

AN ADAPTIVE SUPER-RESOLUTION OF VIDEOS WITH NOISE INFORMATION ON CAMERA SYSTEMS

Toshiyuki Ono, Hiroshi Hasegawa, Isao Yamada, and Kohichi Sakaniwa

Dept. of Communications & Integrated Systems, Tokyo Institute of Technology
2-12-1, Ookayama, Meguroku, Tokyo, 152-8552, Japan.
E-mails: {tono, hasegawa, isao, sakaniwa}@comm.ss.titech.ac.jp

ABSTRACT

We present a novel adaptive super-resolution of videos based on an embedded constraint version of *Adaptive projected subgradient method* [Yamada & Ogura 2004]. The super-resolution image recovery problem is formulated as an estimation of linear time-varying systems, which is a modified version of [Elad & Feuer 1999]. Our method efficiently improves accuracy of estimation by simple iterative operations which can be processed on parallel systems. Robustness to additive noise as well as inaccurate estimation of degradation parameters, is realized by incorporating stochastic information of the noise.

1. INTRODUCTION

Super-resolution recovers a high-resolution image from several degraded low-resolution images which potentially have overlap. Super-resolution system has been demanded in various areas from medical/satellite imaging [1, 2] to HDTV [3, 4]. Elad and Feuer formulated a super-resolution of videos as an adaptive system identification problem [5]. They applied LMS and simplified RLS to the problem.

Adaptive projected subgradient method developed [6] recently as a time-varying generalization of set-theoretic signal recovery schemes [3, 7, 8]. This method includes NLMS, APA and Adaptive parallel subgradient projection algorithm [9] as special examples and free from instability caused by model mismatch unlike RLS. The NLMS, known as a variation of LMS, exhibits potentially faster convergence than LMS [10, P.437]. In the application to super-resolution problem, however, the speed of NLMS is obviously not sufficient. The adaptive filtering for super-resolution problem must process huge data, hence it must utilize multiple data efficiently as well as be realized with low computational complexity.

In this paper, we propose adaptive super-resolution based on an embedded constraint version of *Adaptive projected subgradient method* [6]. This scheme can use multiple data in parallel way. Moreover, by incorporating noise information set-theoretically, the speed of convergence as well as robustness to noise of the adaptive filtering can be greatly improved.

Specially for application to the adaptive super-resolution problem, we first define a model of relation between successive high-resolution images. This model gives more detailed information on noise. Secondly, *stochastic property set*, a set includes the original image in high probability, is defined by using statistical information of the noise. To avoid undesirable estimation, additional constraints are introduced, some of which restrict abrupt intensity change between frames. Finally we present the proposed schemes for super-resolution problem. Numerical example shows excellent performance of the proposed method.

2. PRELIMINARIES

Let \mathbb{R} and \mathbb{Z} be the set of all real numbers and integers, respectively. For all vectors $\mathbf{u} := (u_1, \dots, u_P)$, $\mathbf{v} := (v_1, \dots, v_P)$ in a \mathcal{P} dimensional Euclidean space $\mathbb{R}^{\mathcal{P}}$, its inner product and induced norm are defined by $\langle \mathbf{u}, \mathbf{v} \rangle := \sum_{k=1}^{\mathcal{P}} u_k v_k$ and $\|\mathbf{u}\| := \sqrt{\langle \mathbf{u}, \mathbf{u} \rangle}$. A set $C \subset \mathbb{R}^{\mathcal{P}}$ is convex provided that $\forall \mathbf{u}, \mathbf{v} \in C, \forall \nu \in (0, 1), \nu \mathbf{u} + (1 - \nu) \mathbf{v} \in C$. Given a nonempty closed convex set $C \subset \mathbb{R}^{\mathcal{P}}$, a convex projection $P_C : \mathbb{R}^{\mathcal{P}} \rightarrow C$ maps $\mathbf{u} \in \mathbb{R}^{\mathcal{P}}$ to the unique vector $P_C(\mathbf{u}) \in C$ such that $d(\mathbf{u}, C) := \min_{\mathbf{v} \in C} \|\mathbf{u} - \mathbf{v}\| = \|\mathbf{u} - P_C(\mathbf{u})\|$. For $\mathbf{u} := (u_1, \dots, u_P) \in \mathbb{R}^{\mathcal{P}}$, its m th component u_m is equivalently denoted by $\mathbf{u}(m)$. Let M, N and L be positive integers. Suppose that we have sequences $(\mathbf{x}_k)_{k \in \mathbb{Z}} \subset \mathbb{R}^{L^2 MN}$ and $(\mathbf{y}_k)_{k \in \mathbb{Z}} \subset \mathbb{R}^{MN}$ derived through lexicographically reordering of pixels of $LM \times LN$ high-resolution images and related $M \times N$ low-resolution images, respectively. Each low-resolution image is assumed to be generated by

$$\mathbf{y}_k = D H_k \mathbf{x}_k + \mathbf{n}_k \quad (k \in \mathbb{Z}), \quad (1)$$

where $H_k \in \mathbb{R}^{L^2 MN \times L^2 MN}$ denotes a degradation such as blur, $D \in \mathbb{R}^{MN \times L^2 MN}$ changes resolution by averaging each $L \times L$ region, and $\mathbf{n}_k \in \mathbb{R}^{L^2 MN}$ denotes additive noise.

In this paper, we assume that \mathbf{x}_k and $\mathbf{x}_\ell(k, \ell \in \mathbb{Z})$ are related by

$$\mathbf{x}_\ell = A_{(\ell, k)} \mathbf{x}_k + \mathbf{c}_{(\ell, k)} + \mathbf{e}_{(\ell, k)} \quad (2)$$

where $A_{(\ell,k)} \in \mathbb{R}^{L^2 MN \times L^2 MN}$ represents the geometric warp, $\mathbf{c}_{(\ell,k)} \in \mathbb{R}^{L^2 MN}$ stands for pixels of \mathbf{x}_ℓ having no relation to \mathbf{x}_k , and $\mathbf{e}_{(\ell,k)}$ is an error due to inaccurate registration. In digital imaging systems, CCDs are used for converting observed images to digital signals. Since noise on CCD is often assumed to be Gaussian [11], we will assume \mathbf{n}_k is Gaussian with zero mean and the variance is assumed to be known through testing the camera systems in advance. Moreover, $\mathbf{e}_{(\ell,k)}$ can be approximated by Gaussian noise [12]. Then, by combining (1) and (2), we obtain

$$\mathbf{d}_{(\ell,k)} = B_{(\ell,k)} \mathbf{x}_k + \mathbf{n}_{(\ell,k)} \quad (3)$$

where

$$\begin{aligned} \mathbf{d}_{(\ell,k)} &:= \mathbf{y}_\ell - DH_\ell \mathbf{c}_{(\ell,k)}, \\ B_{(\ell,k)} &:= DH_\ell A_{(\ell,k)} \end{aligned}$$

and a Gaussian noise

$$\mathbf{n}_{(\ell,k)} := DH_\ell \mathbf{e}_{(\ell,k)} + \mathbf{n}_k.$$

We need further assumption that the covariance $E[\mathbf{n}_{(\ell,k)}(m)\mathbf{n}_{(\ell',k)}(m')](\ell, \ell' \in \mathbb{Z}, m, m' \in \{1, \dots, MN\})$ are derived by using a priori information of camera systems and estimated registration error. Eq.(3) can be equivalently expressed by

$$\mathbf{d}_{(\ell,k)}(m) = B_{(\ell,k)}(m) \mathbf{x}_k + \mathbf{n}_{(\ell,k)}(m) \quad (m \in \{1, \dots, MN\}), \quad (4)$$

where $B_{(\ell,k)}(m)$ is m th row vector of $B_{(\ell,k)}$. In addition, some natural and simple requirements would improve the result. For example, additive noise being zero mean, the effect can be simply reduced by averaging over the image. Namely, we have

$$\frac{1}{MN} \sum_{m=1}^{MN} \mathbf{y}_k(m) = \frac{1}{L^2 MN} \sum_{m=1}^{L^2 MN} \mathbf{x}_k(m). \quad (5)$$

The above constraint would suppress undesirable deviation of intensities between frames. Moreover, each component of \mathbf{x} must be nonnegative and would be bounded

$$\mathbf{x} \in [\mathcal{B}_L, \mathcal{B}_U]^{L^2 MN} \quad (6)$$

where $0 \leq \mathcal{B}_L < \mathcal{B}_U < \infty$ are lower and upper bounds of intensity. Then our super-resolution problem is nothing but the following system identification problem of a slowly time-varying system with unknown impulse response.

Problem 1 For each $k \in \mathbb{Z}$, estimate an impulse response \mathbf{x}_k of a slowly time-varying system, subject to (5) and (6), by using its observable inputs $(B_{(\ell,k)}(m))_{\ell \in \{k, \dots, k-K+1\}, m \in \{1, \dots, MN\}}$, outputs $(\mathbf{d}_{(\ell,k)}(m))_{\ell \in \{k, \dots, k-K+1\}, m \in \{1, \dots, MN\}}$ and the covariance of noise $\mathbf{n}_{(\ell,k)}$, where K is the number of low-resolution images used for recovery.

Remark 1 Eq.(4) is similar to an equation that Elad and Feuer employed for their super-resolution. In their paper, $\mathbf{c}_{(\ell,k)} + \mathbf{e}_{(\ell,k)}$ in (2) is regarded as a noise all together. We modified it to (2) so that the variance of $\mathbf{n}_{(\ell,k)}$ is reduced.

3. PROPOSED ADAPTIVE SUPER-RESOLUTION WITH NOISE INFORMATION

For every pair of index sets $I(p) := \{i_1^{(p)}, \dots, i_q^{(p)}\} \subset \{k, k-1, \dots, k-K+1\}$, $J(p) := \{j_1^{(p)}, \dots, j_q^{(p)}\} \subset \{1, \dots, MN\}$, by stacking $q(\geq 1)$ equations of (4), we have

$$\begin{aligned} \boldsymbol{\nu}_{(k,p)} &:= \begin{bmatrix} \mathbf{n}_{(i_1^{(p)},k)}(j_1^{(p)}) \\ \vdots \\ \mathbf{n}_{(i_q^{(p)},k)}(j_q^{(p)}) \end{bmatrix} \\ &= \begin{bmatrix} B_{(i_1^{(p)},k)}(j_1^{(p)}) \\ \vdots \\ B_{(i_q^{(p)},k)}(j_q^{(p)}) \end{bmatrix} \mathbf{x}_k - \begin{bmatrix} \mathbf{d}_{(i_1^{(p)},k)}(j_1^{(p)}) \\ \vdots \\ \mathbf{d}_{(i_q^{(p)},k)}(j_q^{(p)}) \end{bmatrix} \\ &=: U_{(k,p)} \mathbf{x}_k - \mathbf{v}_{(k,p)} \end{aligned}$$

and $Q_{(k,p)} := E[\boldsymbol{\nu}_{(k,p)} \boldsymbol{\nu}_{(k,p)}^T]^{-\frac{1}{2}}$. Then $Q_{(k,p)} \boldsymbol{\nu}_{(k,p)}$ is an i.i.d. process. Note that $Q_{(k,p)}$ is given a priori because of the assumption on covariance. We define a closed convex set $C_{(k,p)}$, called *stochastic property set*, which contains \mathbf{x}_k with high probability, by

$$C_{(k,p)} := \left\{ \mathbf{x} \in \mathbb{R}^{L^2 MN} \mid \|Q_{(k,p)}(U_{(k,p)} \mathbf{x} - \mathbf{v}_{(k,p)})\|^2 \leq \rho_{(k,p)} \right\} \quad (p \in \mathbb{Z}) \quad (7)$$

where $\rho_{(k,p)} \geq 0$ (See Eq.(11) of [9] for detail). The threshold $\rho_{(k,p)}$ controls tradeoff between the probability and tightness of $C_{(k,p)}$. Some examples of $\rho_{(k,p)}$ are given in [9]. Moreover, noise averaging constraint in (5) and bounded intensity constraint in (6) respectively define constraint sets

$$V_{NA} := \{\mathbf{x} \in \mathbb{R}^{L^2 MN} \mid \mathbf{x} \text{ satisfies (5)}\}$$

$$C_{BI} := [\mathcal{B}_L, \mathcal{B}_U]^{L^2 MN}.$$

Our basic idea is this: *since the original image would be included in each $C_{(k,p)}$, V_{NA} , and C_{BI} , we can approach to a solution of Problem 1 by iteratively computing the convex projections onto them.*

Since V_{NA} is a hyperplane and C_{BI} is a hypercubic, convex projections onto them are easy to compute. Unfortunately, the convex projection onto $C_{(k,p)}$ is hard to compute. Thus we employ the next closed halfspace to approximate to $C_{(k,p)}$

$$H_{(k,p)}^-(\mathbf{x}_k^{(n)}) := \left\{ \mathbf{x} \in \mathbb{R}^{L^2 MN} \mid (\mathbf{x} - \mathbf{x}_k^{(n)})^T \nabla f_{(k,p)}(\mathbf{x}_k^{(n)}) + f_{(k,p)}(\mathbf{x}_k^{(n)}) \leq 0 \right\},$$

where

$$f_{(k,p)}(\mathbf{x}) := \|Q_{(k,p)}(U_{(k,p)}\mathbf{x}_k - \mathbf{v}_{(k,p)})\|^2 - \rho_{(k,p)}$$

$$\nabla f_{(k,p)}(\mathbf{x}) = 2(Q_{(k,p)}U_{(k,p)})^T Q_{(k,p)}(U_{(k,p)}\mathbf{x}_k - \mathbf{v}_{(k,p)}).$$

This halfspace is called an *outer approximation* of $C_{(k,p)}$. Then the convex projection onto $C_{(k,p)}$ is efficiently approximated by the convex projection onto $H_{(k,p)}^-(\mathbf{x}_k^{(n)})$, which is nothing but an orthogonal projection onto a boundary of the above outer approximation

$$P_{H_{(k,p)}^-(\mathbf{x}_k^{(n)})}(\mathbf{x}_k^{(n)}) := \begin{cases} \mathbf{x}_k^{(n)} & \text{if } f_{(k,p)}(\mathbf{x}_k^{(n)}) \leq 0 \\ \mathbf{x}_k^{(n)} - \frac{f_{(k,p)}(\mathbf{x}_k^{(n)})}{\|\nabla f_{(k,p)}(\mathbf{x}_k^{(n)})\|^2} \nabla f_{(k,p)}(\mathbf{x}_k^{(n)}) & \text{otherwise} \end{cases}$$

when $\mathbf{x}_k^{(n)} \notin C_{(k,p)}$. Finally the next iterative method, based on an embedded constraint version of *Adaptive projected subgradient method* (See [6, Example 5]), resolves Problem 1 with small computational load for each steps and robustness to noise.

Algorithm 1 For each $k \in \mathbb{Z}$, let

$$\mathbf{x}_k^{(0)} := P_{V_{NA}}(A_{(k,k-1)}\hat{\mathbf{x}}_{k-1} + \hat{\mathbf{c}}_{(k,k-1)})$$

where $\hat{\mathbf{x}}_{k-1}$ and $\hat{\mathbf{c}}_{(k,k-1)}$ are estimates of \mathbf{x}_{k-1} and $\mathbf{c}_{(k,k-1)}$, respectively. Generate a sequence $(\mathbf{x}_k^{(n)})_{n \geq 0} \subset \mathbb{R}^{L^2 MN}$ by the following equation

$$\mathbf{x}_k^{(n+1)} := \begin{cases} \mathbf{x}_k^{(n)} & \text{if } \Theta'_n(\mathbf{x}_k^{(n)}) \in \mathcal{M}_k^\perp \\ \Theta_n(\mathbf{x}_k^{(n)}) & \\ \mathbf{x}_k^{(n)} - \lambda_n \frac{P_{\mathcal{M}_k}(\Theta'_n(\mathbf{x}_k^{(n)}))}{\|P_{\mathcal{M}_k}(\Theta'_n(\mathbf{x}_k^{(n)}))\|^2} & \text{otherwise,} \end{cases}$$

where $\lambda_n \in [0, 2]$, a linear subspace parallel to V_{NA}

$$\mathcal{M}_k := \left\{ \mathbf{x} \in \mathbb{R}^{L^2 MN} \left| \sum_{m=1}^{L^2 MN} \mathbf{x}(m) = 0 \right. \right\},$$

and

$$\Theta_n(\mathbf{x}) := \sum_{t=0}^{J-1} w_t d(\mathbf{x}, H_{(k,p)}^-(\mathbf{x}_k^{(n)})) + w_J d(\mathbf{x}, C_{BI})$$

$$\Theta'_n(\mathbf{x}) := \sum_{t=0}^{J-1} w_t \frac{\mathbf{x} - P_{H_{(k,p)}^-(\mathbf{x}_k^{(n)})}(\mathbf{x})}{\|\mathbf{x} - P_{H_{(k,p)}^-(\mathbf{x}_k^{(n)})}(\mathbf{x})\|} + w_J \frac{\mathbf{x} - P_{C_{BI}}(\mathbf{x})}{\|\mathbf{x} - P_{C_{BI}}(\mathbf{x})\|}$$

with $(w_t)_{t=0}^J$ such that $w_t \geq 0$ and $\sum_{t=0}^J w_t = 1$.

Note : 1)The above projections can be computed in parallel. 2)The information of V_{NA} is fully utilized to improve each update. 3)As for detail of convergence of the proposed method, see [6].

Remark 2 (Application to MPEG Super-Resolution)

A great deal of effort has been devoted to derive higher resolution and quality video sequences from MPEG compressed video sequences [3, 13, 14]. The key of MPEG super-resolution is how to compensate the effect of quantization in compression. It is reported that the quantization noise in spatial domain is a colored Gaussian process [14]. Therefore, the proposed method can also be applied to a resolution enhancement of MPEG compressed video sequences.

4. NUMERICAL EXAMPLE

We generate a sequence of high-resolution images $(\mathbf{x}_k)_{k=1}^{14} \subset \mathbb{R}^{192 \times 256}$ by clipping out larger image. The clipping region at k th image shifts, relative to $k-1$ th image, from bottom-right to top-left by $(1, 2)(k:\text{odd})$ or $(2, 1)(k:\text{even})$. After blurring, downsampling of $(\mathbf{x}_k)_{k=1}^{14}$ and adding noise, low-resolution images $(\mathbf{y}_k)_{k=1}^{14} \subset \mathbb{R}^{96 \times 128}$ are derived. The PSF of blurring filter is a separable 2-D Gaussian truncated to 7×7 and normalized, and the noise is a white Gaussian with zero mean and variance $\sigma^2 = 30$. The motion estimation is assumed to be accurate.

For the proposed method, each stochastic property set is defined by using 3 pixels, and 4 independent sets are used for each iteration. Namely $q = 3$ and $J = 4$. In order to recover k th high-resolution image, 4 low-resolution images $\{\mathbf{y}_{k-3}, \mathbf{y}_{k-2}, \mathbf{y}_{k-1}, \mathbf{y}_k\}$ are used. For every iteration, we use $\lambda_n = 1.5$, $w_t = 1/5$, and $Q_{(k,p)} = I$ where I is an identity matrix.

A low-resolution image \mathbf{y}_{14} is shown in Fig. 1(a). Fig. 1(b) is a recovered image by NLMS, special case where $p = r = 1$ and $\rho_{(k,p)} \equiv 0$. The number of iterations is $4MN (= 49152)$. Application of the proposed method with threshold $\rho_{(k,p)} = 163 \simeq (q + \sqrt{2q})\sigma^2$, which is relatively large threshold in [9], and same number of iterations gives the results in Fig. 1(c). We can see that estimation error in $c_{(\ell,k)}$ makes the right edge of Fig. 1(b) stained, whereas introduction of the stochastic property set resolves the problem. This implies that NLMS based method suffers from being severely stained if there are several moving objects. With several numerical simulations, $\rho_{(k,p)} = 163$ provides fair noise suppression and robustness to estimation error without reducing convergence speed. For clarity, close up versions of Fig. 1 are presented in Fig. 2. A quantitative comparison in PSNR is given in Table 1. These exemplify that the proposed method fairly suppresses noise without losing an edge information as compared with NLMS.



(a) Low-resolution image

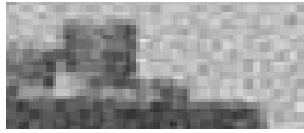


(b) NLMS ($q = r = 1, \rho = 0$)

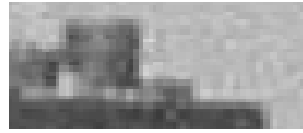


(c) Proposed with $\rho_{(k,p)} = 163$

Fig. 1. Low-resolution image and recovered high-resolution images



(a) NLMS ($q = r = 1, \rho = 0$)



(b) Proposed with $\rho_{(k,p)} = 163$

Fig. 2. Comparison in detail by closer look

Table 1. Comparison between PSNR values at $k = 14$

	NLMS	Proposed
PSNR[dB]	26.0	29.4

5. CONCLUSION

We presented a novel adaptive super-resolution of videos based on an embedded constraint version of *Adaptive projected subgradient method*. We introduced a novel model between successive high-resolution images so that we have detailed information on noise. Then robustness to the additive noise was realized by incorporating stochastic information of the noise.

6. 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] M. K. Ng and N. K. Bose, "Mathematical analysis of super-resolution methodology," *IEEE Signal Processing Magazine*, vol. 20, no. 3, pp. 62–74, May 2003.
- [3] Y. Altunbasak, A. J. Patti, and R. M. Mersereau, "Super-resolution still and video reconstruction from MPEG-coded video," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 12, no. 4, pp. 217–226, Apr. 2002.
- [4] C. A. Segall, R. Molina, and A. K. Katsaggelos, "High-resolution images from low-resolution compressed video," *IEEE Signal Processing Magazine*, vol. 20, no. 3, pp. 37–48, May 2003.
- [5] M. Elad and A. Feuer, "Superresolution restoration of an image sequence: Adaptive filtering approach," *IEEE Trans. on Image Processing*, vol. 8, no. 3, pp. 387–395, Mar. 1999.
- [6] I. Yamada and N. Ogura, "Adaptive projected subgradient method for asymptotic minimization of sequence of nonnegative convex functions," *Numerical Functional Analysis and Optimization*, vol. 25, no. 7&8, pp. 593–617, 2004.
- [7] P. L. Combettes, "Convex set theoretic image recovery by extrapolated iterations of parallel subgradient projections," *IEEE Trans. on Image Processing*, vol. 6, no. 4, pp. 493–506, Apr. 1997.
- [8] P. L. Combettes and J.-C. Pesquet, "Image restoration subject to a total variation constraint," *IEEE Trans. on Image Processing*, vol. 13, no. 9, pp. 1213–1222, Sept. 2004.
- [9] I. Yamada, K. Slavakis, and K. Yamada, "An efficient robust adaptive filtering algorithm based on parallel subgradient projection techniques," *IEEE Trans. on Signal Processing*, vol. 50, no. 5, pp. 1091–1101, May 2002.
- [10] S. Haykin, *Adaptive Filter Theory*, Prentice-Hall, 1996.
- [11] D. Rajan, S. Chaudhuri, and M. V. Joshi, "Multi-objective super resolution: Concepts and examples," *IEEE Signal Processing Magazine*, vol. 20, no. 3, pp. 49–61, May 2003.
- [12] E. S. Lee and M. G. Kang, "Regularized adaptive high-resolution image reconstruction considering inaccurate sub-pixel registration," *IEEE Trans. on Image Processing*, vol. 12, no. 7, pp. 826–837, July 2003.
- [13] H. Hasegawa, T. Ono, I. Yamada, and K. Sakaniwa, "An iterative MPEG super-resolution with an outer approximation of framewise quantization constraint," in *Proc. of IEEE International Workshop on Multimedia Signal Processing*, 2004.
- [14] C. A. Segall, A. K. Katsaggelos, R. Molina, and J. Mateos, "Bayesian resolution enhancement of compressed video," *IEEE Trans. on Image Processing*, vol. 13, no. 7, pp. 898–911, July 2004.