

# Single Image Super Resolution via Sparse Reconstruction

Maarten C. Kruithof<sup>a</sup>, Adam W.M. van Eekeren<sup>a</sup>, Judith Dijk<sup>a</sup>, Klamer Schutte<sup>a</sup>

<sup>a</sup>TNO, Oude Waalsdorperweg 63, Den Haag, The Netherlands

## ABSTRACT

High resolution sensors are required for recognition purposes. Low resolution sensors, however, are still widely used. Software can be used to increase the resolution of such sensors. One way of increasing the resolution of the images produced is using multi-frame super resolution algorithms. Limitation of these methods are that the reconstruction only works if multiple frames are available furthermore these algorithms decreases the temporal resolution. In this paper we use a sparse representation of an overcomplete dictionary to significantly increase the resolution of a single low resolution image. This allows for a higher resolution gain and no loss in temporal resolution. We demonstrate this technique to improve the resolution of number plates images obtained from a near infrared roadside camera.

**Keywords:** Compressive Sensing, Super Resolution, Sparse Reconstruction

## 1. INTRODUCTION

In the field of defense and security it is very important to have high resolution sensors for recognition purposes. However there are many low resolution sensors being used, for instance due to budget limitations or weight limits. One way of increasing the resolution of the images produced by such sensors is using Super-Resolution (SR) reconstruction. During the last decade numerous SR methods have been reported in the literature. Reviews can be found in<sup>1-3</sup>. Multiframe SR reconstruction is the process of combining a set of undersampled (aliased) low-resolution (LR) images to construct a high-resolution (HR) image or image sequence. The disadvantage of such methods is that they need multiple frames. Furthermore these methods need movement of the scene to be able to increase the resolution. These methods also have problems with moving objects, which was partially solved by Eekeren *et al.*<sup>4-6</sup> When only a single frame or multiple frames with complex motion fields are available other methods are needed. One such method is sparse reconstruction using an overcomplete dictionary as described by Yang *et al.*<sup>7</sup> Here a high resolution image is reconstructed from a library with generic low resolution and corresponding high resolution image patches. We chose the method of Yang *et al.*<sup>7</sup> because it was the most robust and straightforward algorithm used in super resolution using sparse reconstruction.

We simplified the algorithm from<sup>7</sup> and used it to increase the resolution of near infrared numberplate images. By choosing a subset of images (only numberplates) we were able to obtain very good results with a small database.

The paper is outlined as follows, in Section 2 we will explain the theory of the reconstruction algorithm as used by<sup>7</sup>, then we will explain the simplifications we made to this algorithm. Next in Section 3 we show in more detail the method used to create a database and reconstruct the images and how we tested the robustness of the method. Then we will show the results of the experiments in Section 4. Finally we discuss the results in more detail in Section 5.

## 2. THEORY

In single image super resolution we need to find the high resolution image,  $\mathbf{X}$ , from a blurred and downsized image,  $\mathbf{Y}$ :

$$\mathbf{Y} = \mathbf{D}\mathbf{H}\mathbf{X}, \quad (1)$$

where  $\mathbf{D}$  is the downsampling operator and  $\mathbf{H}$  is a blurring filter.

We use a simplified version of the method introduced in<sup>7</sup> to solve equation (1). We assume that  $\mathbf{D}$  and  $\mathbf{H}$  are known. Using  $\mathbf{D}$  and  $\mathbf{H}$  we extract a low resolution patch from each high resolution patch. These patches are used to train a low resolution dictionary  $\mathbf{D}_l$  and a high resolution dictionary  $\mathbf{D}_h$  by solving:

$$\min_{\{\mathbf{D}^h, \mathbf{D}^l, \mathbf{Z}\}} \|\mathbf{X}_c - \mathbf{D}_c \mathbf{Z}\|_2^2 + \hat{\lambda} \|\mathbf{Z}\|_1, \quad (2)$$

where  $Z$  is the representation of  $X_c$  in the dictionary  $\mathbf{D}_c$

$$X_c = \begin{bmatrix} \frac{1}{\sqrt{N}} X^h \\ \frac{1}{\sqrt{M}} Y^l \end{bmatrix}, \quad \mathbf{D}_c = \begin{bmatrix} \frac{1}{\sqrt{N}} \mathbf{D}_h \\ \frac{1}{\sqrt{M}} \mathbf{D}_l \end{bmatrix}, \quad (3)$$

$$\hat{\lambda} = \lambda \left( \frac{1}{N} + \frac{1}{M} \right),$$

with  $N$  and  $M$  the dimensions of the high resolution and low resolution image patches,  $X^h$  the vector representation of the high resolution image patches,  $Y^l$  the vector representation of the corresponding low resolution patches and  $\lambda$  the sparseness parameter. This equation is extensively studied.<sup>8-10</sup> Since numberplates only have a limited number of possible patches, we simply obtained two dictionaries containing low resolution and corresponding high resolution patches by taking image patch pairs directly from the high resolution and corresponding low resolution images and normalizing the intensity of the image patches.

Using this combined high resolution and low resolution dictionary equation (1) is solved for each patch using:

$$\min \lambda \|\alpha\|_1 + \frac{1}{2} \|\tilde{\mathbf{D}}\alpha - \tilde{y}\|_2^2, \quad (4)$$

where  $\tilde{\mathbf{D}} = \begin{bmatrix} F\mathbf{D}_l \\ \beta P\mathbf{D}_h \end{bmatrix}$ ,  $\tilde{y} = \begin{bmatrix} Fy \\ \beta yw \end{bmatrix}$ ,  $y$  the vector notation of a low resolution patch,  $w$  the vector notation of a high resolution patch and  $\alpha$  the representation of the patch in the dictionary. The parameter  $\beta$  determines the trade off between matching a low resolution patch and finding a high resolution patch which fits its neighbors. The matrix  $P$  extracts the overlap between the images and  $F$  is a feature extraction operator. For the feature extraction they choose the first and second derivatives of the image in the horizontal and vertical direction.

We found that due to the simple nature of the numberplate images, a light background and dark characters, the overlap fitting was not necessary. So to simplify the algorithm we set  $\beta$  to zero which simplifies equation (4) to

$$\min \lambda \|\alpha\|_1 + \frac{1}{2} \|F\mathbf{D}_l\alpha - Fy\|_2^2 \quad (5)$$

### 3. METHOD

We obtained the dictionary by taking image patch pairs from 287 random number plates, see figure 1, from a library of 6900 Dutch number plates. The high resolution patches were taken from these random images and the low resolution patches were taken from the images downsampled using bicubic interpolation and a lowpass antialiasing filter with a factor of 5 as shown in figure 2. The low resolution patches are always 3 by 3 pixels and have an overlap of one pixel. The high resolution patches have a size determined by the chosen upscaling parameter, in this case 5 and thus have a size of 15 by 15 pixels. This dictionary was used to increase the resolution of random number plate images taken from the library downsampled with a factor of 5 using bicubic interpolation and a lowpass antialiasing filter. These numberplates were not in the dictionary. In addition, we always compare the reconstructed image with an image rescaled using bicubic interpolation.

To test the robustness of this method we also tested increasing the resolution of images which were rescaled using a different method. This simulates the fact that an estimated rescale function may not be the same as the real rescale function.

Finally we tested the method on an visible light image of a numberplate taken with a mobile phone at a large distance.

### 4. RESULTS

An example of a reconstruction is shown in figure 1. Here it is quite clear that the number plate reconstructed with sparse representation is much better readable than the number plate reconstructed using bicubic interpolation. To test the quality of the reconstruction we asked six people to read the reconstructed numberplates. We compared their results with the actual numberplates to get an estimate for the quality of the reconstruction. We

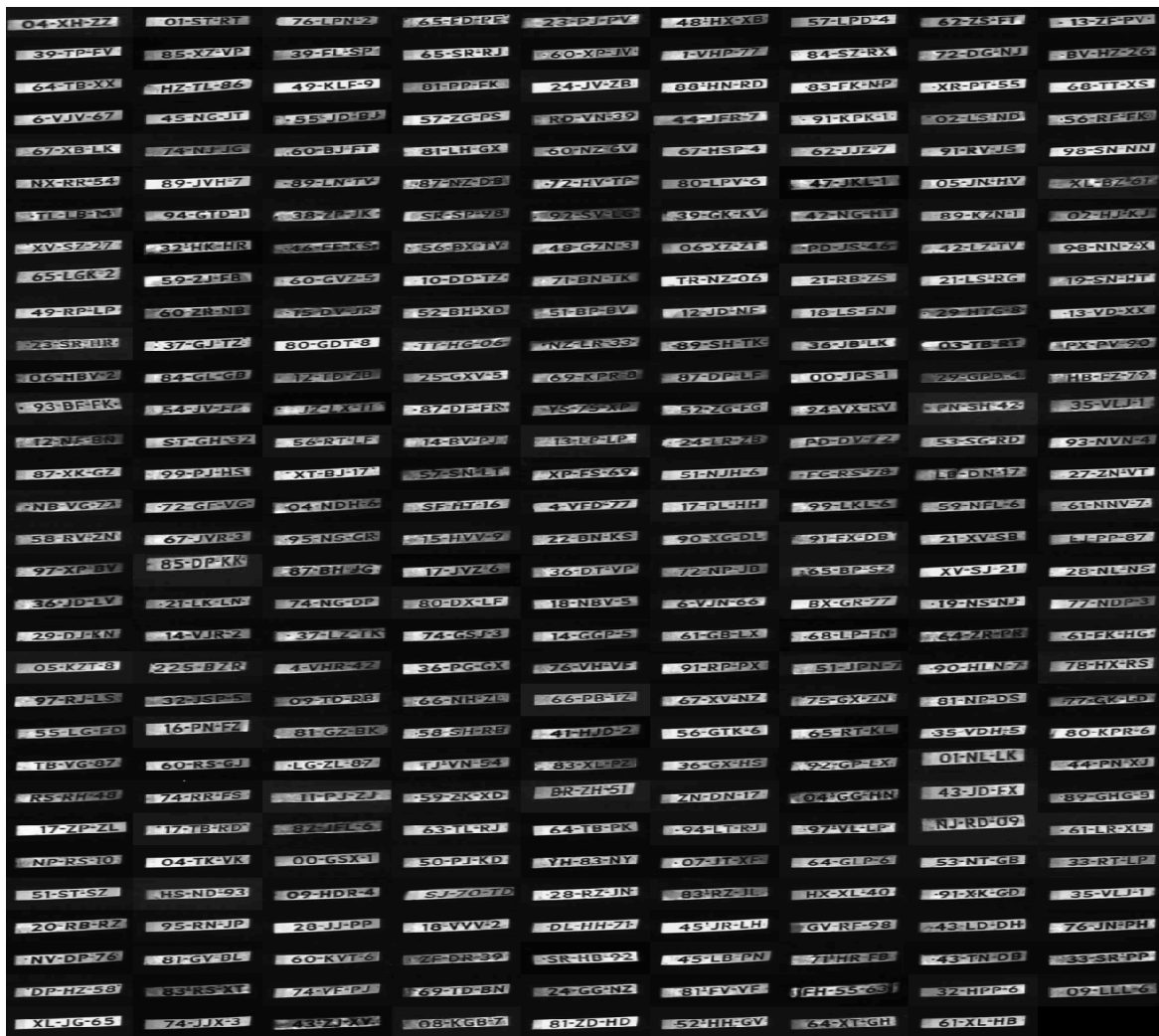
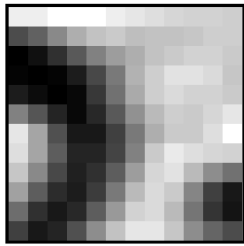


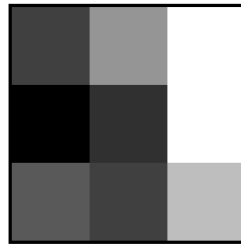
Figure 1. High resolution images of 287 random number plates taken from a library of 6900 number plates.



a



b



c

Figure 2. Example of how the dictionary is built. A patch of 15 by 15 pixels (b) is taken from the high resolution image (a red square) and rescaled to 3 by 3 pixels (c). (b) and (c) are normalized in intensity.



a



b



c



d

Figure 3. Number plate reconstruction. The high resolution image (a) is rescaled (b) and reconstructed using bilinear upsampling (c) and via sparse reconstruction (d). Sparse reconstruction matches the original numberplate (a) much better.

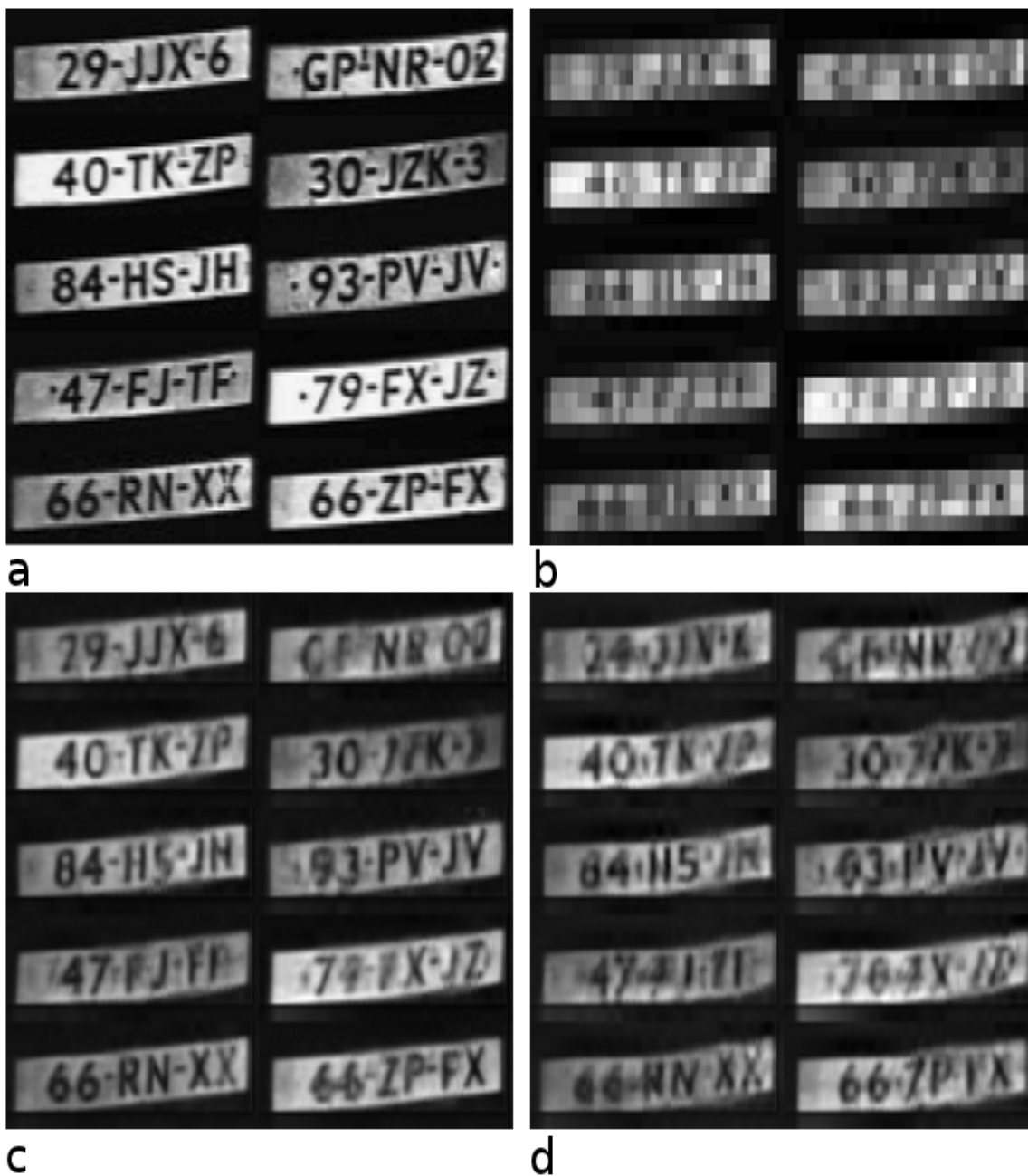


Figure 4. Examples of a different rescale methods and the results. Here the high resolution images (a) were rescaled using bicubic interpolation (b) and upsampled using sparse reconstruction (c). The high resolution images were also blurred with a Gaussian filter with a  $\sigma$  of 2 pixels, every fifth pixel was taken and upsampled using sparse reconstruction (d).

used 10 numberplates with 6 characters each. On average 80% of the characters were readable. This is much better than the result with just bicubic interpolation, where only 27% of the numbers and letters were readable. To test the robustness of the method we used a different rescaling as shown in figure 4. We followed the same procedure and asked several people to read the numberplates. In this case we used 10 numberplates with 6 letters and numbers. Here 62% of the letters were read correctly. So even though our model for the rescaling was different we were still able to reconstruct a large portion of the letters and numbers.

We also used the same procedure to reconstruct an image taken with a mobile phone camera. We created a different library for the reconstruction since the upscaling factor was not 5 but 4. The low resolution image was taken with a mobile phone at a large distance, the resulting reconstruction was comparable to bicubic upsampling, see figure 5. However, the lines of the numbers and letters are clearer which is especially apparent in the 8 at the end of the sequence. However the results are not as good as seen from images of earlier experiments, most likely due to compression artefacts present in this image.



Figure 5. Reconstruction of a numberplate taken with a mobile phone camera. The high resolution image (c) is taken at a distance of 1 meter. The low resolution image is taken at 10 m and converted to grayscale (a). The low resolution image is reconstructed using bicubic upsampling (b) and via sparse reconstruction (d).

## 5. CONCLUSION AND DISCUSSION

To conclude we found that super resolution using sparse reconstruction can be a useful tool in increasing the readability of numberplates. We found that 80% of the characters were readable if the exact model rescaling was known. A huge increase compared to bicubic upsampling where only 27% of the characters were readable. This decreased to 62% if the exact model was not known and the image to be reconstructed came from the same

source. If the image came from a different source (a mobile phone instead of a near infrared traffic camera) the results were not as good as taking images from the same source but still slightly better than a bicubic upscaling.

## REFERENCES

- [1] Park, S. C., Park, M. K., and Kang, M. G., "Super-resolution image reconstruction: a technical overview," *IEEE Signal Processing Magazine* **20**(3), 21–36 (2003).
- [2] Farsiu, S., Robinson, D., Elad, M., and Milanfar, P., "Advances and challenges in super-resolution," *International Journal of Imaging Systems and Technology* **14**, 47–57 (2004).
- [3] van Eekeren, A. W. M., Schutte, K., Oudegeest, O. R., and van Vliet, L. J., "Performance evaluation of super-resolution reconstruction methods on real-world data," *EURASIP Journal on Advances in Signal Processing*, 1–11 (2007). Article ID 43953.
- [4] van Eekeren, A. W. M., Schutte, K., Dijk, J., de Lange, D. J. J., and van Vliet, L. J., "Super-resolution on moving objects and background," in [*Proc. 13th International Conference on Image Processing*], **1**, 2709–2712, IEEE (2006).
- [5] Dijk, J., van Eekeren, A. W. M., Schutte, K., de Lange, D. J. J., and van Vliet, L. J., "Super-resolution reconstruction for moving point target detection," *Optical Engineering* **47**(8) (2008).
- [6] van Eekeren, A. W. M., Schutte, K., and van Vliet, L. J., "Multiframe super-resolution reconstruction of small moving objects," *IEEE Transactions on Image Processing* **19**(11), 2901–2912 (2010).
- [7] Yang, J., Wright, J., Huang, T. S., and Ma, Y., "Image super-resolution via sparse representation," *IEEE Transactions on Image Processing* **19**(11), 2861–2873 (2010).
- [8] Bruno A. Olshausen, D. J. F., "Sparse coding with an overcomplete basis set: a strategy employed by v1," *Vision Research* **37**, 3311–3325 (1997).
- [9] Lee, H., Battle, A., Raina, R., and Ng, A. Y., "Efficient sparse coding algorithms," *Advances in Neural Information Processing Systems* **19**, 801–808 (2007).
- [10] Murray, J. F. and Kreutz-Delgado, K., "Learning sparse overcomplete codes for images," *J. VLSI Signal Process. Syst.* **45**, 97–110 (2006).