

Error Bounds on the Reconstruction of Binary Images from Low Resolution Scans

Wagner Fortes^{1,2} and K. Joost Batenburg^{1,3}

¹ Centrum Wiskunde & Informatica, Amsterdam, The Netherlands

² Mathematical Institute, Leiden University, The Netherlands

³ IBBT-Vision Lab, University of Antwerp, Belgium

Abstract. In this paper, we consider the problem of reconstructing a high-resolution binary image from several low-resolution scans. Each of the pixels in a low-resolution scan yields the value of the sum of the pixels in a rectangular region of the high-resolution image. For any given set of such pixel sums, we derive an upper bound on the difference between a certain binary image which can be computed efficiently, and *any* binary image that corresponds with the given measurements. We also derive a bound on the difference between any two binary images having these pixel sums. Both bounds are evaluated experimentally for different geometrical settings, based on simulated scan data for a range of images.

Keywords: Image reconstruction, Error bounds, Binary image, Rectangular scan.

1 Introduction

Black-and-white images, also called *binary images*, occur in a wide range of imaging applications. In many such applications, the images are actually acquired as grey level images by a scanning device. When scanning text, for example, binary characters are often scanned by a grey level scanner. When taking pictures of numberplates using a low resolution digital camera, the structure of the binary characters may even be unrecognizable in the resulting grey level images. Another example can be found in the single-pixel camera, which has recently been proposed within the framework of compressive sensing. Instead of recording individual fine-resolution pixels, such a camera records the total intensity over various areas of the object being photographed [8,11].

If several such grey level images are available, each representing a low resolution scan of some unknown "original" binary image, one can attempt to reconstruct the binary image by combining the information from multiple scans [2,5,6]. In particular if the relative position of the different scans is well-known, this may lead to a high quality reconstruction. However, if the number of low resolution images available is relatively small in comparison with the resolution needed to properly represent the binary image, this reconstruction problem can be highly underdetermined. In such cases, many binary images can exist that correspond with the same scanned grey level data. At present, no useful bounds are available that can guarantee that the reconstructed image is actually close to the unknown original image.

In a recent paper [1], the authors presented bounds for binary image reconstruction in *tomography* (i.e. from projection data) that allow to bound the error between any two binary solutions, and therefore the error between the reconstructed binary image and the unknown original image. The proposed methodology is quite general and can potentially be extended to other imaging problems. As an intermediate step towards a general framework for bounding errors in binary image reconstruction, we apply the key concepts here to the problem of reconstructing binary images from low resolution scans.

2 Notation and Concepts

Let $A \subset \mathbb{Z}^2$ be a finite set, called the *reconstruction area*. We consider the problem of reconstructing a binary image defined on A , represented by a function $F : A \rightarrow \{0, 1\}$. A high resolution binary image defined on A will be reconstructed from several low resolution scans. The value of each pixel in such a scan corresponds with the summed intensity over all pixels in the corresponding region of the binary image. For simplicity, we assume here that the boundaries of the low resolution pixels coincide exactly with pixel boundaries in the high resolution binary image. We call a set $S \subset A$ a *window* of the reconstruction area. Let $\mathcal{S} = 2^A$, the set of all windows of A . We call a set $\mathbf{S} \subset \mathcal{S}$ of windows a *partition* of A if

(i) $S \cap T = \emptyset$ for all $S, T \in \mathbf{S}$ and

(ii) $\bigcup_{S \in \mathbf{S}} S = A$. We are also interested in the subsets of \mathcal{S} which satisfy the property (i) but do not necessarily satisfy (ii). Such a subset will be called a *partial partition*. For $S \subset A$, define

$$P_F(S) = \sum_{(i,j) \in S} F(i,j). \quad (1)$$

We refer to the values $P_F(S)$ as *window-sums*. Note that our model for computing the window sums does not take certain properties of the imaging system, such as the detector point spread function, into account. However, the proposed methodology can easily be extended to include such effects, as long as they are linear. The *reconstruction problem* consists of finding an image F that has prescribed window-sum for a set \mathbf{S} of windows. The existence and uniqueness of the solution of the general reconstruction problem is not guaranteed, in general.

To simplify the notation, the reconstruction problem can be formulated using linear algebra notation, which will be used in the forthcoming sections. Since there is an one-to-one mapping, say χ , from A to $\{1, \dots, n\}$, the image F can be represented as a vector $\mathbf{x} = (x_j) \in \mathbb{R}^n$, where $n = \#A$ is the cardinality of A . We refer to the entries of \mathbf{x} as *pixels*. A *binary image* on A corresponds with a vector $\bar{\mathbf{x}} \in \{0, 1\}^n$.

For a given set $\mathbf{S} \subset \mathcal{S}$ and an image $\mathbf{x} \in \mathbb{R}^n$, the combined set of window sums results in a vector $\mathbf{p} = (p_i) \in \mathbb{R}^m$, where m represents the number of window-sums taken. As the operator $P_F(S)$ is linear, the mapping from an image to its window sums can be represented by a matrix $\mathbf{W} = (w_{ij}) \in \mathbb{R}^{m \times n}$, which we call the *scan matrix*. The entry w_{ij} represents the weight of the contribution of x_j to the window-sum i .

Then, the general reconstruction problem can be stated as finding a solution of the system

$$\mathbf{W}\mathbf{x} = \mathbf{p}$$

for given window-sum data \mathbf{p} . In the binary image reconstruction problem, one seeks a binary solution of the system. For a given scan matrix \mathbf{W} and a window-sum vector \mathbf{p} , let $\mathcal{T}_{\mathbf{W}}(\mathbf{p}) := \{\mathbf{x} \in \mathbb{R}^n : \mathbf{W}\mathbf{x} = \mathbf{p}\}$, the set of all real-valued solutions corresponding with the given data, and let $\tilde{\mathcal{T}}_{\mathbf{W}}(\mathbf{p}) := \mathcal{T}_{\mathbf{W}}(\mathbf{p}) \cap \{0, 1\}^n$, the set of *binary solutions* of the system.

As the scan matrix is typically not a square matrix, and also does not have full rank, it does not have an inverse. We recall that the *Moore-Penrose pseudo inverse* of an $m \times n$ matrix \mathbf{A} is an $n \times m$ matrix \mathbf{A}^\dagger , which can be uniquely characterized by the two geometric conditions

$$\mathbf{A}^\dagger \mathbf{b} \perp \mathcal{N}(\mathbf{A}) \quad \text{and} \quad (\mathbf{I} - \mathbf{A}\mathbf{A}^\dagger)\mathbf{b} \perp \mathcal{R}(\mathbf{A}) \quad \forall \mathbf{b} \in \mathbb{R}^m,$$

where $\mathcal{N}(\mathbf{A})$ is the nullspace of \mathbf{A} and $\mathcal{R}(\mathbf{A})$ is the range of \mathbf{A} , [4, page 15].

Let $\mathbf{x}^* = \mathbf{W}^\dagger \mathbf{p}$. Then \mathbf{x}^* also has the property (see Chapter 3 of [3]) that it is the minimal Euclidean norm solution of the system $\mathbf{W}\mathbf{x} = \mathbf{p}$, if it exists. We call \mathbf{x}^* the *central reconstruction* of \mathbf{p} . The central reconstruction plays an important role in the bounds we derive for the binary reconstruction problem.

The description of the general reconstruction problem given above is quite broad and we will now specify the scan model by which we define the scan matrix \mathbf{W} and the window-sum vector \mathbf{p} , in order to model the problem of reconstructing high resolution images from low resolution scans.

Put $A = \{(i, j) \in \mathbb{Z}^2 \mid 0 \leq i < l, 0 \leq j < h\}$. Let $1 \leq p \leq l$, and $1 \leq q \leq h$. For $0 \leq i < l$, $0 \leq j < h$, define a rectangular set of pixels of size $p \times q$ by

$$S_{i,j}^{p,q} = \{(i + c, j + r) \mid 0 \leq c < p, 0 \leq r < q\}.$$

Each pixel in a low resolution scan corresponds to a *window* in our framework. It provides information about the summed intensity in a rectangular set of pixels of the scanned high resolution image. Adjacent low resolution pixels are connected and do not overlap. For $0 \leq a < p$ and $0 \leq b < q$, define

$$\mathbf{S}^{a,b} = \{S_{a+ip, b+jq}^{p,q} \mid a + ip < l, b + jq < h\}. \tag{2}$$

Each set $\mathbf{S}^{a,b}$ is a partial partition. Its elements correspond to pixels of the low resolution image. Let us now assume that several such low resolution images are available. Then the total set \mathbf{S} of window-sums consists of the union of partial partitions $\mathbf{S}^d := \mathbf{S}^{a_d, b_d}$ for $d \in \{1, \dots, k\}$. These concepts are illustrated in Fig. 1. Fig. 1a shows a single window $S_{a,b}$, whereas Fig. 1b shows the corresponding partial partition $\mathbf{S}^{a,b}$ formed by a tiling of its translates, where windows that cross the boundary of the image are not allowed. Fig. 1c shows two windows that are in separate partial partitions.

For $1 \leq d \leq k$, define the set of indices of the pixels x_j that were scanned by a partial partition \mathbf{S}^d as $I_d := \{j \mid \chi^{-1}(j) \in \cup_{S \in \mathbf{S}^d} S\}$ and its complement $\bar{I}_d := \{1, \dots, n\} \setminus I_d$.

As already mentioned, this linear scanning model can be modeled by a linear system of equations $\mathbf{W}\mathbf{x} = \mathbf{p}$. The matrix \mathbf{W} and the window-sum \mathbf{p} can be decomposed into k blocks as

$$\mathbf{W} = \begin{pmatrix} \mathbf{W}^1 \\ \vdots \\ \mathbf{W}^k \end{pmatrix}, \quad \mathbf{p} = \begin{pmatrix} \mathbf{p}^1 \\ \vdots \\ \mathbf{p}^k \end{pmatrix}, \tag{3}$$

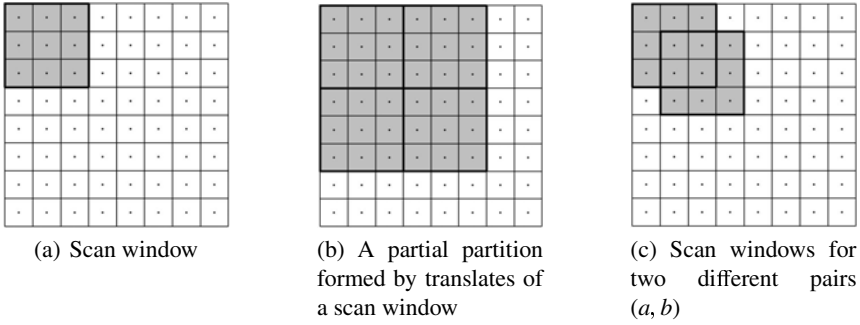


Fig. 1. Rectangular scanning

where each block W^d ($d = 1, \dots, k$) represents the scanning of the image with a rectangular window as defined by S^d and each block p^d represents the corresponding window-sums $P_F(S)$ for $S \in S^d$.

3 Error Bounds

Without loss of generality, we assume that all pixels in A are contained in at least one window. Clearly, no bounds can be given for those pixels that are not scanned at all, and they are removed from the analysis. As each set S^d samples a collection of disjoint subsets of A , the norm of the scanned binary image can be bound from above by the available window-sums:

Proposition 1. *Let $\bar{x} \in \tilde{\mathcal{T}}_W(p)$. Then, $\|\bar{x}\|_2^2 = \|\bar{x}\|_1 \leq \|p^d\|_1 + \#\bar{I}_d$ for all $1 \leq d \leq k$.*

The norm of any binary solution can therefore be estimated by summation of the window-sums in p^d and its accuracy increases with the number of scanned pixels included in the partial partition S^d .

In the next Theorem we will use Prop. 1 to show that all binary solutions of the linear system $Wx = p$ have bounded distance to the central reconstruction x^* .

Lemma 1. *Let $\bar{x} \in \tilde{\mathcal{T}}_W(p)$ and $x^* = W^\dagger p$. Put $R := \min_{1 \leq d \leq k} R_d$, where $R_d := \sqrt{\|p^d\|_1 + \#\bar{I}_d - \|x^*\|_2^2}$. Then, $\|\bar{x} - x^*\|_2 \leq R$.*

Proof. From the definition of x^* we have $(\bar{x} - x^*) \in \mathcal{N}(W)$, and $x^* \perp (\bar{x} - x^*)$. Combining Pythagoras' theorem and Prop. 1 yields the theorem. □

We will now consider the image that is obtained by rounding each entry of x^* to the nearest binary value. Let $\langle \alpha \rangle = \min(|\alpha|, |\alpha - 1|)$ for $\alpha \in \mathbb{R}$, and put $U = \sqrt{\sum_{i=1}^n \langle x_i^* \rangle^2}$, i.e., the Euclidean distance from x^* to the nearest binary vector.

Let $\bar{r} \in \{0, 1\}^n$ such that $\|\bar{r} - x^*\|_2 = U$, i.e., \bar{r} is among the binary vectors that are nearest to x^* in the Euclidean sense. If $R > U$ and $R - U$ is small, it is possible to say

that a fraction of the rounded values are correct, i.e., to provide an upper bound on the number of pixel differences between any solution in $\tilde{\mathcal{T}}_W(\mathbf{p})$ and $\bar{\mathbf{r}}$.

Suppose that $\bar{\mathbf{x}} \in \tilde{\mathcal{T}}_W(\mathbf{p})$ and that $\bar{r}_i = 1$ whereas $\bar{x}_i = 0$. Note that we have $x_i^* \geq \frac{1}{2}$. Put $\bar{\mathbf{r}} := \bar{\mathbf{r}}$ and then set \bar{r}_i to 0. We then have $\|\bar{\mathbf{r}} - \mathbf{x}^*\|_2^2 = \|\bar{\mathbf{r}} - \mathbf{x}^*\|_2^2 - |x_i^* - 1|^2 + |x_i^*|^2 = \|\bar{\mathbf{r}} - \mathbf{x}^*\|_2^2 + 2x_i^* - 1$. Similarly, if $\bar{r}_i = 0$, then the squared Euclidean distance increases by $1 - 2x_i^*$ by setting pixel i to 1. Each time an entry i of $\bar{\mathbf{r}}$ is changed, the squared Euclidean distance to \mathbf{x}^* increases by $b_i := |2x_i^* - 1|$.

As the Euclidean distance from \mathbf{x}^* to $\bar{\mathbf{x}}$ is no greater than R , a bound can be derived on the maximal number of pixels in $\bar{\mathbf{r}}$ that must be changed to move from $\bar{\mathbf{r}}$ to $\bar{\mathbf{x}}$. Let us order the values b_i ($i = 1, \dots, n$) such that $b_i \leq b_{i+1}$ for $1 \leq i \leq n - 1$. Assuming that $\tilde{\mathcal{T}}_W(\mathbf{p}) \neq \emptyset$, we have $R \geq \|\bar{\mathbf{r}} - \mathbf{x}^*\|_2$ and the change of s entries of $\bar{\mathbf{r}}$ would increase the distance between $\bar{\mathbf{r}}$ and \mathbf{x}^* such that $R^2 \geq \|\bar{\mathbf{r}} - \mathbf{x}^*\|_2^2 + \sum_{j=1}^s b_j$.

Theorem 1. *Let $\bar{\mathbf{r}}$, $\bar{\mathbf{x}}$ and b_i ($i = 1, \dots, n$) be as defined above. Choose s such that*

$$\sum_{i=1}^s b_i \leq R^2 - \|\bar{\mathbf{r}} - \mathbf{x}^*\|_2^2 < \sum_{j=1}^{s+1} b_j. \tag{4}$$

Then at most s pixels can have the wrong value in $\bar{\mathbf{r}}$ with respect to $\bar{\mathbf{x}}$ and at least $n - s$ pixels must have the correct value.

Proof. Due to the increasing order of the b_i 's, changing more than s pixels in $\bar{\mathbf{r}}$ will result in a vector $\bar{\mathbf{r}}$ for which $\|\bar{\mathbf{r}} - \mathbf{x}^*\|_2 > R$, which cannot be an element of $\tilde{\mathcal{T}}_W(\mathbf{p})$. \square

Theorem 1 bounds the number of pixel differences between $\bar{\mathbf{x}}$ and $\bar{\mathbf{r}}$, and between $\bar{\mathbf{y}}$ and $\bar{\mathbf{r}}$. When using these two bounds to determine an upper bound on the number of differences between $\bar{\mathbf{x}}$ and $\bar{\mathbf{y}}$, we can assume that these two sets of pixel differences are disjoint, as otherwise the difference between $\bar{\mathbf{x}}$ and $\bar{\mathbf{y}}$ will only be smaller. This observation leads to the following corollary:

Corollary 1. *Let $\bar{\mathbf{r}}$ and b_i ($i = 1, \dots, n$) be as defined above. Let $\bar{\mathbf{x}}, \bar{\mathbf{y}} \in \tilde{\mathcal{T}}_W(\mathbf{p})$. Choose t such that*

$$\sum_{i=1}^t b_i \leq 2(R^2 - \|\bar{\mathbf{r}} - \mathbf{x}^*\|_2^2) < \sum_{j=1}^{t+1} b_j. \tag{5}$$

Then at most t pixels can be different between $\bar{\mathbf{x}}$ and $\bar{\mathbf{y}}$.

4 Experiments and Results

A series of experiments was performed to investigate the of the bounds given in Theorem 1 and Corollary 1, for several test images. The experiments are all based on simulated data obtained by computing the low-resolution scans of a series of test images (called *phantoms*), shown in Fig. 2. All phantoms have a size of 512x512 pixels.

In each experiment, the central reconstruction \mathbf{x}^* was first computed using the CGLS algorithm [9]. Depending on the experiment, this computation took from a few seconds, up to around 50s on a standard PC. The binary vector $\bar{\mathbf{r}}$ was computed by rounding \mathbf{x}^* to

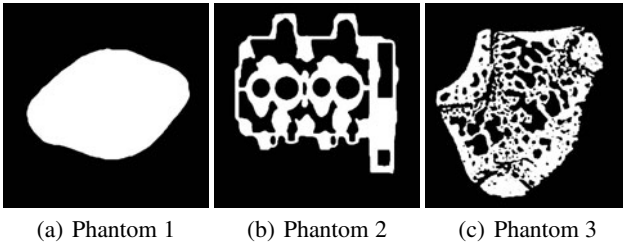


Fig. 2. Original phantom images used for the experiments

the nearest binary vector (choosing $\bar{r}_i = 0$ if $x_i^* = \frac{1}{2}$). The upper bound s from Theorem 1 on the number of differences between \bar{r} and the phantom image \bar{x} was then computed, followed by a bound on the fraction of pixel differences $U := \frac{s}{n}$, and the actual fraction of differences $E := \frac{e}{n}$, where e is the number of pixel differences between \bar{r} and \bar{x} . The upper bound t from Corollary 1 on the number of differences any two binary solutions of $Wx = p$ was then computed, followed by the computation of the fraction of pixel differences $V := \frac{t}{n}$. Due to space limitations, we only show the results for Phantom 3. The results for the other two phantoms are in line with the observations made for the third phantom. In all experiments, a square window was used. Note that the position of a partial partition $S^{a,b}$ with respect to the high resolution image is completely determined by the pair (a, b) , which we call a *starting point*. Each low resolution image of the high resolution binary image corresponds to a different starting point. In the experiments, we distinguish between regularly and randomly distributed starting points, where the regular case corresponds to a low resolution scanner that is gradually shifted across the high resolution image, and the random case corresponds to a device that moves irregularly (or an object that moves in such a way); see Fig. 3. In Fig. 4, the three error measures V , U and E are plotted as a function of the number of starting points for window size of 8×8 and 32×32 and for both regularly and randomly distributed starting points. Note that for a larger window size, more starting points is required to obtain similar error bounds.

Various observations can be made from the graphs in Fig. 4. Even if the number of starting points is much smaller than the number of pixels in the window, meaning that the reconstruction problem is severely underdetermined, it is still possible to guarantee

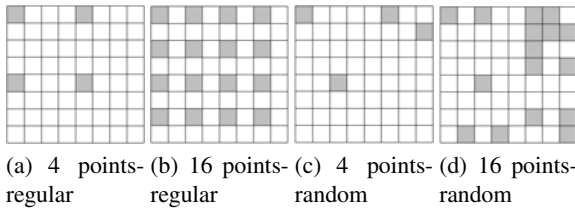


Fig. 3. Distribution of starting points in a first scan-window of size 8×8

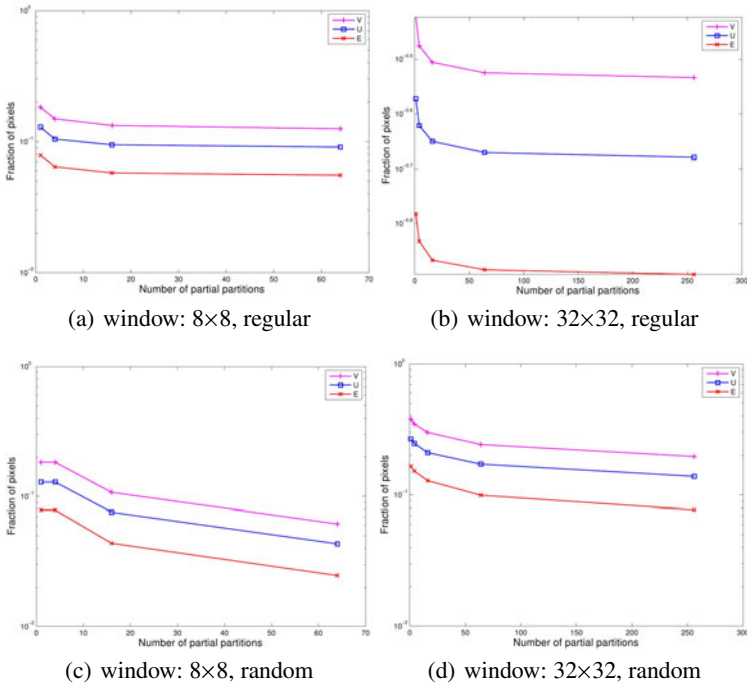


Fig. 4. computed bounds as a function of the number of partial partitions for Phantom 3; V: bound on the distance between any two binary solutions from Cor. 1; U: bound on the distance between any binary solution and the rounded central reconstruction \bar{r} from Thm. 1; E: true error between the rounded central reconstruction and the binary phantom \bar{x}

that only a limited fraction of pixels can be different between binary solutions. Although the given bounds U is clearly not sharp when compared to the real error E , rounding the central reconstruction yields a binary image that is in many cases guaranteed to be rather close to the original image. For example, for window size 8×8 and randomly distributed starting points, having just 16 low resolution images available (resulting in a system of equations that is underdetermined by a factor of 4) can still guarantee that the rounded central reconstruction is within 10% of the original binary image.

Fig. 5 illustrates the key concepts involved in the proposed bounds. The top row shows the central reconstruction for window sizes 8×8 and 32×32 , with regularly and randomly distributed starting points. Here, the number of starting points is chosen as a fixed fraction of $\frac{1}{4}$ times the number of pixels in the window. In this way, all four reconstruction problems can be described by roughly the same number of equations. The middle row shows the difference images between the central reconstruction and the phantom, whereas the bottom row shows the difference images between the rounded central reconstruction and the phantom.

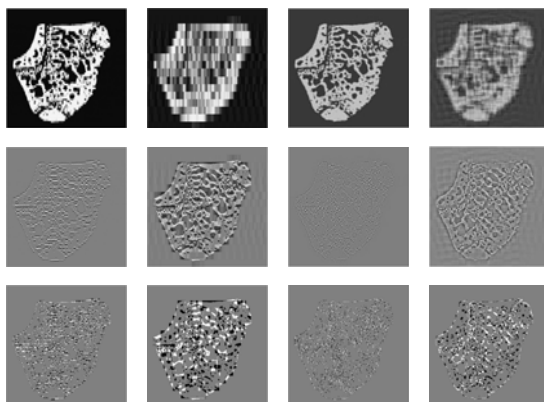


Fig. 5. Illustrations of the key concepts for Phantom 3; **From left to right:** 8×8 window, regular; 32×32 window, regular; 8×8 window, random; 32×32 window, random; **From top to bottom:** central reconstruction; difference between the central reconstruction and the phantom; difference between the rounded central reconstruction and the phantom

5 Outlook and Conclusions

In this article, we have presented general bounds on the accuracy of reconstructions of binary images from several low resolution graylevel scans, with respect to the unknown original image. The bounds can be computed efficiently and give guarantees on the number of pixels that can be different between any two binary reconstructions that satisfy given window-sums, and on the difference between a particular binary image, obtained by rounding the central reconstruction to the nearest binary vector, and any binary image satisfying the window-sums. The experimental results show that by using these bounds, one can prove that the number of differences between binary reconstructions must be small, even when the corresponding real-valued system of equations is severely underdetermined. This work represents an extension of the methodology set up in [1], which is a step towards a set of general bounds for binary image reconstruction problems that allow various forms of image sampling and incorporation of noisy measurements.

References

1. Batenburg, K. J., Fortes, W., Hajdu, L., Tijdeman, R.: Bounds on the difference between reconstructions in binary tomography. In: Proc. of the 16th IAPR International Conference on Discrete Geometry for Computer Imagery, pp. 369–380 (2011)
2. Batenburg, K.J., Sijbers, J.: Generic iterative subset algorithms for discrete tomography. *Discrete Appl. Math.* 157(3), 438–451 (2009)
3. Ben-Israel, A., Greville, T.N.E.: *Generalized inverses: Theory and applications*. Canadian Math. Soc. (2002)
4. Björck, Å.: *Numerical methods for least square problems*. SIAM, Linköping University, Sweden (1996)

5. Frosini, A., Nivat, M.: Binary matrices under the microscope: A tomographical problem. *Theoretical Computer Science* 370, 201–217 (2007)
6. Frosini, A., Nivat, M., Rinaldi, S.: Scanning integer matrices by means of two rectangular windows. *Theoretical Computer Science* 406, 90–96 (2008)
7. Hajdu, L., Tijdeman, R.: Algebraic aspects of discrete tomography. *J. Reine Angew. Math.* 534, 119–128 (2001)
8. Li, L., Stankovic, V., Stankovic, L., Li, L., Cheng, S., Uttamchandani, D.: Single pixel optical imaging using a scanning MEMS mirror. *J. Micromech. Microeng.* 21, 25022 (2011)
9. Saad, Y.: *Iterative methods for sparse linear systems*. SIAM, Philadelphia (2003)
10. Slavinsky, J., Laska, J., Davenport, M., Baraniuk, R.: The compressive multiplexer for multi-channel compressive sensing. In: *Proc. of the International Conference on Acoustics, Speech, and Signal Processing (ICASSP)* (2011)
11. Wakin, M., Laska, J., Duarte, M., Baron, D., Sarvotham, S., Takhar, D., Kelly, K., Baraniuk, R.: An Architecture for Compressive Imaging. In: *Proceedings of the International Conference on Image Processing (ICIP)*, pp. 1273–1276 (2006)