Superresolution of License Plates using Discrete Tomography

Alejandro Rodríguez

Academic year 2010-2011

Promotor: Prof. Dr. Jan Sijbers

Master thesis in computer science
Abstract

DART is a novel super-resolution algorithm that aims in reconstructing a high resolution image from a set of projections or low resolution images. One of the applications of DART is the reconstruction of high resolution images of license plates.

This thesis, studies the performance of DART algorithm in the field of license plate reconstruction. In order to evaluate the robustness of the algorithm, we will perform series of experiments consisting on simulations of a more realistic environment. The results are discussed and compared with commonly used reconstruction methods.

Finally, a graphical interface, developed in Matlab, is presented to support parameterized experiments.
Word of thanks

While writing this bachelor thesis I have enjoyed the support and help of several people. I would like here to take the time to thank these people.

I would like to thank my supervisor Dr. Jan Sijbers. Without his advice, support and encouragement I would have never made my contribution to the field of super-resolution images based on discrete tomography.

I thank my parents for giving me the opportunity to study and for their unconditional support that they have given me over the years.

Finally, I would like to thank to a special person for me, Chryso Georaraki who even at the distance has supported me in developing this thesis.
Index

1. Introduction ........................................................................................................... 5
2. Method ..................................................................................................................... 7  
   2.1 Imaging model..................................................................................................... 7  
   2.2 Downsampling..................................................................................................... 8  
3. DART....................................................................................................................... 9  
4. Simulations and experiments .................................................................................. 11  
   4.1 Reconstruction parameters .................................................................................. 11  
      4.1.1 Number of LR images .................................................................................. 11  
      4.1.2 Number of iterations ................................................................................... 12  
      4.1.3 LR pixel size ................................................................................................. 13  
      4.1.4 Shifting ........................................................................................................ 13  
   4.2 Color ................................................................................................................... 15  
   4.3 Noise .................................................................................................................. 20  
   4.4 Gaussian Blur ..................................................................................................... 25  
5. Interface ................................................................................................................ 28  
6. Conclusions ........................................................................................................... 30  
7. References ............................................................................................................. 31
1. Introduction

Traffic monitoring systems are widely used nowadays, with the aim of detect and recognize vehicle license plates. Commonly, surveillance cameras have limited spatial resolution. For this reason, the obtained frames may present unrecognizable alpha numeric characters.

Several super-resolutions methods were presented in the past, in order to reconstruct a high resolution (HR) image from a set of subpixel-shifted low resolution (LR) images. For example, Zhang et al. suggested a method to enhance only the character pixels while deemphasizing the background pixels [2]. In [3], Li et al. presented a bilinear interpolation scheme to enhance license plates. Cui and Huang [4] described a multiframe scheme for the extraction and enhancement of alpha-numeric characters in license plates. Kang et al. [5] used an iterative image reconstruction scheme to remove motion blur.

Recently, a new reconstruction algorithm for discrete tomography, called DART (Discrete Algebraic Reconstruction Technique) was proposed [9]. The reconstruction methodology is based on discrete tomography, in which prior knowledge on the colors of the license plate is exploited. DART can be applied if the image is known to consist of only a few different compositions, each corresponding to a constant grey value in the reconstruction. Simulation experiments show superior reconstruction quality compared to conventional reconstruction methods.

Originally, the problem of increasing the spatial resolution on license plate images was focused on the reconstruction of (typically binary) images for which the domain was a discrete set. DART deals not only with binary images, but also with images that contain three or more grey levels. There is no fixed upper bound on the number of grey levels.
In previous work [10], simulation experiments demonstrated that the DART algorithm is capable of computing reconstructions of high quality from a small number of low resolution images.

In this work, we will study the performance of DART in different simulated scenarios. After introducing basic notations and concepts in sections 2 and 3, in section 4 we will study the robustness of the algorithm as a function of different reconstruction parameters. Besides, in section 5 we will evaluate DART response in a more real environment, introducing new factors, such as reconstruction of color images, presence of noise, and blurring. Finally, in section 6, a graphical interface is presented, which was developed in Matlab, to support parameterized simulation experiments in a comfortable way.
2. Method

2.1 Imaging model

DART is an algebraic reconstruction algorithm, where the reconstruction problem is represented by a system of linear equations. This section explains the basic equation to compute a two-dimensional HR image for a set of LR images. In the following sections, we will introduce new factors in the equation system, such as noise and blur effect.

Let \( \{y_i\}_{i=1,...,d} \) represent a set of low resolution (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 high resolution (HR) image that we want to reconstruct from \( \{y_i\} \) is represented by \( x \). We model each LR image as a uniformly down sampled version of the HR image, which has been shifted. If \( D \) denotes the down sampling operator and \( A \) the affine transform that maps the HR grid coordinate system to the LR grid system, we have:

\[
DAx = y
\]  

(1)

This can be simplified to:

\[
Wx = y
\]

where \( W = DA \) is the complete system matrix.

Here \( x \) and \( y \) are column vectors. The dimensions of \( W \) and \( y \) depend on the down sampling factor, or so to what extent the image resolution will reduced.

Eq. 1 can be solved using an iterative algebraic reconstruction method such as ART, SART, or SIRT [8].
2.2 Downsampling

In order to gain deeper understanding on the downsampling operator, in this section, a small-scale example is explained. Assuming we have a 4x4 matrix $X$ as HR image, consisting of only white and black values (respectively 1 and 0). We want to down sample the image with a factor of 2. It means for each LR pixel corresponds 4 HR pixels. $W$ is the down sampling matrix, which has many rows as $Y$ pixels, and many columns as $X$ pixels. In that case it is a 4x16 matrix. In the construction of the system, $X$ and $Y$ are expressed as column vectors. The figure below shows the complete matrix system for a given HR image $X$.

\[
X = \begin{pmatrix}
0 & 1 & 1 & 1 \\
0 & 1 & 1 & 1 \\
1 & 0 & 0 & 0 \\
1 & 1 & 0 & 0
\end{pmatrix}
\]

Converting $X$ into a column vector and multiplying for the down sampling matrix we obtain a down sampled image $Y$. Transforming $Y$ in to a 2x2 matrix we have a LR image down sampled with a factor of 2.

\[
X = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix}
\]

\[
W = \begin{pmatrix}
1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0
\end{pmatrix}
\]

\[
Y = \begin{pmatrix} 2 \\ 0 \\ 0 \\ 0 \end{pmatrix}
\]
3. DART

As mentioned before, the reconstruction methodology of DART is based on discrete tomography (DT). This technique is concerned with the tomographic reconstruction of images that consist of only a small number of gray levels. DT reconstruction problems are usually underdetermined. To guide the reconstruction algorithm towards an optimal as well as intuitive solution is necessary heuristic rules. DART was presented as a heuristic DT algorithm that is based on an iterative algebraic reconstruction method (ARM). Starting from a continuous reconstruction, a discrete image is reconstructed by consistent updating of border pixels. Fig. 1 shows a flow-chart of DART. A continuous reconstruction is computed as a starting point, using the ARM. Subsequently, a number of DART iterations are performed.

![Diagram of DART algorithm](image)

Figure 1: Flow chart of the DART algorithm

In the particular case of license plate reconstruction, suppose that we want to reconstruct the binary image shown in Figure 2(a) from only 9 LR images. Figure 2(b) shows the ARM reconstruction after ten 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 Figure 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. 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 (cfr. Figure 2(d)). We now move back to the original gray level ARM reconstruction. All pixels that are not in B are assigned their thresholded value, either black or white. Next, several ARM iterations are performed for the pixels in B only. In this way, we significantly reduce the number of variables in the linear equation system (1), while the number of equations remains the same. Figure 2(e) shows the relative change of the image pixels after one ARM iteration of the boundary pixels, where gray denotes no change. In the ARM step, each of the boundary pixels is 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 ARM. Figure 2(f) shows the result of this filtering operation. Subsequently, the resulting image is again thresholded and each of the steps that we just described is repeated iteratively.

(a) original image     (b) ARM rec.     (c) thresholded rec.

(d) boundary pixels   (e) DART update   (f) final DART rec.

Figure 2: Reconstructing a phantom. The images indicate the various steps of the DART algorithm.
4. Simulations and experiments

In this section, we present the results of a series of experiments, comparing the reconstructions computed by DART with alternative approaches. First we will evaluate the reconstructed image quality as a function of different reconstruction parameters, such as, the number of iterations, the number of LR input images, the pixel size of the LR images, and the amount of shifting. In a second step, we will observe the response of DART in different simulated scenarios, such as, reconstruction of colored license plates, presence of noise and blur. For each experiment, we compare the result of DART with the continuous algorithm SIRT [8]. In order to determine the quality of the reconstruction, we compute the relative error, being the sum of absolute differences between the original LR image and the simulated LR image, divided by the total number of LR pixels.

4.1 Reconstruction parameters

In the first experiments, the influence of certain reconstruction parameters on the DART reconstruction quality is discussed, such as the number of iterations, the number of LR images, the LR pixel size and the shifting.

4.1.1 Number of LR images

One of the principles of discrete tomography is, in order to obtain a quality HR image; each shifted LR image must carry different information. We can assure that the quality of the HR is directly proportional to the number of LR images. Figure 3 shows the relative error, for DART and SIRT, as a function of the number of LR images used in the reconstruction. For this experiment, 40 iterations were executed in the case of DART, and the pixel size of LR images is 8, that is, each LR pixel corresponds to 64 HR pixels.
In this experiment, the DART image reconstruction quality, in terms of the relative error, was studied as a function of the number of iterations used. From Fig. 4, it can be observed that the relative error monotonically decreases with the number of iterations. For this experiment, a set of 10 LR images was used. The LR images were linearly shifted and the pixel size was 8. In addition, in Fig. 4, it can be observed that for a sufficient number of LR images, a small number of iterations are necessary to obtain a result with relative error below 2%.
4.1.3 LR pixel size

As mentioned before, the pixel size indicates how many HR pixels correspond to a single LR pixel. For example, a pixel size of 4 means that, 1 LR pixel corresponds to 4x4 HR pixels. It is clear that for higher values of pixel size the reconstruction quality decreases. However, for lower values the computational cost increases, because the size of the matrixes involved is directly related to the size of the LR images. Figure 5 shows the influence of the LR pixel size on the reconstruction quality in terms of relative error. We can observe that for a LR pixel size of 20, the relative error is still below 5%.

![Figure 5: Relative error in function of the LR pixel size](image_url)

4.1.4 Shifting

Tomography is based on recovering images from their projections [6]. In order to obtain a quality HR reconstruction, each projection must carry different information. In the case of license plates the set of LR images must be shifted with different values at least in one direction; vertical or horizontal. The amount of shifting can be expressed in terms of LR or HR pixels. With higher values of shifting, the LR image can get out of bounds leading to information loss. An example is shown in Figure 6.
In a real environment, we can suppose that the monitoring system can capture the complete license plate for each frame. In order to simulate LR images that provide extra information for each iteration we can measure the shifting in terms of HR pixels. If one LR pixel corresponds to $n$ HR pixels we can express shifting in terms of $xn + y$, where $x$ is the number of LR pixels and $y$ the number of HR pixels. For example, a shifting of 43 HR pixels with a blocksize of 4 can be expressed as $10 \times 4 + 3$. Now suppose we have two low resolution images, down sampled with a factor of 4, the first one has a horizontal shifting of 3 HR pixels and the second one 19 HR pixels (that is, 4LR pixels and 3HR pixels). In that case the second one does not provide extra information to the reconstruction. Hereby we can limit the maximum shifting distance to one LR pixel in all directions. Nevertheless, we added to the interface the functionality of change the amount of shifting measured in terms of LR pixels.
4.2 Color

In previous experiments for reconstruction of license plates, we used binary images, with only two gray levels, 0 for the black pixels and 1 for the whites. DART offered high quality results with binary images in comparison with other reconstruction algorithms. In this section, color is introduced in the reconstruction. Now the input matrix corresponds to a set of LR RGB colored images. The strategy is to process each band separately and then compose a HR color image. Notice that now gray levels and thresholds are different for each band and must be calculated separately. In an image with multiple colors and details, the decision on which are the gray levels for each band is not a trivial problem. The proposed solution is to take as gray levels those which present a number of pixels above a certain threshold of the histogram curve. Other prior knowledge can be included, like the spreading of the gray level in the image or certain shapes.

The quality of the reconstruction of colored license plates depends on the number of colors and the level of detail. In the first experiment, we will evaluate the result obtained for the Belgian license plate, which presents a medium level of detail and an amount of three or four main gray levels per band. Fig. 7 shows the set of downsampled LR images, the reconstruction for each band and the final composition, comparing the result with the SIRT reconstruction.
(a) Set of LR images.

Red band: 3 GL

Green band: 4 GL

Blue band: 4 GL

(a) Reconstructed images for each band using DART.
DART presents a very good reconstruction quality and it can be observed that colored areas are homogenous and do not contain irregularities as in the SIRT reconstruction. That can be observed clearly in the blue area. In addition, DART reconstruction presents the same color tones as the original HR image, whereas the irregularities caused by SIRT reconstruction, lead in variations in the color tones.

It has been observed that DART presents better results with lower values of gray levels, for this reason the LR matrix for each band is scaled to values between 0 and 1. After reconstructing each band, the result is re-scaled to RGB levels, between 0..255, and then the HR color license plate is composed.
In the next experiment the license plate selected has more color and detail. Now the number of gray levels involved is higher. The choice of these gray levels will determine the quality of the reconstruction. For example, by selecting a high number of gray levels, we gain in quality in specific details, but lose quality in the main areas of the reconstructed image. Fig. 8 shows the set of LR images and the DART and SIRT reconstruction with a maximum of 8 gray levels per band.

(a) Set of 10 LR images.

(b) Original HR image.
(c) Reconstructed color image using DART (relative error 1.21%).

(d) Reconstructed color image using SIRT (relative error 4.77%).

Figure 8: Reconstructed images with 10 LR, 25 iterations and pixel size of 8.

Once again, DART offers a quality reconstruction, with a relative error below 1.5. The result presents uniform background colors and well-defined contours. In the case of SIRT the reconstruction is quite good but, we can observe some deficiencies in the background of blue and white areas.

Figure 9 shows the relative error as a function of the threshold value. The license plate used is the same as in the previous experiment. Lower threshold implies more gray levels. We can observe that the optimal value for this image is a threshold of 2000, the amount is 6 or 7 gray levels per band.
4.3 Noise

In this experiment the behavior of DART in the presence of background noise is observed. In order to apply noise in the down sampled LR images, we generate a random noise background and then we add this noise to the LR images. The type of noise selected for this experiment is Gaussian. That is statistical noise that has its probability density function equal to that of the normal distribution, which is also known as the Gaussian distribution. With this type of noise we have two parameters; the variance and the mean. With the variance we can vary the width of the distribution. For this experiment is used a zero-mean distribution, because in that way it is not necessary to scale the LR images in order to maintain the gray values and the thresholds. The matrix system with the addition of noise is:

\[ \mathbf{DA}x + n = y \]

Where D denotes the down sampling operator, \( n \) the Gaussian noise, and A the affine transform that maps the HR grid coordinate system to the LR grid system.
For the first experiment with noise, a binary license plate image is used. With only 2 gray levels, DART presents satisfactory reconstruction quality with noise, with variance values between 0 and 1 the result practically eliminates the effect of noise. Fig. 10 shows the result obtained from DART and SIRT reconstruction for a set of 10 LR images with a noise background with variance of 1.7.

(a) Reconstructed color image using SIRT.

(b) Reconstructed color image using DART.

Figure 10: Reconstructed images with 10 LR, 40 iterations and pixel size of 8.

DART reconstruction almost eliminates the effect of noise and the small characters can be read clearly, for example the phone number. In the case of SIRT the presence of noise clearly remains in the reconstruction.

Note: It has been observed that in the case of noise treatment, the reconstruction quality presents better results with lower values of the DART parameter fixed fraction.

In the next experiment, colored license plates are used, in order to evaluate the response of DART with the presence of noise. In this case the noise is added separately in each band. In this case, the impact of the noise in areas with low values of gray level can lead to irregularities on the reconstruction.
Fig. 11 shows the reconstruction of a color license plate for a set of 10 LR images with a noise background with variance of 1.8.

(a) Reconstructed color image using DART (relative error 3.18%).

(b) Reconstructed color image using SIRT (relative error 6.10%).

Figure 11: Reconstructed images with 10 noised LR, 25 iterations and pixel size of 8.

In this case the amount of noise is higher than in the previous experiment and we can observe that some noise remains in the DART reconstruction. Nevertheless, the reconstruction quality is higher than the result generated by SIRT.

In the case of color images, we have several values of gray levels which could be a problem in the treatment of noise. For example, if the addition of noise in a band produces groups of dark pixels in a clear area, the algorithm could interpret these pixels like part of other adjacent area with a gray level similar to these pixels. In Fig. 12, a sample of the reconstructed red band is shown, where this problem can be observed in the border.
Figure 12: Sample of reconstructed red band.

Figure 13 shows part of the DART and SIRT reconstruction for a high noised LR images. In this case the variance is 3.

(a) DART (relative error 3.35%)   (b) SIRT (relative error 6.18%).

Figure 13: Reconstructed images with 10 noised LR, 25 iterations and pixel size of 8.

In this case the impact of noise is strong and it can be easily observed that the presence of noise remains in both reconstructions. Despite this, DART offers clearly better quality than SIRT.
Figure 14 shows the relative error of DART and SIRT as a function of noise variance. We can observe that DART still manages to keep the relative error below 5% for higher levels of noise. For this experiment, 10 LR images were used.

Figure 14: Relative error obtained with DART and SIRT as a function of noise variance.
4.4 Gaussian Blur

The Gaussian blur is a type of image-blurring filters that uses a Gaussian function (which also expresses the normal distribution in statistics) for calculating the transformation to apply to each pixel in the image. The equation of a Gaussian function in two dimensions, it is the product of two such Gaussians, one in each dimension:

\[ G(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \]

Where \( x \) is the distance from the origin in the horizontal axis, \( y \) is the distance from the origin in the vertical axis, and \( \sigma \) is the standard deviation of the Gaussian distribution. Figure x shows an example of a HR license plate image with presence of Gaussian blur.

To introduce blurring in the matrix system it is needed to apply a new transformation in order to satisfy the following equation:

\[ DGAx + n = y \]

Where \( D \) denotes the down sampling operator, \( G \) the blurring operator, \( n \) the Gaussian noise, and \( A \) the affine transform that maps the HR grid coordinate system to the LR grid system.

To include the blurring operator in our matrix system we applied a sliding window over the down sampling operator in order to simulate Gaussian blur. The size of the sliding window corresponds to the size of the LR pixel with a margins, which corresponds to the parameters \( x \) and \( y \) of the blurring operator. These margins determine the effect of the neighboring HR pixels in a single LR pixel. To obtain the weight of each HR pixel in the sliding window it is added all the occurrences of that pixel as a percentage.
Next figure shows the corresponding sliding window for a down sampling operator of 4 (for each LR pixel correspond 16 HR pixels) with a Gaussian Blurring of 3x3 pixels and a standard derivation of 10.

<table>
<thead>
<tr>
<th>Sliding Window</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gaussian Filter</td>
</tr>
<tr>
<td>0.1113</td>
</tr>
<tr>
<td>0.2220</td>
</tr>
<tr>
<td>0.3328</td>
</tr>
<tr>
<td>0.3328</td>
</tr>
<tr>
<td>0.2220</td>
</tr>
<tr>
<td>0.1113</td>
</tr>
</tbody>
</table>

The cells in gray in the sliding window are the pixels involved in a single block according to the down sampling factor. In this case there are the 16 HR pixels corresponding to 1 LR pixel, the rest of pixels apply the blurring in this LR pixel. Notice that with a blur filter of 3x3 the number of pixels involved to compute each block is more than double. Working with large images and large blurring filters it becomes a problem because the number of indexes of the reconstruction matrix increases exponentially leading memory problems.

In Figure 15, a reconstruction is shown generated with DART and SIRT for a set of 10 LR images with a down sampling factor of 8 and with a blur operator of 7x7 and 10 of σ. Clear difference in reconstruction quality can be noticed. In DART reconstruction blurring almost completely disappears, whereas in SIRT reconstruction is clearly appreciable that blurring remains.
DART presents better reconstruction quality than other proposed algorithms, even in case of blurring problem, which is very common in surveillance cameras for mobile targets. Fig. 16 shows the relative error of the reconstruction obtained with DART and SIRT as a function of a squared Gauss blur filter with $\sigma$ of 10. It can be noticed that DART keeps the relative error below 3%.

Figure 16: Relative error obtained with DART and SIRT as function of Gaussian blur. (The relative error in this case is divided by 10).
5. Interface

The following section demonstrates the functionalities of a graphical interface, developed in Matlab, in order to perform simulation experiments with DART in a comfortable way. The figure below shows a capture of the graphical interface. In what follows, all its functionalities are explained.

Figure 17: Capture of the graphical interface

1. License plate selector: Allows selection among several images.
2. HR image corresponding with the selected license plate. This image will be downsampled in order to generate the set of LR images.
4. Shifting: Allows modify the maximum amount of shifting. For each LR image random values are generated for both directions, horizontal and vertical.
5. Gaussian Blur: Allows blurring the image. The parameters are the size of the mask in X and Y axis, and sigma that corresponds to the standard derivation of the Gaussian distribution.

6. HR images for each band. Shown in order to compare the resulting reconstructions for each band.

7. Reconstructed images for each band.

8. Number of selected gray levels for each band. This amount depends on the value of the threshold of the histogram curve.

9. Number of iterations: Allows select the number of iterations between 1 and 25.

10. Number of LR images: Allows select the number of LR images between 1 and 10.


12. Autorun: Executes a function previously programmed by the user where we can parameterize several followed executions in order to compute statistics or generate graphics.

13. Run: Simple execution with the values of parameters selected in the interface.

14. Reconstructed HR image.
6. Conclusions

We have evaluated the response of DART applied to the problem of license plate reconstruction. The DART algorithm combines the efficiency of iterative algebraic methods from continuous tomography with the power of discrete tomography to compute accurate HR reconstructions from relatively few LR images. Simulation experiments demonstrated that the DART algorithm is capable of computing reconstructions of high quality from a small number of LR images. The efficiency of the algorithm with binary images is high, but also had presented good results with color images. With the addition of Gaussian noise and blurring to the LR images the result obtained had been satisfactory, offering a better result than other algorithms proposed, such as SIRT.
7. References


