RESOLUTION IMPROVEMENT OF DIGITIZED IMAGES[*]

LIBOR VÁŠA[†] AND VÁCLAV SKALA

Abstract. A quick overview of preprocessing performed by digital still cameras is given along with a brief introduction to spatial-domain Super-Resolution methods, i.e. spatial resolution enhancement methods that create one high-resolution image from a series of low-resolution images shifted by a sub-pixel distance. An improvement applicable to some of existing Super-Resolution methods is presented. Principles of digital photography processing techniques are exploited in order to reduce error in the Super-Resolution process. Results of both simulations and real data experiments are shown to consider the improvement and ideas for future research are given.

Key words. resolution enhancement, digital images, demosaicking, image processing, Bayer array, CCD elements, Super Resolution

1. Introduction. Digital still cameras have reached strong position on the field of photographic industry over the past decade. Although parameters of digital equipment constantly improve, there are still areas where analog equipment offers better performance. One of such areas is resolution of gained images.

It is of course possible to employ a sensor element with higher resolution, but such choice is not always available. In the early beginning of digital imaging there has appeared a simple idea that multiple slightly shifted images of unchanged scene contain more information than a single frame and that such information may be exploited to construct one image of the scene with higher resolution. Such process is usually addressed as Super Resolution (SR).

We will show that SR techniques are capable of improving resolution of images taken by commercially available digital cameras. We will show that applying the techniques directly may impair quality of the result due to preprocessing that takes place within the digital cameras. We will show how to exploit knowledge about used sensor to improve the results.

The rest of the paper is organized as follows: Section 2 will briefly describe preprocessing of data gained by digital still camera sensor, focusing on demosaicking, section 3 will give brief introduction to Super Resolution techniques. In section 4 we will derive a technique to improve the algorithms shown in section 3 and in section 5 we will give description and results of experiments we have carried out. Finally, we will conclude in section 6 and give ideas for future work in section 7.

2. Image preprocessing in digital still camera.

2.1. Preprocessing steps. There is a quite complex preprocessing performed within digital cameras that turns data gained from light sensor into a computer image file. Digital still cameras that are presently available usually employ a CCD or CMOS element that linearly transforms incoming light into electric charge. We will be considering cameras with rectangular shape of sensor cell (there is a minority of cameras using hexagonal grid of cells).

[†]Department of Computer Science and Engineering, Faculty of Applied Sciences, University of West Bohemia in Pilsen (`lvasa|skala@kiv.zcu.cz`).

Most sensors used in cameras turn light intensity into electric charge regardless of wavelength of the incoming light. In order to obtain color images, a color filter is placed in front of each cell that only lets through a certain range of wavelengths. Usually red, green and blue filters are used, arranged in a pattern called Bayer array (see figure 2.1). Measured values are usually scaled to comply with computer image representation (logarithmic scaling that is later compensated by exponential response of a computer display). Different translucency of red, green and blue filters and varying lighting conditions are compensated by multiplication of the measured linear values by some constant in a process called White-Balancing.

The main issue with the Bayer-array equipped cameras is the fact that at each pixel location there is only one color channel value available. The remaining two values have to be computed (interpolated from the surrounding values) in a process called demosaicking.

R	G	R	G	R	G	R	G
G	B	G	B	G	B	G	B
R	G	R	G	R	G	R	G
G	B	G	B	G	B	G	B
R	G	R	G	R	G	R	G
G	B	G	B	G	B	G	B
R	G	R	G	R	G	R	G
G	B	G	B	G	B	G	B

FIGURE 2.1. *Bayer color array*

2.2. Demosaicking. There are many methods that perform the demosaicking process ([2], [6], [7], [11], [10], [12], for comparison see [14]), because intuitive approaches like linear interpolation of nearest measured values lead to disturbing artifacts in the resulting image, which may include blurring or sudden jumps in hue on sharp intensity edges.

For our experiments we have implemented the Cok's[11] constant hue algorithm, because it provides satisfactory results and is easy to implement. The algorithm aims to locally preserve hue by keeping the red/green and blue/green ratios constant. In the first step, green channel values are interpolated using some general technique like linear interpolation of their four orthogonal neighbors. This is possible due to nature of the Bayer array layout which ensures that every position where green was not measured has four neighbors where a measured value is available. In the following step, one of following formulas is used according to position in the Bayer array to compute unknown values of red and blue:

$$X = G\frac{\frac{X_{tl}}{G_{tl}} + \frac{X_{tr}}{G_{tr}} + \frac{X_{bl}}{G_{bl}} + \frac{X_{br}}{G_{br}}}{4} \quad \text{for positions where G was not measured,}$$

$$X = G\frac{\frac{X_l}{G_l} + \frac{X_r}{G_r}}{2} \quad \text{for positions where G was measured,}$$

(2.1)
$$X = G\frac{\frac{X_t}{G_t} + \frac{X_b}{G_b}}{2} \quad \text{for positions where G was measured,}$$

where X_{tl} stands for value of red or green to the top left of the currently processed value.

3. Super Resolution methods. Super Resolution (SR) is a term commonly used for techniques producing one high-resolution image from a series of slightly shifted low-resolution images. This research area was first explored by Huang and Tsai ([9]), who have proposed a frequency domain algorithm to solve the problem. Basic disadvantage of their approach is that their algorithm assumes that input data are point samples of original continuous function. This assumption fails for the case of CCD and CMOS digital cameras, where the sampling process is always integral. Therefore it is not possible to use this (or any similar) approach in our task.

Space domain methods appeared later, exploiting various approaches to the problem ([3], [4], [13], [8]). Common notation for various methods was proposed by Elad and Feuer[3], relation between high resolution image X (represented by a lexicographically ordered column vector of size L^2) and measured set of N low-resolution images Y (represented by lexicographically ordered column vectors of size S_k^2) is expressed by equation (3.1).

$$(3.1) \qquad\qquad Y_k = D_k C_k F_k X + E_k$$

where
$\qquad F_k$ is a $L^2 \times L^2$ matrix representing a geometric warp
$\qquad C_k$ is a $L^2 \times L^2$ matrix representing blur in the degradation process
$\qquad D_k$ is a $S_k^2 \times L^2$ matrix representing decimation by the integral sampling
$\qquad E_k$ is a $S_k^2 \times S_k^2$ matrix representing additional noise

Using such notation makes it possible to express one of the basic methods of space domain SR known as Iterative Back-Projection [13] (IBP). To find X^* such that

$$(3.2) \qquad X^* = \underset{X}{ArgMin}\Big(\sum_{k=1}^{n} \|D_k C_k F_k X - Y_k\|\Big),$$

under L_2 norm we will iteratively apply formula (3.3) to some initial approximation of the high-resolution image. This equals to steepest descent solution of (3.2) [13].

$$(3.3) \qquad X_{n+1} = X_n - \beta\Big[\frac{1}{N}\sum_{k=1}^{N}(F_k^T C_k^T D_k^T(D_k C_k F_k X_n - Y_k)\Big]$$

Several modifications of the basic IBP algorithm were proposed (for example in works of Zomet [15] and Farsiu [5]), aimed to improve robustness of the algorithm to noise and errors in image registration. Zomet's algorithm can be expressed by formula (3.4), the basic difference is in using pixelwise median of error images instead of averaging the errors.

$$(3.4) \qquad X_{n+1} = X_n - \beta.median\Big\{\big(F_k^T C_k^T D_k^T(D_k C_k F_k X - Y_k)\big)\Big\}_{k=1}^{N}$$

Similar approach is taken in the work of Farsiu [5], the author states that optimization expressed by equation (3.2) under L_1 norm yields following iterative equation:

$$(3.5) \qquad X_{n+1} = X_n - \beta\Big[\frac{1}{N}\sum_{k=1}^{N}(F_k^T C_k^T D_k^T sign(D_k C_k F_k X_n - Y_k)\Big]$$

3.1. Smoothness prior. It is shown in [1] that Super resolution is generally an ill-posed inverse problem. This means that there are many almost optimal solutions of equation (3.2), some of which may be for example very noisy. It is therefore useful to include some further prior knowledge about the original image, which will help the algorithm to choose the most suitable solution of equation (3.2).

One of things we usually know about the input image is that it is not very noisy. Including such assumption to the algorithm is called smoothness prior, authors of [5] suggest using optimization equation (3.6) to gain iterative formula (3.7). Constants P, α and λ are introduced as tweaking constants that influence the nature of smoothness assumed for the image, S_x^n and S_y^n represent matrices shifting image by n pixels in direction of x resp. y axis. The additional term in (3.6) therefore represents a measure of similarity of negbouring pixels, while additional term in (3.7) enforces this similarity.

$$(3.6) \quad X^* = \underset{X}{ArgMin}\Big(\sum_{k=1}^{n}\|D_kC_kF_kX - Y_k\| + \lambda\sum_{r=1}^{P}\sum_{s=1}^{P}\alpha^{s+r}\|X - S_x^rS_y^s\|\Big)$$

$$(3.7) \quad \begin{aligned} X_{n+1} = \quad & X_n - \beta\Big[\tfrac{1}{N}\sum_{k=1}^{N}(F_k^TC_k^TD_k^T sign(D_kC_kF_kX_n - Y_k) \\ & +\lambda\sum_{r=1}^{P}\sum_{s=1}^{P}\alpha^{s+r}[I - S_x^{-r}S_y^{-s}]sign(X - S_x^rS_y^sX)\Big] \end{aligned}$$

4. Enhanced algorithm derivation. In previous sections we have shown that an image captured by common digital still camera undergoes a complicated pre-processing that includes demosaicking, white-balancing and value scaling. We have also shown that there are powerful algorithms capable of restoring a high resolution image from multiple degraded and slightly shifted images. We will now combine the knowledge acquired and propose an improvement of the algorithms.

Our basic idea is that not all data contained within final image file are directly related to the original scene. This is caused by the nature of demosaicking, which represents interpolation of measured data. We can improve the results by removing such interpolated data from the input.

Removal of interpolated data can be done by two approaches, both of which assume knowledge of concrete color filter array layout. First possibility is to use only those R, G and B values from the image file that were actually measured. In this case we will be using values altered by white-balancing, value scaling and in some cases also by demosaicking (some algorithms alter even the measured values).

The other possibility is to use a camera capable of producing RAW data file. Such file only contains values measured by the CCD element, not altered by either of demosaicking, white balancing or value scaling.

Any of the algorithms presented for the space domain SR can be used in almost unchanged manner. Each input image will only require a boolean mask that will tell the algorithm whether or not to take such pixel into account. Computation of the back-projected error can remain unaltered, only using respectively lower number of input values.

It is possible that there will be areas of the image that were not measured by any of input images. This would lead the algorithm to keeping the values of the first approximation at these spots throughout the whole computation, which is misleading. A better solution is offered by incorporating the smoothness assumption presented

above into the algorithm. Doing so will effectively lead to interpolation of the current approximation at the spots not measured by either of input images.

Rewriting equations (3.3), (3.4) and (3.5) we get improved formulas for basic IBP (3.3), Zomet SR (4.2) and Farsiu SR (4.3).

$$(4.1) \qquad X_{n+1} = X_n - \beta[Q \otimes \sum_{k=1}^{N}(F_k^T C_k^T D_k^T (M \otimes (D_k C_k F_k X_n - Y_k)))]$$

$$(4.2) \qquad X_{n+1} = X_n - \beta[median(F_k^T C_k^T D_k^T (M \otimes (D_k C_k F_k X_n - Y_k)))_{k=1}^{N}]$$

$$(4.3) \qquad X_{n+1} = X_n - \beta[Q \otimes \sum_{k=1}^{N} sign(F_k^T C_k^T D_k^T (M \otimes (D_k C_k F_k X_n - Y_k)))]$$

where the matrix M of size $S \times S$ represents a binary mask, i.e. contains ones in positions of measured values and zeros in positions containing values obtained by demosaicking (or containing no value in the case of RAW data file). The matrix Q of size $L \times L$ contains value $1/n$ where n is the number of measurements for each hi-res pixel, or zero at positions where no measurement was performed. The $\otimes$ operator represents multiplication of matrix elements at corresponding positions. Note that the median in (4.2) is computed only from non-zero values of the argument.

The proposed enhancement of algorithms is easy to implement and if implemented carefully, it may even decrease running time of the SR process. Generally the enhancement doesn't change the computational complexity of SR algorithms.

The algorithm may require additional memory for the representation of binary mask, but for the most usual case of Bayer array mask layout it is possible to determine the mask value according to position in the picture without using any extra memory. Therefore the only additional memory is required to store number of measurements for each hi-res pixel.

We have tested the algorithm for the case of simple translational registration, but our enhancement is generally independent on used registration algorithm.

5. Experiments. We have carried out a series of tests in order to compare the SR methods mentioned above and to evaluate the improvement gained by incorporating the proposed binary mask. We have tested accuracy of all the methods mentioned above on images simulating the degradation by digital camera. We have also performed tests with real images, but no measure of accuracy can be given for such tests because the original image is in such case not available for comparison with the SR results.

First, we have tested accuracy of the original IBP, Farsiu and Zomet methods with variable size of the β parameter. We have used one input image for all the tests and we have obtained optimal values of the iteration step β. As a measure of accuracy a pixelwise mean squared error (MSE) was used, comparing the original image to the ones received from the SR algorithms. The dependency of MSE on iteration step size is shown in figures 4.1, 4.2 and 4.3.

These tests show that the Zomet robust method gives best results. It also showed best robustness to noise in input data and therefore we have chosen this method for testing of our modification.

In the second presented experiment, we have only used the Zomet Robust method modified according to equation (4.2). We have used one set of 20 simulated input
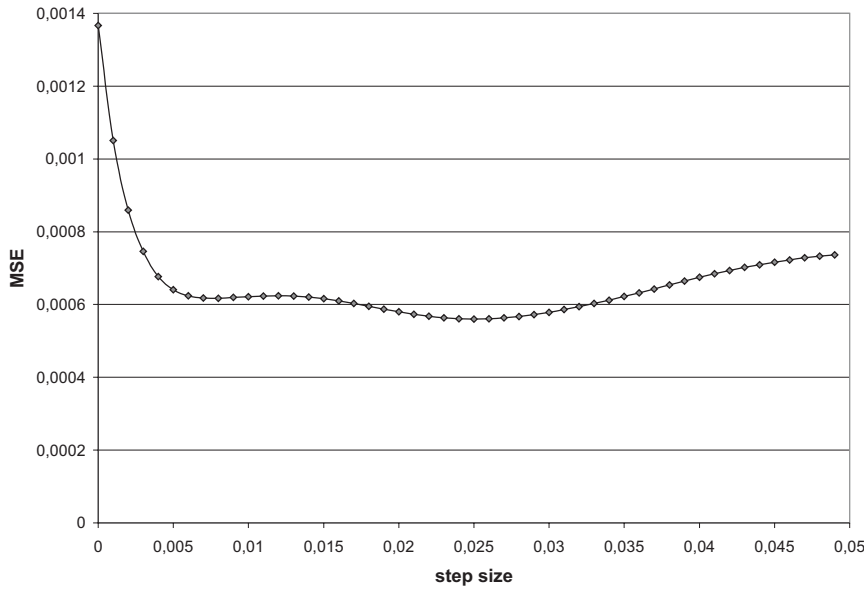
FIGURE 4.1. *Accuracy of original IBP method*

images. In contrast to the previous experiment we have included a simulation of demosaicking of green color channel, i.e. we have removed one half of the samples according to their position in Bayer array and we have recomputed the missing values using the Cok's algorithm [2] explained above.

Our aim was to compare two situations: in the first case, we have used a mask representing full sampling, i.e. containing value 1 at all positions. This is equal to using all data including the demosaicked values and the algorithm is equal to unaltered Zomet method. In the second case, we have used a binary mask representing positions of measured values of green channel in Bayer array. In such case only one half of the samples were taken into account during the processing, but on the other hand no interpolated data was used.

We have again measured a dependence of MSE on size of iteration step (β), the results are shown in figure 4.4.

We have performed similar tests for various input images, obtaining equivalent results. It is obvious from the graph that excluding the interpolated values from the input data may reduce the MSE by about 50% for the case of chosen optimal step size (we have used the value of 1.4 for the β parameter of Zomet algorithm). Such improvement is clearly visible from resulting images shown in figure 7.1 c) and figure 7.1 d).

We have tested our enhancement with real camera images. In this case, we cannot provide any numbers showing the accuracy, because we have no original image to compare with. Figure 7.2 shows our results. It may seem that the improvement is not as clearly visible as it is for the simulated data. This is probably caused by noise that is present in the input images (a noise with standard deviation about 2-5% is usually present in images gained by commercially available digital still cameras).

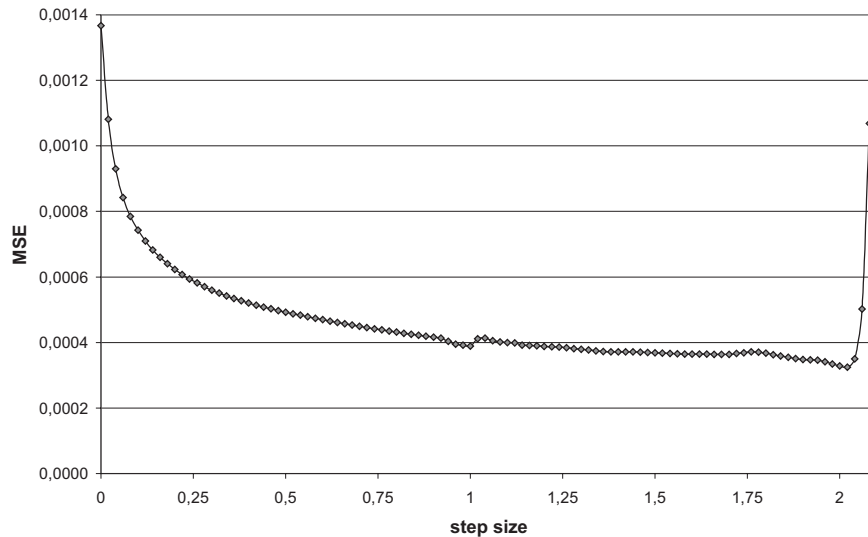The other reason may be the simple registration model we have used, as we

FIGURE 4.2. *Accuracy of Zomet robust method*

considered only translation and not rotation for image registration in our experiments.

The problem is that SR is extremely sensitive to exactness of relative position of image pixels, an error in positioning of the pixel of more than one half of hi-res pixel size leads to completely wrong results. It is easy to see that when the input is an image of size 200×200 pixels and we want to increase the resolution $3\times$, then the rotation should never exceed $0.04°$. This is of course hard to achieve in practice and therefore for a real application the rotation must be considered.

6. Conclusions. We have implemented and tested SR algorithms suitable for enhancing resolution of images gained by usual digital still camera. We have explored preprocessing that is performed on the measured data and implemented a simulation of such preprocessing.

We have proposed an enhancement of spatial SR methods that is applicable on all three implemented SR algorithms, which exploits the knowledge about preprocessing that is performed within the camera. We have tested this enhancement and presented results showing that it may reduce MSE by almost 50%.

7. Future work. We would like to consider an advanced registration algorithms including rotation estimation of input images. We hope that such improvement will help to bring SR methods closer to practice and enable an easy SR processing of images taken by digital camera held in hand.

REFERENCES

[1] S. BAKER, T. KANADE, *Limits on Super-Resolution and How to Break Them*, Proc.CVPR00, 2000
[2] D. R. COK, *Signal Processing method and apparatus for producing interpolated chrominance values in a sampled color image signal*, U.S. Patent No. 4,642,678 (1987)
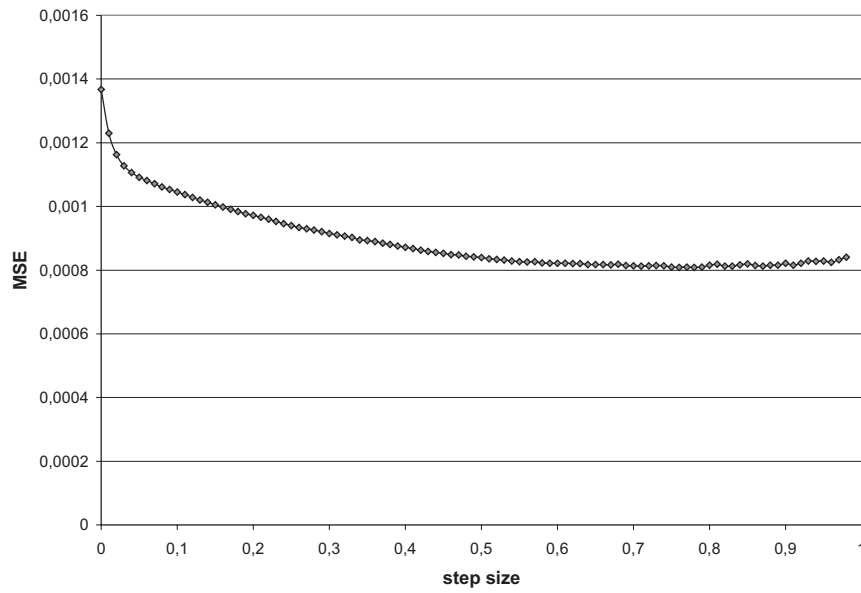
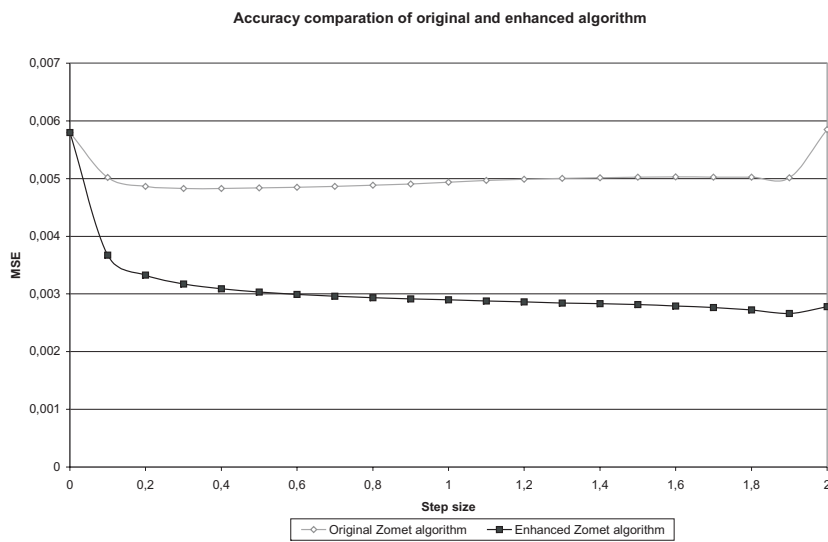FIGURE 4.3. *Accuracy of Farsiu robust method*

FIGURE 4.4. *Accuracy comparison between original end enhanced Zomet method*

[3] M. ELAD, A. FEUER, *Restoration of a Single Superresolution Image from Several Blurred, Noisy, and Undersampled Measured Images*, IEEE Transactions on Image Processing, Vol. 6, No. 12, December 1997

[4] M. ELAD, Y. HEL-OR, *A Fast Super-Resolution Reconstruction Algorithm for Pure Translation Motion and Common Space-Invariant Blur*, IEEE Transactions on Image Processing, Vol. 10, No. 8, pp. 1187-1193, August 2001.

a) original image

b) degraded image

c) result of Zomet method

d) result of enhanced
Zomet method

e) enlarged result of Zomet method

f) enlarged result of enhanced
Zomet method

FIGURE 7.1. *Accuracy testing images*

[5] S. FARSIU, D. ROBINSON, M. ELAD, P. MILANFAR, *Fast and Robust Super-Resolution*, Proceedings of ICIP 2003

[6] W. T. FREEMAN, *Median filter for reconstructing missing color samples* U.S. Patent No. 4,724,395 (1988)

[7] J. F. HAMILTON, J. E. ADAMS, *Adaptive color plane interpolation in single sensor color electronic camera*, U.S. Patent No. 5,629,734 (1997)

[8] R. C. HARDIE, K. J. BARNARD, E. E. ARMSTRONG, *Joint MAP Registration and High-Resolution Image Estimation Using a Sequence of Undersampled Images*, IEEE Transactions on Image Processing, Vol. 6, No. 12, December 1997

[9] T. HUANG, R. TSAI, *Multi-frame image restoration and registration*, Advances in Computer Vision and Image Processing, volume 1, pages 317-339. JAI Press Inc., 1984
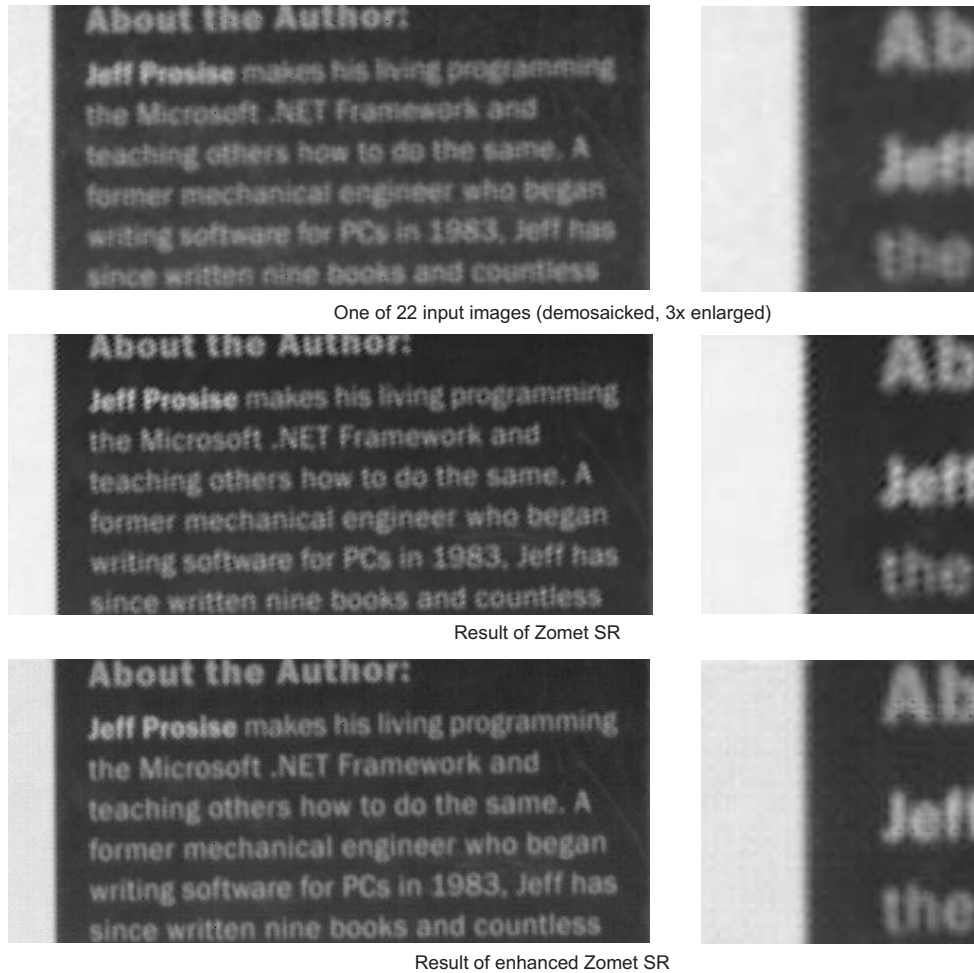
One of 22 input images (demosaicked, 3x enlarged)

Result of Zomet SR

Result of enhanced Zomet SR

FIGURE 7.2. *Real images testing results*

[10] R. KIMMEL, *Demosaicking: image reconstruction from CCD samples*, Proc. Trans. Image Processing, vol. 8, pp. 1221-1228, 1999

[11] C. A. LAROCHE, M. A. PRESCOTT, *Apparatus and method for adaptively interpolating a full color image utilizing chrominance gradients*, U.S. Patent No. 5,373,322 (1994)

[12] A. LUKIN, D. KUBASOV, *An Improved Demosaicking Algorithm*, to appear in proceedings of Graphicon 2004

[13] S. PELEG, D. KEREN, L. SCHWEITZER, *Improving image resolution using subpixel motion*, Pattern Recognition Letter, vol. 5, pp. 223-226, March 1987

[14] R. RAMANATH, W.E. SNYDER, G.L. BILBRO, AND W.A. SANDER, *Demosaicking methods for bayer color arrays*, Journal of Electronic Imaging, vol. 11, no. 3, pp. 306-315, Jul. 2002.

[15] A. ZOMET, A. RAV-ACHA, S. PELEG, *Robust Super-Resolution*, Proceedings of the Int. Conf. on Computer Vision and Pattern Recognition (CVPR), vol. 1, pp. 645-650, Dec. 2001