Discrete algebraic reconstruction technique: a new approach for superresolution reconstruction of license plates

Karim Zarei Zefreh
Wim van Aarle
K. Joost Batenburg
Jan Sijbers
Discrete algebraic reconstruction technique: a new approach for superresolution reconstruction of license plates

Karim Zarei Zefreh
Wim van Aarle
University of Antwerp (CDE)
Department of Physics
iMinds–Vision Lab
Universiteitsplein 1
B-2610 Antwerpen, Belgium
E-mail: karim.zarei@gmail.com

K. Joost Batenburg
Centrum Wiskunde & Informatica
Science Park 123
NL-1098XG Amsterdam, The Netherlands

Jan Sijbers
University of Antwerp (CDE)
Department of Physics
iMinds–Vision Lab
Universiteitsplein 1
B-2610 Antwerpen, Belgium

Abstract. A new superresolution algorithm is proposed to reconstruct a high-resolution license plate image from a set of low-resolution camera images. The reconstruction methodology is based on the discrete algebraic reconstruction technique (DART), a recently developed reconstruction method. While DART has already been successfully applied in tomographic imaging, it has not yet been transferred to the field of camera imaging. DART is introduced for camera imaging through a demonstration of how prior knowledge of the colors of the license plate can be directly exploited during the reconstruction of a high-resolution image from a set of low-resolution images. Simulation experiments show that DART can reconstruct images with superior quality compared to conventional reconstruction methods.

1 Introduction
Vehicle license plate detection from surveillance cameras is widely used in traffic monitoring and control systems. Deciphering license plates based on video sequences is challenging. Surveillance cameras have limited spatial resolution, which may not always suffice to resolve the alphanumeric characters from the license plates. Super-resolution methods are often required to reconstruct a high-resolution (HR) image from a set of subpixel-shifted low-resolution (LR) images. Fundamentally, such a task involves dealiasing and deblurring. To improve the readability of the plates, several methods have been suggested in the past. For example, Zhang et al. suggested a method to enhance only the character pixels while deemphasizing the background pixels. In Ref. 3, Li et al. presented a bilinear interpolation scheme to enhance license plates. Cui and Huang described a multi-frame scheme for the extraction and enhancement of alphanumeric characters in license plates. The authors in Ref. 5 proposed a robust maximum a posteriori–based method with discontinuity-adaptive Markov random field (DAMRF) prior for enhancing edges in reconstruction process. A generalized DAMRF model has been also used in Ref. 6 to make license plate numbers more legible. Kang et al. presented an iterative image reconstruction scheme to remove motion blur.

Iterative reconstruction schemes are based on minimizing the difference between the simulated (i.e., computed) LR images and the observed (measured) LR images. The simulated LR images are computed from an imaging model including blur and subsampling. In each iteration, the error (difference) between simulated LR images and observed LR images is redistributed across the current estimate of the HR image. This process is repeated iteratively to minimize the energy of the error.

If the HR image to be reconstructed is composed of a larger number of pixels than the total number of LR image pixels, an infinite number of HR images can be found that adhere to the acquired set of LR images. To select among the large number of possible solutions the HR image that

Paper 13208SS received Apr. 15, 2013; revised manuscript received Jun. 26, 2013; accepted for publication Jul. 17, 2013; published online Aug. 28, 2013.

0091-3286/2013/$25.00 © 2013 SPIE and IS&T [DOI: 10.1117/1.JEI.22.4.041111]
corresponds to the true HR image, prior knowledge about the HR image to be reconstructed is required.

In the field of license plate reconstruction, various types of prior knowledge have been proposed to make the inverse problem well posed. For example, smoothness constraints suggesting that the HR image is (globally or piece-wise) smooth are popular choices. Alternatively, knowledge that the plate’s text consists of alphanumeric characters only may prove to be highly useful as well. In this work, a complementary kind of prior knowledge is incorporated into the reconstruction process, namely that a license plate consists of a limited number of colors (gray levels). To efficiently incorporate such prior knowledge, the concept of discrete algebraic reconstruction and, in particular, a recently proposed algorithm [discrete algebraic reconstruction technique (DART)] will be introduced.

Discrete algebraic reconstruction, as introduced in this paper, stems from discrete tomography (DT), which is concerned with the problem of recovering images from their projections, where the images are assumed to consist of a small number (2 to 5) of gray values only. If the set of gray levels is known (or can be estimated) in advance, this prior knowledge can be exploited by the DT reconstruction algorithm.

Potential benefits of DT are an increase of the reconstruction quality and a reduction of the required number of projection images. The DT reconstruction problem, however, is generally underdetermined and the number of possible solutions can be substantial. To guide the reconstruction process toward an optimal as well as intuitive solution, the DART algorithm was proposed. It allows one to efficiently reconstruct a high-quality image from a limited number of projections. The DART algorithm has been successfully applied to produce accurate reconstructions in several domains such as computed tomography (CT) and electron tomography.

In this paper, DART is introduced in the field of license plate reconstruction to increase the spatial resolution of a license plate image from a set of LR images by exploiting the discreteness of the license plate gray levels in an iterative reconstruction framework. While DART has been applied in CT and electron tomography, to the authors’ knowledge, the concept of DT has not yet been transferred to the domain of reconstructing an HR image from a set of LR camera images. While in tomography the projection values are modeled as line integrals of the unknown object, in camera imaging the LR camera pixel values are modeled by a combination of an orthographic projection, spatial averaging (blurring), and subsampling of the unknown HR image.

After introducing basic notations and concepts in Sec. 2.1, an overview of the DART algorithm is given in Sec. 2.2, which will be focused on the reconstruction of an HR image from a set of LR images. In Sec. 3, results are presented and discussed for simulated as well as experimental datasets. Finally, in Sec. 4, conclusions are drawn.

2 Method

2.1 Imaging Model

Let \( \{ y_i \}_{i=1}^d \) represent a set of \( d \) LR images of size \( M \times N \). It is assumed that these images are acquired under orthographic projections and that individual scene motions can be modeled as affine transformations. The HR image that we want to reconstruct from \( \{ y_i \} \) is represented by \( x \). We model each LR image as a noisy, uniformly downsampled version of the HR image, which has been shifted and blurred. If \( D \) denotes the downsampling operator, \( G \) the blurring operator, and \( A \) the affine transform that maps the HR grid coordinate system to the LR grid system, we have

\[
DGAx + n = y. \tag{1}
\]

This can be rewritten to

\[
Wx + n = y, \tag{2}
\]

where \( W = DGA \) is the complete system matrix. The reconstruction \( x \) is computed on a rectangular pixel array of width \( w \) and height \( h \). Hence, the total number of pixels in the reconstruction is given by \( n = wh \). For each LR image, we assume that the number of pixels is \( l \). The total number of available LR pixels is denoted by \( m = ld \). The entries of the \( n \times 1 \) column vector \( x \) correspond to the pixel values of the reconstruction. The \( m \times 1 \) column vector \( y \) contains the LR image pixels, ordered column-wise.

Finally, the \( m \times n \) system matrix \( W \) defines the transformation from \( x \) to \( y \).

Ignoring noise, the reconstruction problem can then be formulated as a system of linear equations (see, e.g., Chapter 7 of Ref. 14).

\[
Wx = y. \tag{3}
\]

An (approximate) solution of Eq. (3) can be found using an iterative algebraic reconstruction method such as ART, simultaneous ART, or simultaneous iterative reconstruction technique (SIRT). The DART algorithm, which will be discussed in the next section, can be used in conjunction with each of these algorithms. In the context of this paper, we refer to algebraic reconstruction method (ARM) as a specific iterative reconstruction algorithm. In our experiments, we used SIRT as the ARM. SIRT computes a solution for which the norm of the difference \( \| Wx - y \| \) between the computed set of LR pixels and the measured data is minimal with respect to a certain vector norm, i.e., a least-squares solution. In each ARM iteration, all LR images are enumerated in random order, each time updating the current reconstruction.

2.2 Discrete Algebraic Reconstruction Technique

In this section, we will give a brief overview of DART applied to the reconstruction of an HR, discrete gray-level image from a small number of LR images.

2.2.1 Principle of DART

DART was introduced in the field of CT, where an image needs to be reconstructed from a set of x-ray projection images. DART tries to find a solution of Eq. (3) under the constraint that each of the \( x_i \) can only take values in a prescribed set. First, a conventional reconstruction is computed. Then, a number of DART iterations are performed, each containing the following steps:

1. The current reconstruction is subdivided into partitions by thresholding and each of the partitions is assigned the known gray-level value.
2. The difference between the actual projection data and the forward projection of the segmentation is computed.


3. Pixels that are on the boundary of a partition are identified.

4. A regular iterative technique such as SIRT is applied to reconstruct the projection difference in the pixels on the boundaries. The other pixels are kept fixed at the gray level of their partition.

More formally, in each DART iteration, nonboundary pixels are assigned a given gray level in the prior known set of gray levels. Let $B$ be the boundary of the object in the thresholded image, which is defined as the set of all pixels that are adjacent to at least one pixel having a different gray level. Let $I$ denote the remaining set of pixels. All pixels in $I$ are assigned their thresholded value. The idea of DART is now to exclude the pixels in $I$ from the regular ARM iterations and, hence, only update the boundary pixels. That is, several ARM iterations are performed on the pixels in $B$ only. In this way, the number of variables in the linear equation system [Eq. (3)] is significantly reduced, while the number of equations remains the same. In each iteration, the boundary pixels are allowed to vary independently, which may result in large local variations of the pixel values. To regularize the reconstruction algorithm, the boundary pixels are locally smoothed after applying an ARM iteration on $B$. Subsequently, the resulting image is again thresholded, a new set of boundary pixels $B$ is determined, and the next ARM iteration is run on the pixels in $B$. These steps are iterated until a certain convergence criterion is met.

Figure 1 shows an overview of the DART algorithm. Note that DART is an integrated approach in which algebraic reconstruction iterations are interwoven with segmentation steps. Hence, DART is not just an extension of SIRT but an integrated method that uses an arbitrary algebraic reconstruction method along with an efficient way to impose the discreteness constraint of the gray levels.

### 2.2.2 DART for license plate reconstruction

In this subsection, DART is explained for the reconstruction of an HR image from a set of LR images. Suppose that we want to reconstruct the binary image shown in Fig. 2(a) from a small number of LR images. Figure 2(b) shows the ARM reconstruction after 10 iterations. From the ARM reconstruction, it is difficult to decide where the edges of the object are exactly. Moreover, the small letters at the bottom are totally unreadable. Yet, the thresholded reconstruction in Fig. 2(c) shows that pixels that are not close to the boundary of the original HR image are assigned the correct gray level in the thresholded image. The observation that generally only the boundaries in the image are incorrectly classified is the core of DART, an iterative reconstruction technique that combines conventional iterative reconstruction techniques and segmentation into a single algorithm. While an ARM redistributes the error between the simulated and the observed LR images over the complete image, DART redistributes this error only in the border regions.

Figure 2(d) shows the set of boundary pixels. All remaining pixels are assigned their thresholded value, in this example, either black or white. Then, several ARM iterations are performed on the boundary pixels only. Figure 2(e) shows the relative change of the image pixels after one ARM iteration of the boundary pixels, where gray denotes no change. The result of the local smoothing of the boundary pixels is shown in Fig. 2(f). Subsequently, the resulting image is again thresholded, a new set of boundary pixels is determined, and the next ARM iteration is run on the boundary pixels.

### 2.3 Motion Estimation Within DART

To reconstruct an HR image from a set of LR images, knowledge of the transformation $W$ between the HR image and each of the LR images is crucial. To estimate the shift with subpixel accuracy, phase correlation is often employed, which uses the fast Fourier transform of the shifted LR images to obtain a measure of correlation.\(^{15-17}\) In our work, a more general approach is employed to estimate the transformation parameters by integrating the estimation of the transformation between the HR image and the LR images within the proposed iterative reconstruction scheme (shown in Fig. 3). The motion estimation procedure is as follows:

1. Initialize the transformation parameters (an initial estimate for the shift can be obtained using the cross-power spectrum as in Ref. 5).
2. Reconstruct an HR image from the set of measured LR images using the estimated transformation parameters.
3. Simulate the LR images from the reconstructed HR image by geometrically transforming the HR image to the grid of each of the LR images and subsequently blurring and downsampling the transformed images.

4. Compute the mean squared error (MSE) between the acquired LR images and the simulated LR images.

5. Find new estimates of the transformation parameters by minimizing the MSE between the acquired LR images and the simulated LR images as a function of the transformation parameters.

6. Go to step 2 until a certain convergence criterion is reached.

3 Experiments and Results

Simulation experiments were set up to test the performance of the proposed license plate reconstruction methodology. For the experiments, two HR images were created, one binary HR image and one color HR image, both of size $256 \times 1024$, shown in Fig. 4(a) and 4(b), respectively. Furthermore, an HR real license plate picture was captured with a simple digital camera [Fig. 4(c)]. The HR real image was of size $120 \times 570$ pixels captured by a digital camera (Canon PowerShot A480). One pixel of the HR image corresponded to a physical dimension of 0.90 mm. From each HR image, a set of LR images was created by scaling, shifting, blurring, and, subsequently, downsampling the HR image. The blurring involved a convolution of a Gaussian kernel of size $5 \times 5$ with a standard deviation equal to two times the HR pixel width. The weights of the kernel were computed by the value of the Gaussian function in the center of each pixel in the kernel. Next, from the set of simulated LR images, an HR image was reconstructed using some conventional reconstruction methods and DART, our proposed discrete algebraic reconstruction method.

For each experiment described below, the reconstruction quality was calculated in terms of MSE and peak signal-to-noise ratio (PSNR). MSE was computed as the sum of squared differences between the original HR image and the reconstructed HR image, divided by the total number of HR pixels.

3.1 Comparative Experiment: Ideal Conditions

First, the performance of DART was compared to that of bicubic interpolation, iterated back-projection (IBP), projections onto convex sets (POCS), and robust super-resolution under ideal conditions, that is, no noise, change of lighting, etc. and the exact knowledge of the point spread function and the subpixel translations exist. For this comparative experiment, we used eight LR images of size $16 \times 66$ pixels, generated from Fig. 4(a). The HR image shown in Fig. 4(a) was used as the ground truth image. The results of the employed reconstruction methods are presented in Fig. 5. For DART, a computationally optimized implementation of SIRT algorithm was used as a subroutine ARM and its result after 20 iterations was used as an initial estimate. For each DART iteration, the well-known Otsu thresholding was used to subdivide the current reconstruction into partitions and each of the partitions were assigned the known gray-level value (e.g., for this experiment, gray levels are 0 and 1). To compare the results of other methods that yield gray-level images with the results of other methods that yield gray-level images with the results of other methods that yield gray-level images with the results of other methods that yield gray-level images with the results of other methods that yield gray-level images with the results of other methods that yield gray-level images with the results of other methods that yield gray-level images with the results of other methods that yield gray-level images with the results of other methods that yield gray-level images with the results of other methods that yield gray-level images with the results of other methods that yield gray-level images with the results of other methods that yield gray-level images with the results of other methods that yield gray-level images with the results.
of DART, their reconstructed images also were segmented using the Otsu segmentation, yielding the required discrete set of gray levels. Figure 5(a) shows one of the LR images. Figure 5(b) is the result after bicubic interpolation (by averaging the interpolated LR images). Next, the reconstructed image using the POCS is shown in Fig. 5(c). Figure 5(d) shows the result of IBP. Finally, Fig. 5(e) and 5(f) shows the result of robust super resolution (SR) and DART, respectively. The examined SR algorithms converge rapidly (usually within <10 iterations) and are very stable, so in all iterative methods 10 iterations were employed. When visually comparing the reconstructed license plates, a clear difference in reconstruction quality can be noticed, especially when focusing on the boundary pixels.

Apart from visual comparison, MSE and PSNR were computed for all approaches. To further validate our results, the experiment was repeated five times, and the average MSE and PSNR are shown in Table 1.

### 3.2 Experimental Conditions

In this section, the performance of DART as a function of (simulated) experimental conditions will be discussed, such as the number of LR input images, the pixel size of the LR images, and the robustness against noise. Since robust SR was observed to outperform bicubic interpolation as well as the POCS and IBP methods, in the remaining experiments, DART was compared to robust SR only. Similar to the previous section, each experiment was repeated five times, and the average MSE was used to compare the methods.

#### 3.2.1 Number of LR images

Each LR image generally provides new information that can be used in the reconstruction. It is intuitively clear that the quality of the HR image reconstruction should improve with the number of LR images. Figure 6 shows MSE as a function of the number of LR images used in the reconstruction. For this experiment, we used a set of LR images of size $57 \times 34$ pixels, generated from Fig. 4(c). The LR images were randomly shifted and the shift was assumed to be known. Figure 6 shows that the optimal number of LR images for DART to reconstruct an HR image with minimum MSE is eight images. The figure shows that even with a smaller number of LR images as low as three, a significant improvement in reconstruction quality can be obtained with DART. We remark that the contribution of each LR image to the reconstruction quality depends on its shift with respect to the grids of the other LR images. It is clear that when two LR images are shifted by the same amount, or if the difference in their shift values is a whole number times the LR pixel size, no additional information is added. However, since, in practice, the time sampling of the LR images is independent of the spatial sampling, this is unlikely to occur.

#### 3.2.2 Subsampling factor of the LR images

The subsampling factor of the LR images is naturally of importance with respect to the quality of the reconstructed HR image. In this experiment, MSE of DART and robust SR as a function of the subsampling factor used to generate the LR images was considered. To this end, 10 sets of LR images were used in which each set contained four LR images with the same pixel size. To illustrate, the subsampling with a factor 10 for an LR image means that the size of the LR pixel was 10 times larger than the HR pixel of the original image in both dimensions. All LR images were generated from Fig. 4(c) based on the explained imaging model. Not surprisingly, it is clear from Fig. 7 that for both the robust and DART reconstruction, MSE increases with increasing subsampling factor. However, it is also clear that MSE for robust SR is significantly larger than that of DART, especially for large subsampling factors.

Figure 8 shows a visual comparison between robust SR and DART with a subsampling factor of 10, in which the LR images are very small [e.g., $7 \times 34$ pixel in Fig. 8(a)] and alphanumeric characters are hardly readable. Applying DART, a completely readable HR image can be appreciated, whereas with the robust SR method, some characters are not recognizable.

#### 3.2.3 Noise level

The license plate reconstruction performance for both robust SR and DART was tested as a function of the noise level of the LR images. To that end, eight LR images of size $57 \times 34$ pixels were generated from Fig. 4(c) and then polluted with Gaussian distributed noise with increasing variance. The results shown in Fig. 9 show that, independent of the applied noise level, the performance of DART was significantly better with respect to MSE, compared to robust SR. Moreover, it shows that DART is more robust against noise than robust SR method.
3.2.4 Motion estimation

In practice, the transformation between the HR image and the LR images is unknown and needs to be estimated. As explained in Sec. 2.3, the transformation parameters can be estimated within the reconstruction scheme. Conventionally, shifts are estimated based on phase correlation.

For this experiment a set of eight LR images with random shifts was generated. Each LR image was of size $12 \times 57$ [Fig. 10(d)]. MSE was computed from a DART reconstruction in which the shifts were known [Fig. 10(a)], from a DART reconstruction in which the shifts were estimated using cross-correlation [Fig. 10(b)], and from a DART reconstruction in which the shifts were estimated using the proposed iterative method [Fig. 10(c)]. The number of DART iterations was 10 for all experiments.

3.2.5 Color license plates

DART can easily be applied to the reconstruction of color license plates as well. Indeed, each band represents a gray-level image with only a very small (three or four) gray levels. Hence, from the sets of LR images for each band, an HR image band can be reconstructed. In our experiments, the gray levels of the bands were assumed to be known. However, they can also be estimated in a semi-automatic way as described in Ref. 24. After the reconstruction of the separate bands, the HR image is composed. An example is shown in Fig. 11, where four LR images were used with a downsampling factor 8 in both directions. In DART, 25 iterations were used. Note that if, instead of each band separately, the brightness of the RGB color space is known to be composed of only a few intensity levels, this prior knowledge can as well be exploited in a DART-based reconstruction.

4 Conclusions

We have presented a new application of DART, a recently proposed iterative algebraic reconstruction algorithm, to reconstruct an HR license plate image from a series of LR
images. The DART algorithm combines the efficiency of iterative algebraic methods from continuous tomography with the power of discrete algebraic reconstruction technique to compute an accurate HR image from a small number of LR images.

We compared the DART approach to conventional super-resolution techniques: bicubic, POCS, IBP, robust SR. The experiments demonstrated the superiority of DART compared to conventional SR methods. The DART algorithm is capable of reconstructing a high-quality license plate image from a small number of LR license plate images and can be applied to reconstruct binary, gray-level, as well as color license plates.

Acknowledgments
This work was partially financially supported by the Strategic Basic Research-project QUANTIVIAM (060819) of the Institute for the Promotion of Innovation through Science and Technology in Flanders (IWV Vlaanderen).

References

Wim van Aarle received his master’s degree in computer sciences at the University of Antwerp, Antwerp, Belgium, in 2007. In 2012, he obtained his PhD degree in science from the same university. He is currently a postdoctoral researcher at the Visionlab, University of Antwerp, and at the Intel ExaScience Lab at the IMEC Institute, Leuven, Belgium. His research interests include high-performance computing and discrete tomography in a wide spectrum of applications.

K. Joost Batenburg graduated in mathematics (2002) and computer science (2003) from the University of Leiden, The Netherlands. From 2002 to 2006 he was a PhD student at the University of Leiden and CWI in Amsterdam. He obtained his PhD in mathematics in 2006 with the thesis “Network flow algorithms for discrete tomography.” He then moved to the Vision Lab at the University of Antwerp, where he worked as an Fonds Wetenschappelijk Onderzoek (FWO) postdoctoral fellow (2006 to 2010) and from March 2010 as a part-time professor. With Jan Sijbers, he leads the ASTRA group, which focuses on developing new algorithms and techniques for tomographic reconstruction at all scales. Since March 1, 2010, he has been a scientific staff member at the Centrum Wiskunde & Informatica in Amsterdam, where he performs research in the field of computational imaging.