ড কে বলে 'গুড নাইট', আবার 'রস'কে বলে 'সর'। ওর বাবার নাম নজরুল, তাই সে নিজেই তার বাবাকে ব্যঙ্গ করে বলে, 'নজরুল করেছে ভুল, দাড়ি না রেখে রেখেছ চুল।' আমি একদিন আয়না পরিষ্কার করছি দেখে ও বলল, 'আপু, আপনি আয়না পরিষ্কার করছেন কেন, স্টপশো দেখা যাবে বলে?' প্রথমে ওর কথার অর্থ বুঝতে না পারলেও পরে যখন বুঝলাম, তখন তো হাসতে হাসতে আমার অবস্থা খারাপ। কারণ, ও আসলে বলতে চেয়েছে 'স্পষ্ট'। আবার টেবিলের ওপর আকবুর চশমা রাখা দেখে আমাকে বলল, 'আপু, আপনার আকবুর চশমার পাহাড় কত?' আমি তো অবাক! পরে বুঝলাম ও পাওয়ারের কথা জিজ্ঞেস করছে। আরেক দিন আমার আপু বাসার সামনের হলুদগাছ থেকে হলুদ তুলছে দেখে বলল, 'আপু, আমাকে কতগুলো হলুদ দেবেন?' আপু বলল, 'আচ্ছা দেব, কিন্তু তুমি হলুদ দিয়ে কী করবে?' পরক্ষণেই বলে, 'আচ্ছা এখন লাগবে না, আমার যেদিন গায়েহলুদ হবে, সেদিন দিইন।' আরেক দিন আমাদের এক প্রতিবেশীর বিয়ে উপলক্ষে তাদের বাড়ি 'লাইটিং' করা হয়েছে। এটা আবার পূর্ণতার খুব ভালো লেগেছে, তাই ও ওর বাবাকে বলেছে, 'আবু, কবে তুমি আমার বিয়ে দেবে?' ওর আবু নাকি বলেছেন, 'তুমি বড় হও, তারপর তোমার বিয়ে দেব।' ও এখন অপেক্ষা করছে বড় হওয়ার।

● ঘোপ নওয়াপাড়া রোড, যশোর

Fig. 18 Comparative results on show-through correction (from 2nd example)

$\gamma = 0.9$. For the adaptive filtering method in [1], the step-size, $\mu = 0.8 \times 10^{-3}$ and a filter-size of $[11 \times 11]$ were selected.

Some regions of the show-through corrected images are selected for illustration. Figure 14a shows the comparison between the results obtained by our method and the adaptive filter method of [1]. Figure 14a shows the selected regions from the scanned image of Fig. 13a. Figures 14b,c show the results obtained by the proposed method and the method of [1], respectively. As marked in Figs. 14ci and cii, the image has been damaged while correcting for show-through noise using the algorithm of [1]. In these cases, the adaptive filter coefficients yielded overestimated PSF values. On the other hand, our proposed method yields visibly much improved results as can be seen in Figs. 14bi and bii.

Figures 15 and 16 show results obtained by our proposed method for some more selected regions of the scanned images. Almost no artifact of cleaning process is visible at actual size of the images for this data set.

The last set of images deserve special annotation. Figure 15a shows a portion of the *side A* document, which contains image and text on the *side A* as well as on the corresponding *side B*. Figure 15b shows the results. Here, we can clearly see the effectiveness of the proposed algorithm in such a complicated situation.

Figure 16a shows a portion of the document, which is much simpler than the previous one containing only text on *side A* and *side B*. Figure 16b shows the results of the proposed algorithm in such a case. As evident, the show-through effect is removed without introducing any visible artifact.

Figures 17a,b show the show-through contaminated version or the original states of the *side A* and *side B*, respectively, of the 2nd set of scanned documents. The images were scanned from $8'' \times 10.5''$ size pages, with a scanning resolution of 300 dpi, as before. These images were processed using our proposed algorithm and [1].

Figure 18 shows comparative results obtained by our method and that of [1] taken from a selected region of *side B* of the 2nd data set. Figure 18a is the show-through corrupted image. Figure 18b, c show the cleaned image using our technique and that obtained by applying [1], respectively. The algorithm of [1] has left noticeable artifacts as clearly visible in Fig. 18c. On the other hand, the proposed method gives much better result in this case.

Figures 19 and 20 show results obtained by our proposed method for some more selected regions of the 2nd data set. Almost no artifact of cleaning process is visible at actual size of the images for this data set. In this particular case, the background of the original image is dark. The show-through reflectance is not prominently darker than the background as was the case in the previous example. We need to choose the white paper reflectance to be high as possible. Therefore, γ should be high in the present case. We adopted $\gamma = 0.95$.

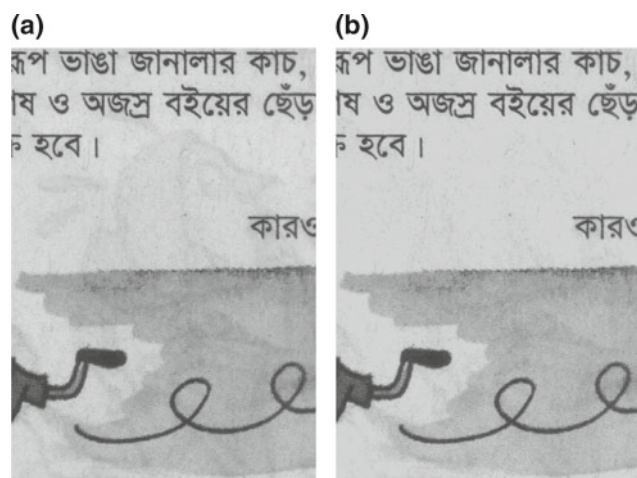


Fig. 19 a Image distorted by show-through, b results obtained using the proposed method

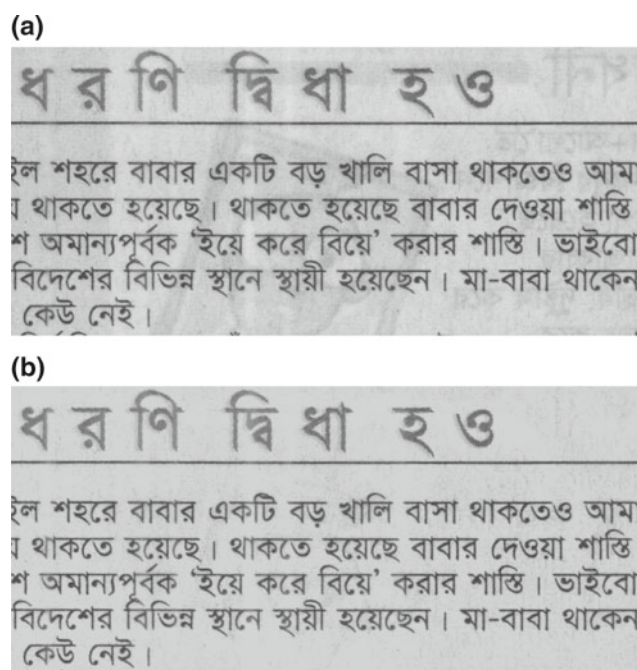


Fig. 20 a Image distorted by show-through, b results obtained using the proposed method

A preference-based informal perceptual test was performed to compare the results of our proposed algorithm to that of [1]. The scanned document, the cleaned document using our algorithm and [1] were shown to 15 neutral subjects. Each were provided with 10 sets (i.e., 3 documents in each set) of documents. The observers were asked to identify the better document from the visual perspective from each set of data. The results are shown in Table 1. Cleaned images obtained by using our algorithm were preferred in almost all cases.

Table 1 Comparative results of preference test

| | Preferred (%) |
|---|---------------|
| Scanned document | 3.33 |
| Cleaned document using the proposed algorithm | 94.67 |
| Cleaned document using [1] | 2 |

There was very little or no blurring effect in all the data sets that were used in the experiments. The proposed algorithm performs satisfactorily for little or no blurring as observed in the presented and other experimental results.

8 Conclusions

In this paper, a novel physics-based mathematical model for show-through has been derived through detailed analysis of the phenomenon. A probabilistic approach has been adopted to calculate the scattering of photons in the ink-paper-backing scanning system. The relationship between the reflectances of the two sides of the paper has been established using the photon fluxes, which have been calculated by taking into account the effect of both the printed sides of the document. The mathematical model for show-through is represented by the above-mentioned relationship between the reflectances. The model has been verified by showing how well it follows the practical show-through data. We have also proposed, a show-through correction algorithm in this paper where we estimated the model parameters from the joint histogram formed from the scans of the two sides of a duplex printed page and then cleaned the scanned document using the proposed model. The results of experimental tests have demonstrated superior performance of our proposed show-through correction technique when compared to [1] for scans of documents with complex distribution of texts and shades.

Appendix A: Calculating $\bar{\chi}_{AB}(f_A, f_B)$

Consider that the light is incident on *side A* and is emerging from *side B*. The photon flux will pass through as explained in the following 3 steps:

1. Light is incident on Σ_{2A} and Σ_{1A} with probabilities of $(1 - f_A)$ and f_A , respectively. Therefore, the light will go through *side A* whose transmittance is $((1 - f_A) + (f_A)T_{iA})$.
2. Then the light passes through the paper of transmittance T_w .
3. Finally, the photon flux will go through *side B* whose transmittance is $((1 - f_B) + (f_B)T_{iB})$.

Therefore, the transmission ratio of the combined ‘paper and ink system’ of our concerned region can be expressed as,

$$\begin{aligned} \bar{\chi}_{AB}(f_A, f_B) &= ((1 - f_A) + (f_A)T_{iA}) \times T_w \times ((1 - f_B) + (f_B)T_{iB}) \\ &= T_w [(1 - f_A) + (f_A)T_{iA}][(1 - f_B) + (f_B)T_{iB}] \\ &= T_w [1 - f_A(1 - T_{iA})][1 - f_B(1 - T_{iB})] \end{aligned} \tag{A.1}$$

Appendix B: Joint and conditional PDFs in terms of f_A and \bar{p}_A

We intend to express the joint Probability Density Functions (PDFs), $P(m, n)$, and the conditional PDFs, ${}_m P_n$, in terms of f_A and \bar{p}_A . Here, $m, n = 1A$ or $2A$.

From (36), we can write,

$$\bar{p}_A = \frac{{}_2A P_{1A}}{f_A} = \frac{{}_1A P_{2A}}{1 - f_A} \tag{B.1}$$

Now, the *self-reflectance* of *side A* with no printing on that side, can be expressed as,

$$R_{AA}^w = R_{AA}(T_{iA} = 1) \tag{B.2}$$

From [13], the probability of photons returning directly from both Σ_{1A} and Σ_{2A} of *side A* (i.e., these photons do not pass through *side A*) would be equal to the *self-reflectance* of *side A*, R_{AA}^w . Therefore,

$${}_1A P_{1A} + {}_1A P_{2A} = R_{AA}^w \tag{B.3}$$

$${}_2A P_{1A} + {}_2A P_{2A} = R_{AA}^w \tag{B.4}$$

Using (B.1), (B.3) and (B.4), we have,

$$\left. \begin{aligned} {}_1A P_{1A} &= \frac{P(1A,1A)}{f_A} = \bar{p}_A f_A + (R_{AA}^w - \bar{p}_A) \\ {}_2A P_{2A} &= \frac{P(2A,2A)}{1 - f_A} = -\bar{p}_A f_A + R_{AA}^w \\ {}_1A P_{2A} &= \frac{P(1A,2A)}{f_A} = -\bar{p}_A f_A + \bar{p}_A \end{aligned} \right\} \tag{B.5}$$

We can now replace the conditional probabilities ${}_1A P_{1A}$, ${}_2A P_{2A}$ and ${}_1A P_{1A}$ with expressions containing f_A and the constants \bar{p}_A and R_{AA}^w .

Appendix C: Expressing f_A as a function of *side A* reflectance (R_A^0)

Rewriting (45), we have,

$$\begin{aligned} R_A^0 &= R_A^w - R_A^w(1 - T_{iA}^2)f_A - f_A(1 - T_{iA})^2 k_{pA} \\ &\quad + f_A^2 k_{pA}(1 - T_{iA})^2 \end{aligned} \tag{C.1}$$

Now (C.1) can be expressed as follows,

$$af_A^2 + bf_A + c = 0 \tag{C.2}$$

where,

$$\left. \begin{aligned} a &= k_{pA}(1 - T_{iA})^2 \\ b &= -R_A^w(1 - T_{iA}^2) - (1 - T_{iA})^2 k_{pA} \\ c &= R_A^w - R_A^0 \end{aligned} \right\} \tag{C.3}$$

Here, b is a negative quantity. a and c are positive quantities. c will be positive because reflectance of side A will not be larger than white paper reflectance of side A, i.e., $R_A^w \geq R_A^0$. Now, from (C.2), we have,

$$\begin{aligned} f_A &= \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} \\ &= \frac{b_1 \pm \sqrt{b_1^2 - 4a_1c_1}}{2a_1} \\ &= \frac{b_1}{2a_1} \pm \sqrt{\left(\frac{b_1}{2a_1}\right)^2 - \frac{c_1}{a_1}} \end{aligned} \tag{C.4}$$

Here, $a_1 = a, b_1 = -b$ and $c_1 = c$. Now, a_1, b_1 and c_1 are all positive quantities which will simplify further calculations. Now,

$$\begin{aligned} \frac{b_1}{2a_1} &= \frac{k_{pA}(1 - T_{iA})^2 + R_A^w(1 - T_{iA}^2)}{2k_{pA}(1 - T_{iA})^2} \\ &= \frac{1}{2} + \frac{R_A^w(1 + T_{iA})}{2k_{pA}(1 - T_{iA})} \end{aligned} \tag{C.5}$$

Again,

$$\frac{c_1}{a_1} = \frac{R_A^w - R_A^0}{k_{pA}(1 - T_{iA})^2} \tag{C.6}$$

Therefore,

$$\begin{aligned} \left. \frac{c_1}{a_1} \right|_{\max} &= \frac{R_A^w - (R_A^0|_{\min})}{k_{pA}(1 - T_{iA})^2} \\ &= \frac{R_A^w - T_{iA}^2 R_A^w}{k_{pA}(1 - T_{iA})^2} \quad [\because (R_A^0|_{\min}) = T_{iA}^2 R_A^w, \\ &\hspace{10em} \text{as explained below}] \\ &= \frac{R_A^w(1 + T_{iA})}{k_{pA}(1 - T_{iA})} \end{aligned} \tag{C.7}$$

and,

$$\begin{aligned} \left. \frac{c_1}{a_1} \right|_{\min} &= \frac{R_A^w - (R_A^0|_{\max})}{k_{pA}(1 - T_{iA})^2} \\ &= \frac{R_A^w - R_A^w}{k_{pA}(1 - T_{iA})^2} \quad [\because (R_A^0|_{\max}) = R_A^w, \\ &\hspace{10em} \text{as explained below}] \\ &= 0 \end{aligned} \tag{C.8}$$

In the above calculations for $\left(\frac{c_1}{a_1}\right)_{\max}$ and $\left(\frac{c_1}{a_1}\right)_{\min}$, the reflectance $R_A^0|_{\min}$ corresponds to the lowest value of R_A^0 which is found when the whole *specific region* of *side A*

is covered with ink, i.e., when $f_A = 1$. And, $R_A^0|_{\max}$ corresponds to the highest value of R_A^0 which is found when there is no printing on *side A*, i.e., $f_A = 0$. Therefore, using (C.1), we have,

$$\begin{aligned} R_A^0|_{\min} &= R_A^0(f_A = 1) = T_{iA}^2 R_A^w \\ R_A^0|_{\max} &= R_A^0(f_A = 0) = R_A^w \end{aligned}$$

Now, for $0 \leq f_A \leq 1$, the following cases are required to be satisfied:

1. For $R_A^0 = (R_A^0|_{\min})$, calculations yield $\frac{c_1}{a_1} = \left(\frac{c_1}{a_1}\right)_{\max}$.
And, $f_A = (f_A|_{\max}) = 1$ should be found when $R_A^0 = (R_A^0|_{\min})$.
2. For $R_A^0 = (R_A^0|_{\max})$, calculations yield $\frac{c_1}{a_1} = \left(\frac{c_1}{a_1}\right)_{\min}$.
And, $f_A = (f_A|_{\min}) = 0$ should be found when $R_A^0 = (R_A^0|_{\max})$.

To satisfy the above-mentioned conditions for $0 \leq f_A \leq 1$, in (C.4), only the $-ve$ sign will be taken,

$$f_A = \frac{b_1 - \sqrt{b_1^2 - 4a_1c_1}}{2a_1} \tag{C.9}$$

Therefore, we have expressed f_A as a function of the *side A* reflectance R_A^0 .

References

1. Sharma, G.: Show-through cancellation in scans of duplex printed documents. *IEEE Trans. Image Process.* **10**, 736–754 (2001)
2. Ophir, B., Malah, D.: Improved cross-talk cancellation in scanned images by adaptive decorrelation. In: *Proceedings of 23rd IEEE convention of electrical and electronics engineers in Israel*, pp. 388–391, Sept 2004
3. Ophir, B., Malah, D.: Show-through cancellation in scanned images using blind source separation techniques. In: *IEEE international conference on image process*, vol. 3, pp. 233–236, Sept 2007
4. Merrikh-Bayat, F., Babaie-Zadeh, M., Jutten, C.: A nonlinear blind source separation solution for removing the show-through effect in the scanned documents. *European signal processing conference, EUSIPCO*, Sept 2008
5. Besag, J.: On the statistical analysis of dirty pictures. *J. R. Stat. Soc. B* **48**(3), 259–302 (1986)
6. Hosseini, S., Deville, Y.: Blind maximum likelihood separation of a linear-quadratic mixture. In: *Proceedings of 5th international conference on independent component analysis and blind source separation (ICA'04)*, pp. 694–701, Sept 2004
7. Almeida, M.S.C., Almeida, L.B.: Wavelet-based separation of non-linear show-through and bleed-through image mixtures. *Neuro-computing*, ISSN:0925-2312, vol. 72, no. 1–3, pp. 57–70, Dec 2008
8. Khan, M.R., Imtiaz, H., Hasan, M.K.: Show-through correction in scanned images using joint histogram. *Signal, Image and Video Processing (SIViP)* **4**, 337–351 (2010)

9. Yule, J.A.C., Nielsen[sic] W.J.: The penetration of light into paper and its effect on halftone reproduction. In: TAGA Proc., pp. 65–76 (1951)
10. Ruckdeschel, F.R., Hauser, O.G.: Yule-Nielsen effect in printing: A physical analysis. *Appl. Opt.* **17**(21) (1978)
11. Maltz, M.: Light scattering in xerographic images. *J. Appl. Photogr. Eng.* **9**(3), 83–89 (1983)
12. Yang, L., Kruse, B.: Ink penetration and its effect on printing. In: SPIEs 12th annual international symposium, vol. 3963, pp. 365–375 Dec 1999
13. Yang, L., Kruse, B., Lenz, R.: Light scattering and ink penetration effects on tone reproduction. *J. Opt. Soc. Am. A Opt. Image Sci. Vis.* **18**(2), 360–366 (2001)
14. Gonzalez, R.C., Woods, R.E.: *Digital Image Processing*. Prentice Hall of India, New Delhi (2002)
15. Brown L., G.: A survey of image registration techniques. *ACM Comput. Surv.* **24**(4), 325–376 (1992)
16. Tian, Q., Huhns, M.N.: Algorithms for subpixel registration. *CVGIP: Graph Models Image Process* **35**(2), 220–233 (1986)
17. Wang, J., Brown, M.S., Tan, C.L.: Accurate alignment of double-sided manuscripts for bleed-through removal. In: The eighth IAPR international workshop on document analysis systems. DAS '08, pp. 69–75, Sept 2008