A NEW TECHNIQUE OF NON-ITERATIVE SUPER-RESOLUTION WITHOUT BOUNDARY DISTORTION

Izumi Ito

Tokyo Institute of Technology
Graduate School of
Decision Science and Technology
2-12-1, Ookayama, Meguro-ku, Tokyo, Japan

Hitoshi Kiya

Tokyo Metropolitan University
Graduate School of System Design
6-6, Asahigaoka, Hino-shi, Tokyo, Japan

ABSTRACT

We propose a new technique of non-iterative super-resolution image reconstruction. A closed form solution to reconstruction is derived in the discrete cosine transform (DCT) domain from a MAP-based cost functional. An average image is used in order to avoid iterative operations. Symmetric convolution with appropriate types of DCT suppresses boundary distortion. Experimental results demonstrate the effectiveness of the proposed technique.

Index Terms— Super resolution, MAP estimator, discrete cosine transform, restoration

1. INTRODUCTION

Super-resolution (SR) image reconstruction is to create a high-resolution (HR) image from multiple low-resolution (LR) images captured from the same scene. A variety of approaches to SR image reconstruction exist [1].

Maximum a posteriori (MAP) estimator is one of the approaches, which provides a flexible and convenient way to use a priori knowledge. Generally, MAP-based SR requires high computational load in the spatial domain [2]. One way to the acceleration is to use an average image, into which the values of multiple registered LR images are integrated [3][4]. Tanaka and Okutomi showed a MAP-based SR using the average image with the weight can be calculated fast with high performance [3]. Kudoh et al. tried to restore the average image without iterative operations [4]. However, boundary distortion is occurred on the reconstructed image.

Symmetric convolution is a convolution between symmetrically extended sequences using discrete cosine transform (DCT) or discrete sine transform (DST) [5]. The authors in the present paper showed that a linear convolution can be calculated using symmetric convolution in the DCT domain with lower computational complexity than using discrete Fourier transform (DFT) [6].

In the present paper, we propose a non-iterative SR image reconstruction technique. We use the average image for the closed form solution. The MAP-based cost functional is defined using symmetric convolution in the DCT domain so that the reconstructed image holds the smoothness around the boundary. We discuss the appropriate type of DCT to be applied to each sequence for symmetric convolution. Experimental results show the effectiveness of the proposed technique.

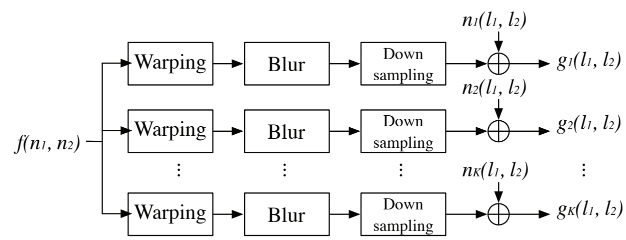


Fig. 1. Image acquisition model. Multiple LR images $g_i(l_1, l_2)$ for $i = 1, \dots, K$ are captured through a degradation process from an HR image $f(n_1, n_2)$.

2. PRELIMINARY

2.1. Image acquisition model

Observed LR images are degraded by warping, blurring, down-sampling an HR image, and they are corrupted by additive noise as shown in Fig. 1. The i-th LR image $g_i(l_1, l_2)$ is expressed in matrix-vector form as

$$\overrightarrow{g_i} = [D][B_i][W_i]\overrightarrow{f} + \overrightarrow{n_i} \tag{1}$$

where $\overrightarrow{g_i}$, $\overrightarrow{f}$, $\overrightarrow{n_i}$ denote the lexicographically ordered i-th LR image vector, HR image vector, i-th noise vector, respectively, and $[D]$, $[B_i]$, and $[W_i]$ represent a decimation matrix, the i-th blur matrix, and the i-th warp matrix.

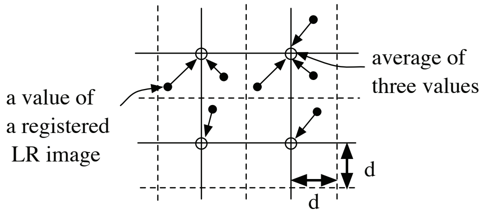


Fig. 2. Average image with HR grid. A black circle expresses a value of a registered LR image, and a white circle represents the average of the values of multiple registered LR images.

2.2. MAP-based SR and constrained least squeas

The cost functional of a MAP-based SR is defined as

$$E(f) = \sum_{i=1}^{K} ||[H_i]\overrightarrow{f} - \overrightarrow{g_i}||^2 + \lambda||[C]\overrightarrow{f}||^2 \qquad (2)$$

where $[H_i] = [D][B_i][W_i]$, λ denotes the regularization parameter, $||x||$ denotes the l_2 norm of x, and $[C]$ is a priori knowledge matrix. Commonly, a high-pass filter is used as a priori knowledge, which suggests that most images are smooth. Since the cost functional in (2) is convex and differentiable, a unique estimate can be found by iterative techniques such as steepest-descent algorithm.

2.3. Average image

An average image is introduced for fast SR algorithm [3][4].

The average image, $g_a(n_1, n_2)$, is defined as

$$g_a(n_1, n_2) = \frac{1}{w(n_1, n_2)} \sum_{i=1}^{K} \sum_{(t_1,t_2)\in D_{(n_1,n_2)}} \hat{g}_i(t_1, t_2) \qquad (3)$$

where $D_{(n_1,n_2)}$ represents a $2d \times 2d$ range centered on (n_1, n_2), d denotes a half sample of the HR image, $\hat{g}_i(t_1, t_2)$, $t_1, t_2 \in \mathbb{R}$, is the i-th registered LR image, and $w(n_1, n_2)$ is the number of elements in $D_{(n_1,n_2)}$ as shown in Fig. 2.

2.4. Symmetric convolution

Symmetric convolution yields a linear convolution of symmetrically extended sequences [5]. It is achieved using DCT or DST without augmenting the original sequences.

Symmetric convolution of $h(n)$ with $x(n)$ is defined as

$$y(n) = h(n) \circledast x(n) = (\epsilon_a[h(n)] * \epsilon_b[x(n)])\, R(n) \qquad (4)$$

where the operator '$\circledast$' denotes symmetric convolution, ϵ_a and ϵ_b denote the symmetric extension operators for inputs $h(n)$ and $x(n)$, respectively, the operator '$*$' denotes convolution, and $R(n)$ is a rectangular window that extracts representative samples.

Equation (4) can be calculated as

$$y(n) = T_c^{-1}[T_a[h(n)] \times T_b[x(n)]] \qquad (5)$$

where T_a and T_b are the corresponding DCT or DST for $h(n)$ and $x(n)$, respectively, the operator '$\times$' is element-by-element multiplication, and T_c^{-1} is the appropriate inverse transform that is uniquely determined from the combination of T_a and T_b. DCT and DST are subdivided and 40 distinct combinations of T_a and T_b with their inverse transforms are derived.

3. PROPOSED TECHNIQUE

We propose an SR image reconstruction technique.

The assumption is that the average image is available and that we know the information about a common point spread function (PSF) that is modeled as a low-pass filter and noise characteristics.

3.1. Cost functional and closed form solution

The cost functional of the proposed technique is defined using the average image and symmetric convolution in which only the reconstructed image is extended symmetrically not both inputs in order to suppress boundary distortion.

Let $f(n_1, n_2)$ be an HR image of size $N \times N$ that we desire to reconstruct. The cost functional $E(\mathbf{f})$ is defined as

$$\begin{aligned} E(\mathbf{f}) = ||p_q(n_1, n_2) \circledast f(n_1, n_2) - g_a(n_1, n_2)||^2 \\ + \alpha||c_q(n_1, n_2) \circledast f(n_1, n_2)||^2 \end{aligned} \qquad (6)$$

where α denotes the regularization parameter, $p_q(n_1, n_2)$ is a quarter of PSF $p(n_1, n_2)$ of size $L_p \times L_p$, $g_a(n_1, n_2)$ expresses the average image of the same size as $f(n_1, n_2)$ according to (3), and $c_q(n_1, n_2)$ is a quarter of a high-pass filter $c(n_1, n_2)$ of size $L_c \times L_c$.

The use of a quarter of a PSF and a high-pass filter provides the desirable effect of symmetric convolution. If a PSF of size 5×5 is

$$p(n_1, n_2) = \begin{bmatrix} Y & Y & Y & Y & Y \\ Y & X & X & X & Y \\ Y & X & C & X & Y \\ Y & X & X & X & Y \\ Y & Y & Y & Y & Y \end{bmatrix} \qquad (7)$$

then, the quarter of the PSF is defined as

$$p_q(n_1, n_2) = \begin{bmatrix} C & X & Y \\ X & X & Y \\ Y & Y & Y \end{bmatrix} \qquad (8)$$

where $C = p_q(0, 0)$ denotes the value at the center of PSF, and X and Y denote arbitrary values. $c_q(n_1, n_2)$ is obtained in the same way as $p_q(n_1, n_2)$ is in (8).

Table 1. Type of DCT for each sequence.

sequence ($N \times N$)	zero-padded $p_q(n_1, n_2)$	zero-padded $c_q(n_1, n_2)$	$g_a(n_1, n_2)$	$f(n_1, n_2)$
proposed 1	DCT-1	DCT-1	DCT-1	DCT-1
proposed 2	DCT-1	DCT-1	DCT-2	DCT-2
corresponding DCT coeff.	$P_C(k_1, k_2)$	$C_C(k_1, k_2)$	$G_C(k_1, k_2)$	$F_C(k_1, k_2)$

From (4) and (5), the cost functional in (6) can be expressed in the DCT domain as

$$E(F_C) = \sum_{k_1=0}^{N-1} \sum_{k_2=0}^{N-1} \{(P_C(k_1, k_2)F_C(k_1, k_2) - G_C(k_1, k_2))^2 + \alpha(C_C(k_1, k_2)F_C(k_1, k_2))^2\} \quad (9)$$

where $F_C(k_1, k_2)$, $G_C(k_1, k_2)$, $P_C(k_1, k_2)$, and $C_C(k_1, k_2)$ denote the corresponding DCT coefficients of $f(n_1, n_2)$, $g_a(n_1, n_2)$, $p_q(n_1, n_2)$, and $c_q(n_1, n_2)$, respectively. The details are discussed in the next subsection.

To find the solution that minimizes $E(F_C)$, the derivative of $E(F_C)$ must be

$$\partial E(F_C)/\partial F_C(k_1, k_2) = 0. \quad (10)$$

It yields for $k_1, k_2 = 0, 1, \ldots, N-1$,

$$F_C(k_1, k_2) = \frac{P_C(k_1,k_2)}{P_C(k_1,k_2)^2 + \alpha C_C(k_1,k_2)^2} G_C(k_1,k_2). \quad (11)$$

Thus, we can obtain the closed form solution. Therefore, applying the inverse DCT to $F_C(k_1, k_2)$ in (11), we can reconstruct an HR image without iterative operations.

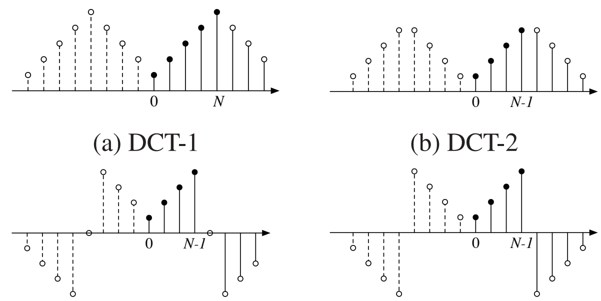
3.2. Appropriate type of DCT

Let us consider which type of DCT should be applied to each sequence.

DCT is the collective term of discrete transforms using cosine as bases. There are four distinct types of DCT according to underlying extension which relates to DFT as shown in Fig.3. Derivation of (5) is based on this relation.

From the aspect of PSF, $P_C(k_1, k_2)$ should be Type 1 DCT (DCT-1) coefficients. Accordingly, $P_C(k_1, k_2)$ is generated by applying DCT-1 to zero-padded $p_q(n_1, n_2)$ of size $N \times N$. In this way, the whole PSF is formed when zero-padded $p_q(n_1, n_2)$ is extended symmetrically.

In symmetric convolution, there is a constraint that any pair of types within a class can be convolved but not between classes [5]. Since a class consists of DCT-1 and DCT-2, the remaining sequences should be DCT-1 or DCT-2. In addition, DCT-2 cannot inherently generate high-pass filters. Therefore, $C_C(k_1, k_2)$ should be DCT-1 coefficients, which are generated by applying DCT-1 to zero-padded $c_q(n_1, n_2)$.



(a) DCT-1 (b) DCT-2 (c) DCT-3 (d) DCT-4

Fig. 3. Symmetric periodic sequence corresponding to four types of DCT. The DCT coefficients of a sequence expressed as black circles correspond to the DFT coefficients of the symmetrically extended sequence expressed as solid lines.

Under the above constraint, $G_C(k_1, k_2)$ should be DCT-1 or DCT-2 coefficients. When $G_C(k_1, k_2)$ is DCT-1, $F_C(k_1, k_2)$, the product of DCT-1 and DCT-1, is defined as DCT-1 by (5). Accordingly, the estimated HR image, $f(n_1, n_2)$, is obtained by applying the inverse DCT-1 to $F_C(k_1, k_2)$. When $G_C(k_1, k_2)$ is DCT-2, $F_C(k_1, k_2)$ is DCT-2. $f(n_1, n_2)$ is obtained from the inverse DCT-2 of $F_C(k_1, k_2)$. The type of DCT (see the Appendix) applied to each sequence is summarized in Table 1.

4. SIMULATIONS

The proposed technique is performed to evaluate the reconstructed images.

A total of 32 frames of LR images of size 128×128 were shifted by one of the shifts $\{(0, 0), (0, 0.5), (0.5, 0), (0.5, 0.5)\}$, blurred with a common PSF of a 9×9 Gaussian kernel of zero mean and a variance of 1.3, and down-sampled by a factor of two in both the horizontal and vertical directions. In addition, Gaussian noise of zero mean and the standard deviation of 0.01 was added to each LR image.

Figure 4 shows HR images reconstructed by bicubic interpolation of one of the LR images, a fast Fourier transform (FFT)-based technique, and the proposed technique in Table 1, in which the regularization parameter was set to 0.0003 and a four-point neighborhood Laplacian was used as

(a) Bicubic interpolation
PSNR = 24.00[dB]

(b) FFT-based
PSNR = 24.74[dB]

(c) proposed 1
PSNR = 32.32[dB]

(d) proposed 2
PSNR = 32.42[dB]

Fig. 4. Reconstructed HR image (256×256 Barbara).

$c(n_1, n_2)$. In the FFT-based technique, the HR image was restored by constrained least squares method [7] as

$$F(k_1, k_2) = \frac{P^*(k_1, k_2)}{(|P(k_1 k_2)|^2 + \alpha |C(k_1, k_2)|^2)} G_a(k_1, k_2) \quad (12)$$

where $F(k_1, k_2)$, $P(k_1, k_2)$, $G_a(k_1, k_2)$, and $C(k_1, k_2)$ are the corresponding DFT coefficients of $f(n_1, n_2)$, $p(n_1, n_2)$, $g_a(n_1, n_2)$, and $c(n_1, n_2)$, respectively, and $P^*(k_1, k_2)$ is the complex conjugate of $P(k_1, k_2)$. Low PSNR in FFT-based technique is caused by the ringing around the image boundary, although the reconstructed image is far improved compared to the bicubic-interpolated image. Conversely, an edge-preserving image is obtained without the boundary distortion in the proposed technique. Table 2 shows the PSNR of other reconstructed images. We can confirm that the proposed technique is effective for SR image reconstruction.

5. CONCLUSION

We have proposed a non-iterative SR image reconstruction technique. The use of the average image reduces the problem of SR image reconstruction to that of image restoration. We use symmetric convolution in order to suppress the boundary distortion. We have considered the appropriate type of DCT to be applied to each sequence for symmetric convolution. The use of DCT is advantageous in that symmetric convolution can be calculated simply and quickly without extending the image and that DCT has fast algorithms with real numbers. We have shown the effectiveness of the proposed technique.

Table 2. PSNR [dB] of HR image for the original image.

256×256 image	bicubic	FFT-based	**proposed 1**	**proposed 2**
couple	28.13	25.80	**33.06**	**33.18**
boat	25.18	23.33	**31.51**	**31.51**
airplane	27.81	25.47	**34.31**	**34.28**
sailboat	28.09	24.11	**32.73**	**32.79**
Lena	26.94	23.98	**32.92**	**32.95**
baboon	25.66	25.62	**29.37**	**29.33**

6. APPENDIX

The matrix form of DCT-1, DCT-2, and their inverse denoted as [C1], [C2], and $[\cdot]^{-1}$, respectively, is given below [5].

$$[\mathrm{C1}] = 2c_n \cos\left(\frac{\pi k n}{N}\right), k, n = 0, 1, \ldots, N$$

$$[\mathrm{C1}]^{-1} = \frac{1}{2N}[C1]$$

$$[\mathrm{C2}] = 2\cos\left(\frac{\pi k(n + \frac{1}{2})}{N}\right), k, n = 0, 1, \ldots, N-1$$

$$[\mathrm{C2}]^{-1} = 2c_k \cos\left(\frac{\pi k(n + \frac{1}{2})}{N}\right), k, n = 0, 1, \ldots, N-1$$

$$c_p = \begin{cases} 1/2, & p = 0 \quad \text{or} \quad N \\ 1, & p = 1, 2, \ldots, N-1 \end{cases}.$$

7. REFERENCES

[1] S. C. Park, M. K. Park, and M. G. Kang, “Super-resolution image reconstruction: A technical overview, ” IEEE Signal Processing Magazine, vol. 20, no. 3, pp. 21-36, May 2003.

[2] D. Zhang, H. F. Li, and M. H. Du, “ Fast MAP-based multi-frame super-resolution image reconstruction, ” Image and Vision Computing, vol. 23, no. 7, pp. 671-679, Jul. 2005.

[3] M. Tanaka, and M. Okutomi, ”A fast MAP-based super-resolution algorithm for general motion,” Proc. of SPIE-IS & T Electronic Imaging, SPIE vol.6065, 60651B, pp.1-12, 2006.

[4] S. Kudoh, T. Nakachi, and N. Hamada, “ Fast and Enhanced Super Resolution for DCT-Compressed Images, ” Journal of Signal Processing, vol. 14, no. 6, pp. 451-458, 2010.

[5] S. A. Martucci, “Symmetric convolution and the discrete sine and cosine transforms,” IEEE Trans. Signal Process., vol. 42, no. 5, pp. 1038-1051, May 1994.

[6] I. Ito and H. Kiya, “A computiong method for linear convolution in the DCT domain,” 19th European Signal Processing Conference (EUSIPCO 2011), pp. 323-327, Aug. 2011.

[7] R. C. Gonzalez and R. E. Woods, *Digital image processing, 3rd eddition*, Prentice Hall, 2007