Super-resolution of bar codes

Donald G Bailey[†]

Institute of Information Sciences and Technology, Massey University, New Zealand

ABSTRACT

This paper describes how under-resolved images of bar codes may be read by suitable processing. A two-dimensional image of a bar code with insufficient resolution to be able to resolve the individual bars is processed to give a high-resolution image. For this to work, the bar code (or camera) must be slightly rotated to give a fraction of a pixel offset between rows. Since the bars are straight, the offset relative to the first complete row of the bar code increases linearly with vertical position in the image. This offset between rows results in a shift in phase that is proportional to both offset and frequency. A phase image is formed by Fourier transforming each row in the image, and retaining the phases. By subtracting the first row from subsequent rows of the phase image, a surface is fitted to give the offset between rows. A high-resolution image is then formed by interleaving the pixel values from rows where the offset is nearest to the new pixel spacing. This image appears blurred because of the area sampling caused by the sensor, combined with the low pass response of the camera electronics. By modelling the image capture system, the point spread function may be estimated and then removed by using inverse filtering in the frequency domain. The offset between the rows is then removed by using a linear phase filter. This allows the rows within the resultant image to be averaged to reduce noise.

Keywords: bar code reading, image fusion, super-resolution, aliasing, registration, image reconstruction, inverse filtering

1. INTRODUCTION

Image fusion involves the combining of data from a number of image sources in such a manner that the output image contains more information than that contained in any of the individual images. The image sources are often different in their properties or characteristics. An important use of image fusion is to improve image quality by using multiple images taken from a single source. One example of this is the averaging a number of images of a static scene to improve the signal to noise ratio[1]. The improvement in signal to noise ratio comes at the expense of temporal resolution. A related image fusion problem is to trade temporal resolution for spatial resolution, by combining several low-resolution images to construct a high-resolution image[2,3].

The pixels of an image provide a series of samples of the world. The sampling density, or the spacing between the pixels, limits the achievable resolution. However if a series of images is captured, each with the samples in slightly different locations, the combined sample density is higher that that of any single image. Constructing a high-resolution image consists of registering all of the individual low-resolution images to one another, and then resampling the ensemble using a single sampling grid.

To investigate some of the issues involved in this process, a simplified problem was considered—that of reconstructing a high-resolution bar code from a low-resolution image. While the ability to read a low-resolution bar code image is a useful problem in its own right, images of bar codes offer a number of simplifications that allow the difficulties in constructing a high-resolution image to be determined, without being confounded by difficulties inherent in the data itself. These simplifications are:

1) A bar code is really a one-dimensional image. The reduction in dimensionality simplifies both the registration and the resampling steps considerably.

2) A two-dimensional image of a bar code provides an ensemble of independent one-dimensional bar code images. Therefore rather than capturing a series of images, a single two-dimensional image provides all of the required input data.

[†] E-mail address: D.G.Bailey@massey.ac.nz

Part of the SPIE Conference on Machine Vision Systems for Inspection and Metrology VII
Boston, Massachusetts • November 1998 SPIE Vol. 3521 • 0277-786X/98/$10.00

3) The registration step is simplified because the rows of the two-dimensional bar code image all have the same relative offset. Determining the offset between the one-dimensional images from any two rows gives the offset for any row within the image.

4) Bar code images are well defined and have high contrast making it easier to evaluate the quality of the high resolution reconstruction.

For there to be more information in the ensemble than in a single image, a single image must not contain all of the information. If each low-resolution image was sampled at the Nyquist rate (or higher) then each image contains sufficient information to be able to recover the original data exactly by using sinc interpolation. This means than any single image would be able to provide all of the information required for reconstruction at any desired resolution. Having multiple independent images would not provide an improvement. Therefore, for multiple images to provide additional information, the sample frequency for each individual image must be below the Nyquist rate, and the images be subject to aliasing. The process of constructing a higher resolution image untangles the aliased information, so that the output image contains more information than that available from any of the individual input images.

Before manipulating the bar code images, it is useful to know about their information content. Bar codes consist of a series of alternating bars and spaces of varying widths. All of the widths are integer multiples of the unit bar width. The images considered here are of UPC (Universal Product Code) bar codes[4], which encode 12 decimal digits. Each digit is encoded by a pattern of two bars and two spaces occupying a width of 7 units. There is also a guard band of alternating bars and spaces occupying 3 units at each end of the bar code, and another guard band in the centre occupying 5 units. The total width of a bar code is 95 units. A typical bar code pattern is shown in figure 1.

Figure 1: An "ideal" bar code image.

The bar code can be considered to consist of a series of unit width rectangular pulses. This may be decomposed into a single rectangular pulse of unit width convolved with a modulated train of delta functions. The frequency of the delta train, ω_0, corresponds to the unit bar width, and the modulation pattern consists of a 0 for a space and a 1 for a bar. Figure 2 shows the image decomposition and corresponding frequency spectra. In the frequency domain, the bar code image consists of the broad band signal scaled by $\sin\omega_0/\omega_0$. Since the rectangular pulse in the spatial domain corresponds exactly to the spacing of the delta train, in the frequency domain the nulls of the $\sin\omega_0/\omega_0$ coincide exactly with the peaks at $n\omega_0$, cancelling them.

Since the sinc function decays only slowly with frequency, there will be some aliasing even when an apparently adequate sample rate is used. The high frequency information comes from the sharp edges of the bars. Not all of this information is actually required to represent the data contained within the bars. The highest frequency of the modulating waveform occurs when there are alternating bars and spaces of unit width. The corresponding frequency is $\omega_0/2$, which is half of the width of the main lobe of the sinc multiplier in the frequency domain. Band-limiting the bar code image to this frequency gives a bar code with rounded edges and ripples in the wider bars, but it can still be unambiguously restored using simple thresholding (figure 3).

Since UPC bar codes are 95 units wide, a minimum of 95 samples (at a sample frequency of ω_0) is required across the width of the bar code to enable the bars to be resolved unambiguously. This minimum assumes that the information in the bars is not distorted as a result of aliasing, and that the samples are in the middle of the bars not on the edges. In practice, since the main lobe of the $\sin\omega_0/\omega_0$ extends to ω_0, it is suggested that a minimum sample frequency of $2\omega_0$ will be necessary unless a low-pass filter is employed to prevent aliasing.

To demonstrate super-resolution, it is necessary for the bar code to occupy less than 95 pixels in each row of the image. Figure 4 shows two synthetic images, with 107 and 77 samples respectively across the width of the bar code. The bar codes are tilted slightly to enable the bar patterns to be recovered. There is noticeable aliasing in figure 4(a), even though there are more than the minimum number of samples across the width of the bar code. This is because of the interference caused by the main lobe of the $\sin\omega_0/\omega_0$ function. The bar codes are unresolveable in figure 4(b) when each row of the image is viewed independently.

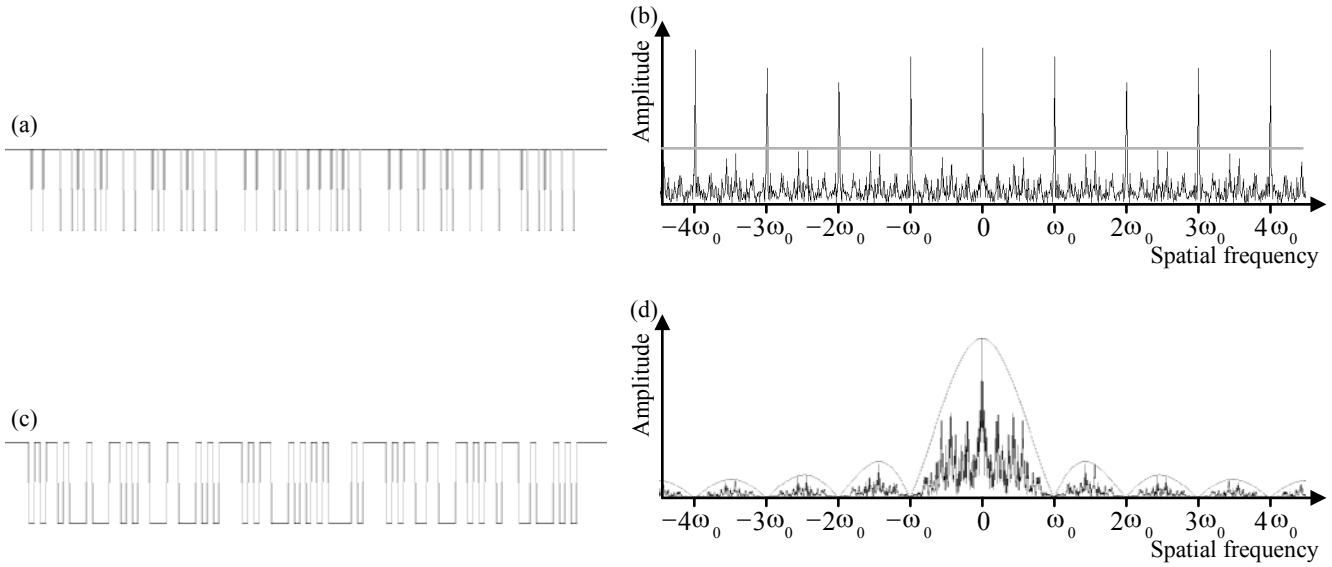

Figure 2: The frequency content of a bar code. (a) Representing the bar code as a modulated delta train, and (b) the corresponding frequency spectrum. The grey line is a reference level corresponding to the sinc envelope in (d). (c) Profile of the actual bar code, and (d) its frequency spectrum with a sinc envelope in grey. Note the DC value has been truncated in (d) for clarity.

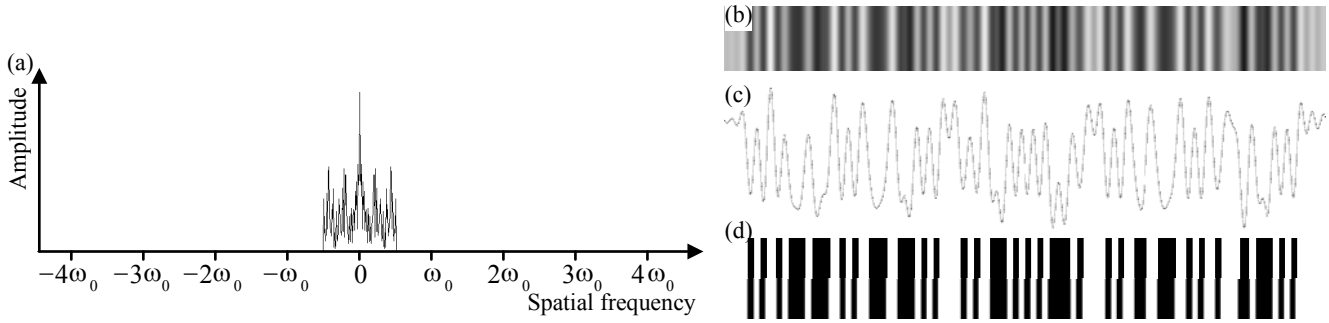

Figure 3: The effect of band-limiting the frequency spectrum. (a) The band-limited spectrum (the DC value has been truncated for clarity), and (b) the resulting bar code image. (c) The profile of the band-limited bar code. (d) After thresholding to recover the bars. The bottom part of the image shows the original bar code, showing a slight shifting in bar positions.

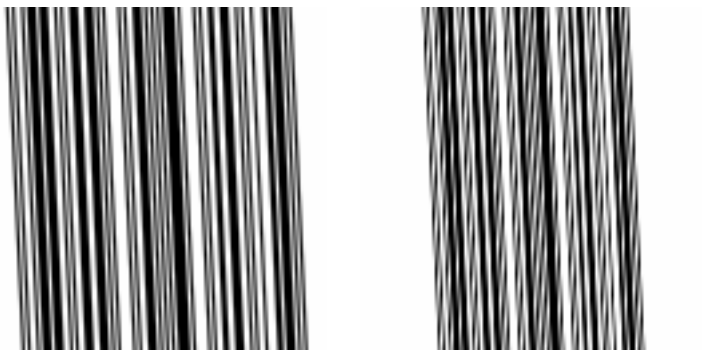

Figure 4: Two synthetic images of bar codes. (a) 107 samples taken over the width of the bar. (b) 77 samples taken over the width of the bar.

2. REGISTRATION

An ensemble of one-dimensional images may be obtained from a single two-dimensional image if the bar code is slightly tilted within the image. Each row in the image is a one-dimensional image of the bar code. The slight rotation means that the samples in each row are offset slightly from those in the other rows. Therefore they form an independent set of samples, each at a slightly different location from the others. To combine the low-resolution images, the individual images must first be registered to one another. For the bar code image, this involves determining the horizontal offset or shift between successive rows within the two-dimensional image.

While there are several approaches to registration[5], it was decided to use the phase information in the frequency domain. A spatial shift corresponds to a phase shift in the frequency domain proportional to both frequency and the size of the shift in the image. If an image is shifted by a pixels, the shift in the frequency domain is given by[6]

$$f(x+a) \Leftrightarrow F(\omega)e^{j\omega a} \tag{1}$$

where $f(x)$ the pixel value as a function of position x, $F(\omega)$ is spatial frequency content of the image, and $\Leftrightarrow$ indicates a Fourier transform relationship. To make use of this relationship, a phase shift image was obtained by taking the Fourier transform of each row in the bar code image, retaining only the phase. The phase of the first row in the image was then subtracted from all subsequent rows to obtain the phase shift relative to the first row in the image. Let r be the row number and ϑ be the phase shift:

$$\vartheta = \angle F_r(\omega) - \angle F_0(\omega) = \angle e^{j\omega ra} = \omega ra \tag{2}$$

The phase shift image contains the frequency dependent phase relative to the first row in the image (figure 5). As the phase of each image is in the range $\pm\pi$, the difference is in the range $\pm 2\pi$. Starting at the DC component of the first row, the phase is unwrapped by adding or subtracting 2π to each value to minimise the difference in phase with both the same frequency on the row above and the immediately lower frequency on the same row (figure 5(b)). This simple approach works well for the low frequencies where the amplitude is large, although at the higher frequencies where the amplitude is smaller, the phase is more easily perturbed by aliasing and noise.

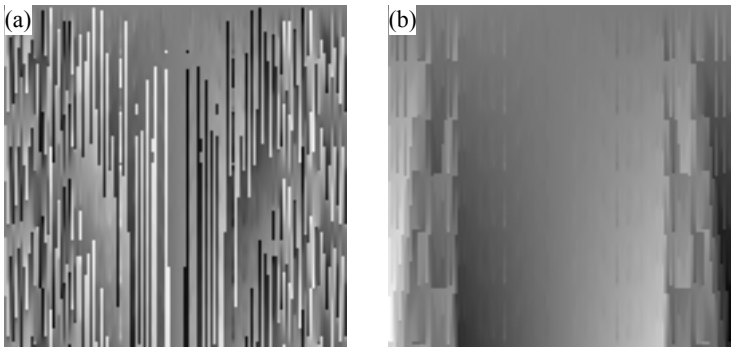

Figure 5: The phase shift image obtained from the under-resolved bar codes in figure 4(b).
(a) The raw phase shift image and (b) after unwrapping the phase.

The offset per row, a, is found by doing a least squares fit of the phase surface predicted by equation (2) to the phase image.

$$a = \frac{\sum_{r,\omega} \vartheta(r,\omega) r\omega}{\sum_{r,\omega} r^2\omega^2} \tag{3}$$

In doing the fit, only the central 25% of the phase image is used because the higher frequency phase terms are subject to error. For this image, the measured offset is $a = 0.0455$ pixels per row.

3. RESAMPLING

Since each row of the image has a slight offset, by selecting and interleaving the samples from the different rows, we can increase the sampling rate. While in principle, any increase in sampling rate up to $1/a$ may be obtained in this way, it is convenient to have an integer multiple of the original sampling rate because that allows a small number of complete rows to be selected and interleaved (see figure 6). If α is the increase in sampling rate, then samples from rows

$$r_n = \text{round}(n / \alpha a) \quad 0 < n < \alpha - 1 \tag{4}$$

are interleaved. This effectively selects the rows with offsets nearest to the desired sample spacing. While interpolation between adjacent rows could be used, the position error in selecting the nearest row is less than $\alpha a / 2$, which is small for small offsets. For α=4, with the current image, rows 0, 5, 11, and 16 are interleaved (with sample position errors of 0, 0.09, 0.002, and 0.09 respectively).

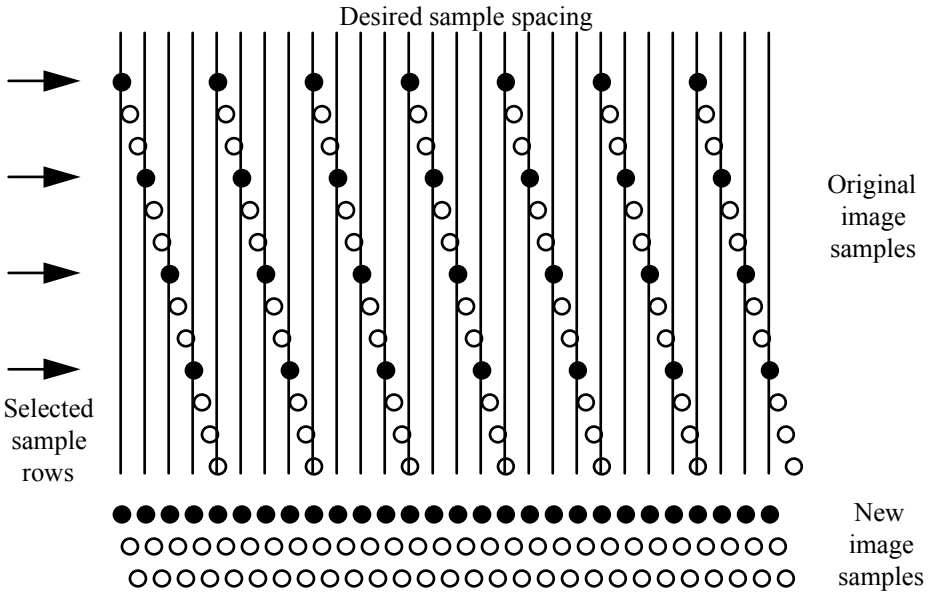

Figure 6: Increasing the sampling rate by interleaving pixels from rows with offsets nearest to the desired sample spacing.

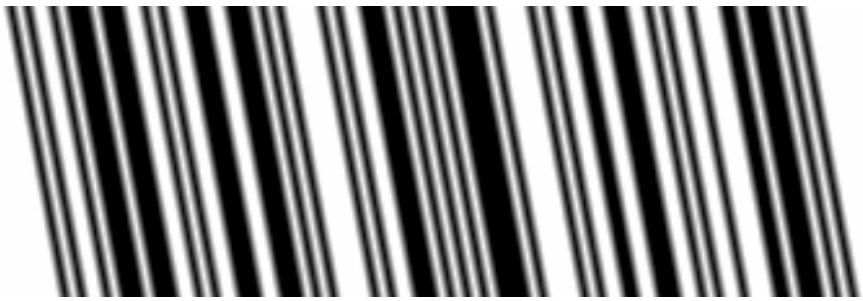

Figure 7: The resampled high-resolution image.

Rather than construct a single high-resolution image of the bar codes, a series of high-resolution images is constructed using the r_n as offsets (see figure 7). This gives additional data that may be averaged later to improve the signal to noise ratio. The slope of the bar codes within the constructed image is increased by a factor of α because of the horizontal sample rate is increased but the vertical sample rate remains unchanged. The slight blurring in the image results from the interpolation used when synthesizing the original low-resolution image.

4. PRACTICAL CONSIDERATIONS

The preceding sections demonstrate the technique of obtaining a high-resolution bar code image from a low-resolution two-dimensional bar code image. In practice, the process is not that simple—there are two complicating factors. The first is that the bar codes are not ideal, and the second results from complexities in the image capture process.

The actual widths of the bars and spaces will not be exact because of ink spreading during the printing process. This affects the frequency content of the bars in the following way. Since each bar is slightly wider than it should be, and the bars vary in width up to 4 units wide, the $\sin\omega_0/\omega_0$ envelope is smeared. This has two effects (compare figure 8 with figure 1(d)). The first is that the nulls at $n\omega_0$ do not go completely to zero, so the peaks from the delta train are not completely cancelled out. The resulting peaks, particularly those at $\pm\omega_0$, result in increased aliasing. The second effect is that the higher frequency side lobes are reduced in amplitude, reducing potential aliasing from that cause.

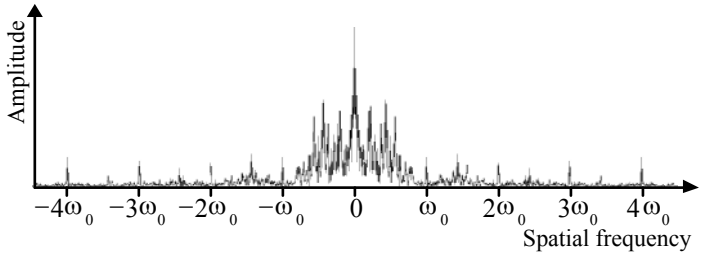

Figure 8: Effect of ink spread (synthetic image). Note the DC value has been truncated for clarity.

Capturing an image for processing is not a simple process. There are several steps involved, and each step results in a deterioration of the quality of the data.

1) **Lens system:** Lenses never focus perfectly so there is always some blur even when in best focus. This will be exacerbated if the image is slightly out of focus. To a first approximation, the two-dimensional point spread function (PSF) of the lens may be considered as a circular disc. The projection of this disc along the direction of the bars gives the one-dimensional point spread function (figure 9a). The effect of the lens PSF is to attenuate the high frequencies, effectively acting as a low pass filter. Determining the width of the PSF is not a straightforward procedure since it depends on the lens and aperture settings. This is further complicated by the fact that the PSF is also spatially variant, being slightly wider on the edge of the field of view than in the centre. Any radial distortion introduced by the lens would also need to be taken into account in the resampling process.

The quality of the lens is often chosen to match the sensor resolution, although for most solid state sensors, this is not likely to be the limiting factor. However it can be difficult to determine the best focus when the (low-resolution) sensor is used to evaluate the focus. By reducing the illumination and opening the aperture to give a reduced depth of focus, focusing becomes more critical enabling the location of best focus to be determined more easily. After focusing, the illumination and aperture can be restored.

2) **Image sensor:** Images are not point-sampled by the sensor, but area-sampled. Since the bar code is at an angle relative to the sensor, the point spread function of the image sensor will be slightly trapezoidal. The trapezoid may be considered as a convolution of two rectangular functions, each contributing a $\sin\omega/\omega$ low pass characteristic. However, for small tilts in the bar code, one of the rectangular functions will be quite narrow. For modest improvements in resolution, the response of the narrow rectangular function can be neglected. The type of sensor may have significant bearing on the point spread function. Not all of the available area is active for sensing and there can be significant variations in sensor geometry between different sensor types. These factors affect the width of the PSF. Information on the sensor geometry is readily available from specification sheets, so determining the sensor PSF is straight forward. Both the lens PSF and the sensor PSF act as low pass filters, and reduce the aliasing caused by the sampling process. This will limit the extent to which the resolution may be improved by the process outlined in this paper. To enhance the high-resolution images, inverse filtering may be used to partially compensate for the effects of the sensor PSF.

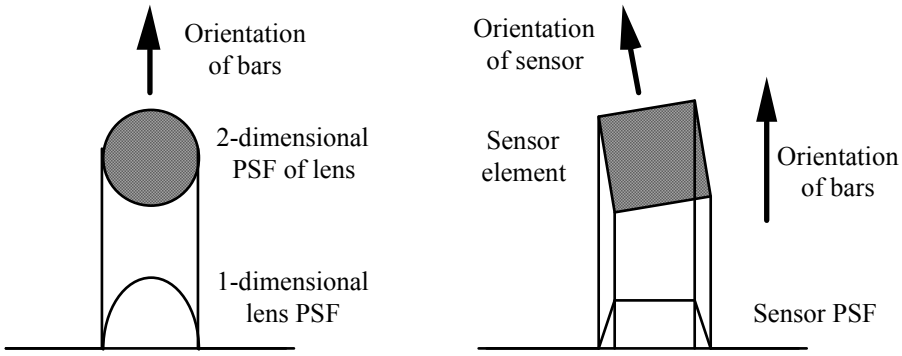

Figure 9: Derivation of the point spread functions (PSFs) of (a) the lens system, and (b) an area sensor.

3) **Camera electronics:** Unless it is a digital camera, the output from the sensor is converted into an analogue video signal. This may involve a sample and hold circuit to reduce clock noise, followed by a low pass filter to interpolate the samples and produce a continuous signal. Some cameras have a high frequency emphasis on the low pass filter to reduce the blur to any sharp edges. If the frequency response of the camera system is known, it may be removed by inverse filtering.

4) **Video frame grabber:** The analogue video signal is then sampled at the video frame grabber, after possibly passing through a low pass filter to reduce aliasing. The sample rate of the video frame grabber will almost certainly be different from that at the sensor. The multiple sampling will introduce another set of aliasing artifacts because the replications resulting from the sensor sampling will not have been completely eliminated by the low pass filter.

The overall response of the image capture process is to strongly attenuate the high frequency content of the image. This has the consequence that the reconstruction process yields a blurred image of the high-resolution bar codes. As described earlier, the aliasing introduced by the sampling process is actually essential to enable a higher resolution image to be constructed. A bar code image was captured (with 77 samples across the width of the bar) and processed to observe the effects of the image capture process. The low-resolution image and its high-resolution reconstruction are shown in figure 10. In comparing figure 10 with figure 7, the low pass response of the system may be observed as a distinct blurring of the bars. In comparing the frequency content of the captured image with that of the synthetic image (figure 11), the loss of contrast is clearly a result of the severe attenuation of high frequency information in the image.

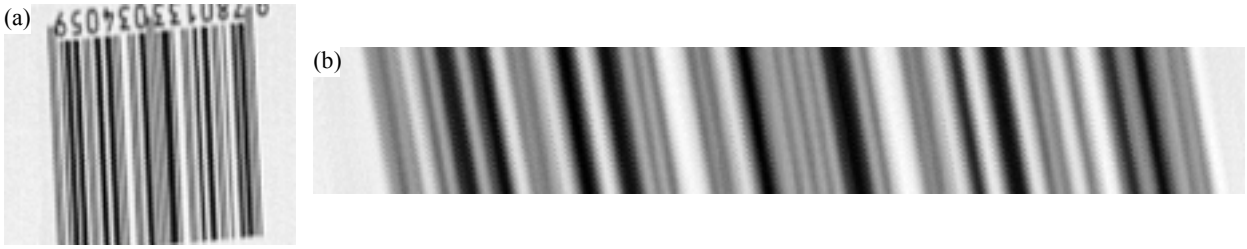

Figure 10: Applying the reconstruction described in sections 2 and 3 to a real bar code image.
(a) The original image, and (b) the resampled high-resolution image.

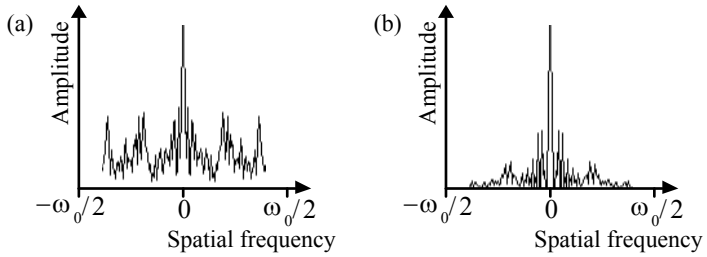

Figure 11: Frequency content of low resolution bar code images. (a) Spectrum of the synthetic image from figure 4(b). (b) Spectrum of the captured image (figure 10(a)). Note the DC value has been truncated in both plots for clarity.

5. REMOVING SYSTEM EFFECTS

To improve the quality of the reconstructed high-resolution image, the low pass characteristic of the imaging system needs to be estimated and its effects removed. By comparing the frequency content of the synthetic image and the real image after reconstruction (figure 12), the transfer function between them may be estimated. In comparing the amplitudes between figure 11 and 12, it can be seen that the lower frequencies (below about $\omega_0 / 4$) have been recovered successfully, but the higher frequencies are severely attenuated. As indicated in the introduction (and illustrated in figure 3) it is necessary to restore the frequency content up to at least $\omega_0 / 2$ to recover the bars.

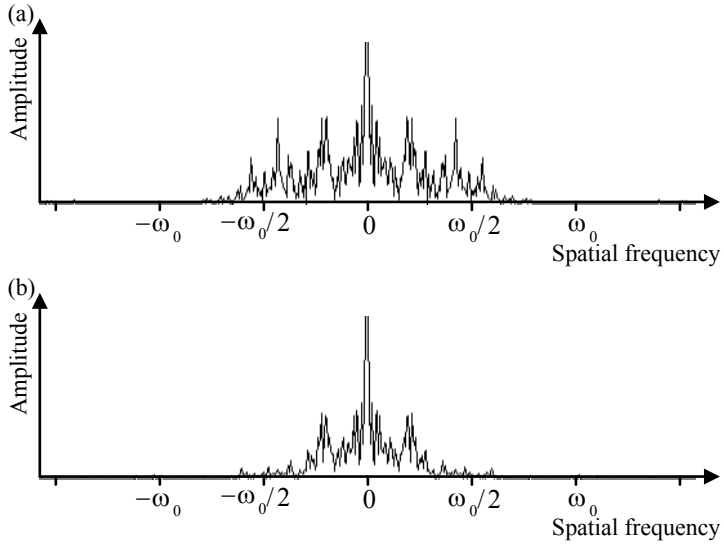

Figure 12: Frequency content of the reconstructed images. (a) Spectrum from the synthetic image (figure 7). (b) Spectrum from the real image (figure 10(b)). Note the DC value has been truncated in both plots for clarity.

If the system response is linear then the effect of the imaging system transfer function, H_{system}, may be modelled as

$$F_{captured} = F_{synthetic} \times H_{system} \tag{5}$$

This assumes that the synthetic image provides a good estimate of how the image would appear if the limitations of the imaging system were not present. Equation (5) may be inverted to estimate the system transfer function in terms of the frequency content of the high-resolution captured and synthetic images (from figure 12).

$$H_{system} \approx F_{captured} \times \frac{F^*_{synthetic}}{\left| F_{synthetic} \right|^2 + k} \tag{6}$$

The constant k is necessary to prevent division by zero. If k is set too low, the estimated system response will be dominated by noise, and if set too high, the system response will not be estimated accurately. An optimum value was found by trial and error. The estimated system response is shown in figure 13.

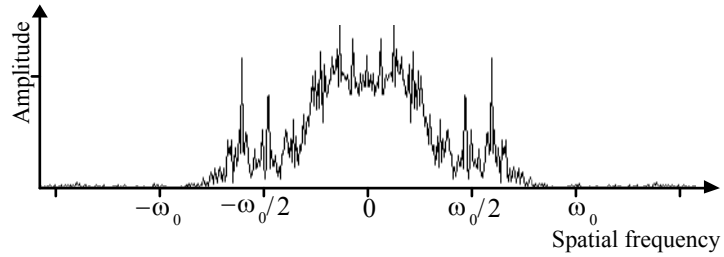

Figure 13: Estimated frequency response of the imaging system as calculated from equation (6).

This transfer function is not a simple product of the individual transfer functions of the imaging system components because of aliasing. The central lobe corresponds to a low pass filter with a cut frequency of approximately 4 MHz, which is typical for a video camera[7]. The secondary lobes result from the higher frequency information being aliased into the pass band of the low pass filter by the sampling at the sensor. This information is later unfolded by the reconstruction process, giving humps in the frequency response outside the main pass band. This makes accurate modelling of the system transfer function more difficult. In the experiments here, the estimated system response from figure 13 was used.

The system transfer function may be used to construct an inverse filter[8] for removing the system response from the image:

$$F_{reconstructed} \approx F_{captured} \times \frac{H^*_{system}}{\left\|H_{system}\right\|^2 + k_2} \tag{7}$$

When this filter is applied to the constructed high-resolution bar code, the image of figure 14(a) results. This considerably reduces the blur, but also amplifies any noise present in the image. The noise may be reduced by averaging along the length of the bars. This is easiest if the tilt in the bars is removed. For this, a linear phase filter based on equation (1) was constructed to shift each row by its calculated offset. The rows are then averaged giving the image in figure 14(b). The final step is to remove the slight blur introduced into the image by interpolation when constructing the synthetic image used to estimate the system response. The final output is shown in figure 14(c). While shown here as a series of steps, in practice these three filters are all combined into a single filter.

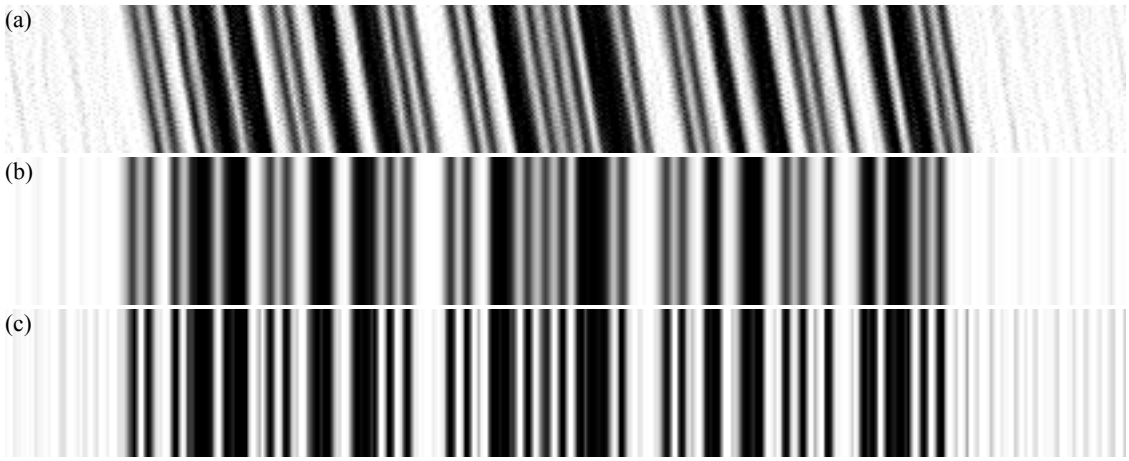

Figure 14: Removing the system transfer function from the bar codes. (a) The result of applying the inverse filter using equation (7). (b) After removing the tilt in the bars and averaging along the length of the bars to reduce noise. (c) After removing the blur resulting from interpolation in the construction of the synthetic image.

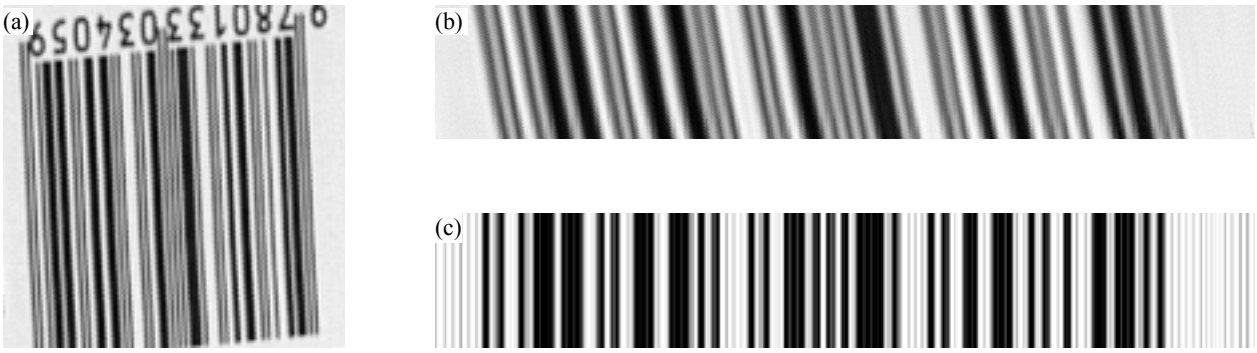

Figure 15: Applying the reconstruction process to another image. (a) The captured image with 107 samples across the width of the bar code. (b) The reconstructed high resolution image. (c) The reconstructed image after removing the system transfer function.

This complete process was applied to another image captured with 107 samples across the width of the bar code (see figure 15). While this image is sampled at greater than the absolute minimum of 95 samples across the width of the bars, there is still considerable aliasing resulting from the high frequency content within the image. This aliasing is completely removed in the high-resolution construction, and the contrast has been considerably restored by removing the system response.

If necessary, the resultant images (figures 14(c) and 15(c)) can be tidied further by using prior knowledge about the type of bar codes. For UPC bar codes, we know that the complete bar code is 95 units wide. By detecting the left and right end of the complete bar code and dividing the width between into 95 equal regions, each region can be classified as either bar or space based on the pixel values within the region.

6. CONCLUSIONS

In this paper, a technique has been demonstrated for improving the resolution of bar code images. Even if the bar code has been under-sampled or is severely aliased, a high-resolution image may be reconstructed. There are several prerequisites for this process to be successful.

1) A two-dimensional image of the bar code must be captured, with the bar code tilted slightly relative to the sensor. This provides an ensemble of related low-resolution one-dimensional images of the bar codes. By combining the information from the ensemble, it is possible to reconstruct a high-resolution image of the bar code.

2) The original image must be aliased. The reconstruction process effectively unscrambles the aliased high frequency information giving an improvement in resolution. If the original image is not aliased, each of the individual lines of the image correspond to $\sin x / x$ interpolations of one another, and no higher frequency information is available to recover.

3) To achieve super-resolution, it is necessary to compensate for the limitations of the image capture system when reconstructing the high-resolution image. If this is not done, the imaging system limits the achievable gains in resolution. The process involves being able to accurately model the system transfer function. The image capture system has a strong low pass response with multiple aliasing caused by multiple sampling. This complicates the modelling of the compensation filter. The low pass responses may be removed by inverse filtering because the high frequency terms go to zero only slowly (apart from at specific frequencies). However, while modest gains in resolution are achievable, large gains would be difficult to obtain because of noise limitations.

While the technique described here is effective at enhancing the resolution of one-dimensional bar codes, it cannot be easily extended to two dimensional images. A two-dimensional bar code image conveniently provides a large number of equally spaced one-dimensional images making resampling straightforward. In two dimensions, the sample density from multiple images will be considerably lower, limiting the extent to which the resolution may be improved.

7. REFERENCES

1. A.R. Weeks, *Fundamentals of Electronic Image Processing*, pp 102-106, SPIE Press, Bellingham, Washington, 1996.
2. R.C. Hardie, T.R. Tuinstra, J. Bognar, K.J. Bernard, and E. Armstrong, "High resolution image reconstruction from digital video with global and non-global scene motion", *Proc. IEEE Inter. Conf. On Image Processing*, Vol I, pp 153-156, 1997.
3. M.C. Hong, M.G. Kang, A.K. Katsaggelos, "An iterative weighted regularized algorithm for improving the resolution of video sequences", *Proc. IEEE Inter. Conf. On Image Processing*, Vol II, pp 474-477, 1997.
4. R. Adams, *Universal Product Code (UPC) and European Article Numbering (EAN),* http://www.adams1.com/pub/russadam/upccode.html
5. A.K. Jain, *Fundamentals of Image Processing*, pp 400-407, Prentice Hall, Englewood Cliffs, New Jersey, 1989.
6. E.O. Brigham, *The Fast Fourier Transform*, Prentice Hall, Englewood Cliffs, New Jersey, 1974.
7. K.G. Jackson, G.B. Townsend, *TV and Video Engineers Reference Book,* Butterworth-Heinemann, Oxford, 1991.
8. K.R. Castleman, *Digital Image Processing*, Prentice Hall, Englewood Cliffs, New Jersey, 1979.