L_0 norm restricted LIC with ADMM

Shohei Kubota* , Ryoichiro Yoshida* , and Yoshimitsu Kuroki *

* National Institute of Technology, Kurume College, Komorino 1-1-1, Kurume-shi, Fukuoka, 833-8555 Japan

E-mail: a3405sk@kurume.kosen-ac.jp and kuroki@kurume-nct.ac.jp

***Abstract*—LIC (Local Intensity Compensation) is an intra-frame motion compensation for video coding, and was a candidate for HEVC. To obtain reference blocks and their coefficients in linear combination, Inoue et al. have applied AHIT (Accelerated Hard Iterative Thresholding) as a solver. However, this algorithm is heuristic; then, this paper proposes to use a sophisticated convex optimization solver ADMM (Alternating Direction Method of Multipliers). Experimental results demonstrate that our method provides higher PSNR values in the same reference block numbers than the conventional method.**

I. Introduction

The motion compensation is one of the encoding technologies of moving pictures to reduce temporal redundancy among inter frames. After the motion compensation, general standards as MPEG series reduce special redundancy of motion compensation errors by orthogonal transforms; for example, DCT (Discrete Cosine Transform). In the motion compensation stage, vectors which indicate horizontal and vertical differences between a target block and reference blocks, namely motion vectors, have to be sent to receivers or to stored in media. Recent standards decrease prediction errors by generalizing the compensation scheme: bi-directional prediction, half and quarter pixel interpolation, and various sizes of reference blocks. LIC (Local Intensity Compensation [1]) was a candidate for HEVC (High Efficiency Video Coding), and approximates a current block by linear combination of already coded reference blocks.

Sparse Representation for LIC is a reasonable approach to achieve high PSNR with low bit-rate encoding since it approximates a target vector with as few basis blocks as possible. The optimal solution of sparse representation defines the number of basis blocks with L_0 norm of their coefficient vector, and is regarded as a NP (Non-deterministic Polynomial time) problem. However, some sub-optimal solutions for sparse representation with L_0 norm is proposed in recent years. AIHT (Accelerated Iterative Hard Thresholding [2]) is a sub-optimal solver, but this solution is heuristic; thus, this paper tackles a more sophisticated solver for LIC by applying ADMM (Alternating Direction Method of Multipliers [3]).

II. Sparse representation and Accelerated Iterative Hard Thresholding(AIHT)

A. Definition

Sparse representation tries to approximate a original signal $y \in R^M$ with linear combination of basis signals $a_i \in R^M (i = 1, 2, \cdots, N > M)$, where the coefficients x_i's of a_i's are also denoted by a vector $x \in R^M$. Furthermore, the basis is expressed as a matrix A, of which columns are given by a_i's; then, sparse representation approximates y as $y \approx Ax$ with less number of a_i's. The number of non-zero elements of a vector is expressed by L_0 norm of the vector. When L_0 norm of x is less than or equal to K, the problem is called K-Sparse Problem, and the solution $\hat{x}$ is obtained by the following equation:

$$\hat{x}_0 = \underset{x}{\operatorname{argmin}} ||y - Ax||_2^2 \text{ subject to } ||x||_0 \leq K. \quad (1)$$

B. LIC using sparse representation

LIC approximates a target block b_p with linear combination of reference block b_{ci}'s, which is formulated as sparse representation with coefficient vector $w = (w_1, w_2, \cdots, w_N)^T$:

$$\hat{w}_0 = \underset{w}{\operatorname{argmin}} \frac{1}{2}||b_p - Bw||_2^2 \text{ subject to } ||w||_0 \leq K; \quad (2)$$

where $B = (b_{c1}, b_{c2}, \cdots, b_{cN})$, and each block is represented as a vector in the raster scanning order.

C. Accelerated Iterative Hard Thresholding (AIHT)

AIHT is a solver to the above-mentioned K-Sparse Problem and obtains a sub-optimal coefficient vector. AIHT is an iterative method and the solution of (2) is summarized as

$$\hat{w}^{(n+1)} = h_K(w^{(n)} + B^T(b_p - Bw^{(n)})), \quad (3)$$

$$\hat{w_1}^{(n+1)} = \hat{w}^{(n+1)} + a_1(\hat{w}^{(n+1)} - w^{(n)}), \text{and} \quad (4)$$

$$w_2^{(n+1)} = \hat{w_1} + a_2(w_1^{(n+1)} - w^{(n-1)}); \quad (5)$$

where $w^{(n)}$ is a temporal solution at iteration n, $h_K(w)$ replaces the elements of w of which absolute values are less than the K-th largest elements with zero. The values a_1 and a_2 are closed-form to minimize $||b_p - B\hat{w}_2^{(n+1)}||_2^2$. After the computation from equation (3) through (5), AIHT updates $w^{(n+1)}$ with the condition:

$$||b_p - Bh_K(w_2^{(n+1)})||_2^2 > ||b_p - B\hat{w}^{(n+1)}||_2^2 \quad (6)$$

if inequation (6) is fulfill , $w^{(n+1)}$ is replaced as $\hat{w}^{(n+1)}$ and otherwise, $w^{(n+1)}$ is updated to $h_K(\hat{w}_2^{(n+1)})$.

III. Alternating Direction Method of Multipliers (ADMM)

ADMM is a convex optimization algorithm to solve the following formula:

$$\min_{x,y} f(x) + g(y) \text{ subject to } y = Lx, \quad (7)$$

978-1-5386-2615-3/18/$31.00

©2018 IEEE

where $f(x)$ and $g(y)$ are convex functions. This solver obtains the optimal solution by computing the following iteration:

$$x^{(n+1)} = \operatorname*{argmin}_{x} f(x) + \frac{1}{2\gamma}||y^{(n)} - Lx - z^{(n)}||_2^2, \quad (8)$$

$$y^{(n+1)} = \operatorname*{argmin}_{y} g(y) + \frac{1}{2\gamma}||y - Lx^{(n+1)} - z^{(n)}||_2^2, \text{and} \quad (9)$$

$$z^{(n+1)} = z^{(n)} + Lx^{(n+1)} - y^{(n+1)}. \quad (10)$$

IV. Proposed Method

A. Overview

As mentioned in the section II-B, AIHT obtains a sub-optimal coefficient vector of LIC, but the solver is heuristic; thus this paper tries to establish a more sophisticated solution by applying ADMM for LIC. In order to apply ADMM, we assume that the reference blocks with non-zero coefficients are around the target block. To measure the assumption, our method involves a distance function of w as expressed in the next subsection.

B. Distance function

The distance function $d(w)$ means a distance between a target block and its reference blocks; namely $d(w)$ is the squared sum of Manhattan distance of motion vectors with non-zero coefficients. Let denote motion vector of i-th reference block as (α_i, β_i); then, the distance function $d(w)$ is defined as

$$d(w) = \sum_{j \in J} (|\alpha_j| + |\beta_j|)^2, \quad (11)$$

where $J = \{i \in \mathbb{Z} | 1 \leq i \leq N, w_i \neq 0\}$. To balance the approximation criterion of a target block and the distance of non-zero reference blocks, the proposed formula uses parameter λ to multiply $d(w)$.

C. Indicator function

To formulate K-Sparse Problem, we also involve an indicator function $\iota_{S_K}(w)$ defined as

$$\iota_{S_K}(w) := \begin{cases} 0, & ||w||_0 \leq K \\ \infty, & otherwise. \end{cases} \quad (12)$$

Via the above discussion, we define the cost function of K-sparse LIC to be solved with ADMM as the following equation:

$$\min_{w} \frac{1}{2}||b_p - Bw||_2^2 + \lambda d(w) + \iota_{S_K}(w). \quad (13)$$

D. How to solve the formula with ADMM

In the comparison of equation (13) to (7), f and g are replaced:

$$\begin{aligned} f(u) &= \frac{1}{2}||b_p - Bu_1||_2^2 + \lambda d(u_2), & (14)\\ g(w) &= \iota_{N_K}(w), & (15)\\ u &= (u_1^T \; u_2^T)^T, \text{and} & (16)\\ L &= (I \; I)^T \text{ subject to } u = Lw; & (17) \end{aligned}$$

thus, equation (13) becomes the following equation:

$$\min_{u,w} f(u) + g(w) \text{ subject to } u = Lw, \quad (18)$$

and the iteration process form equation (8) through (10) can be applied to obtain the solution as described in the next subsection.

E. Update Expression

The iteration process for equation (18) is given as

$$u_1^{(n+1)} = \operatorname*{argmin}_{u_1} \frac{1}{2}||Bu_1 - b_p||^2 + \frac{1}{2\gamma}||w^{(n)} - u_1 - d_1^{(n)}||_2^2, \quad (19)$$

$$u_2^{(n+1)} = \operatorname*{argmin}_{u_2} \lambda d(w) + \frac{1}{2\gamma}||w^{(n)} - u_2 - d_2^{(n)}||_2^2, \quad (20)$$

$$w^{(n+1)} = \operatorname*{argmin}_{w} \iota_{S_K}(w) + \frac{1}{2\gamma}||w^{(n)} - u^{(n+1)} - d^{(n)}||_2^2, \text{and} \quad (21)$$

$$d^{(n+1)} = d^{(n)} + u_1^{(n+1)} + u_2^{(n+1)} - Lw^{(n+1)}, \quad (22)$$

where $d = (d_1^T \; d_2^T)^T$. Furthermore, $u_1^{(n+1)}$, $u_2^{(n+1)}$, and $w^{(n+1)}$ are solved as following.

The solution of $u_1^{(n+1)}$ is given by the following:

$$u_1^{(n+1)} = (B^T B + \frac{1}{\gamma} I)^{-1} \left\{ B^T b_p + \frac{1}{\gamma}(w^{(n)} - d_1^{(n)}) \right\}. \quad (23)$$

Next, $u_2^{(n+1)}$ is obtained as

$$u_{2_i}^{(n+1)} = \begin{cases} v_i, & \lambda d_i < \frac{1}{2} v_i^2 \\ 0, & otherwise, \end{cases} \quad (24)$$

where $u_{2_i}^{(n)}$ and the v_i are the i-th element of $u_2^{(n)}$ and that of $w^{(n)} - d_2^{(n)}$ respectively, and d_i is the squared Manhattan distance of i-th reference block. Before we conclude $w^{(n+1)}$, vector t and s are used to obtain $w^{(n+1)}$:

$$t_i = v_i^2 + v_{i+N}^2 - \frac{1}{2}(v_i - v_{i+N})^2, \quad (25)$$

where t_i is the i-th element of t, and s is defined as

$$s = h_K(t); \quad (26)$$

namely, s replaces elements of t smaller than the K-th largest element of t with zero. At last, w is obtained as follows:

$$w_i^{(n+1)} = \begin{cases} \frac{1}{2}(v_i + v_{i+N}) & (s_i \neq 0) \\ 0 & (otherwise), \end{cases} \quad (27)$$

where s_i is the i-th element of s.

F. Algorithm

From the above, the K-Sparse Problem solver is summarized as the Algorithm 1.

Algorithm 1 LIC with ADMM

Input: b_p is a target block, a matrix B contains reference blocks in column, λ is a parameter of distance function d, and we set $L = (I\ I)^T$, $\gamma > 0$, $\epsilon > 0$, and $0 < \eta < 1$.

Output: w (a K-sparse linear combination coefficient vector)

Initialize: i-th and $i+N$-th elements of $w^{(n)}$ and $d^{(0)}$ are set to one, vector $u^{(0)} = 0$

while $||w^{(n)} - w^{(n-1)}||_2^2 > \epsilon$ **do**

  $u_1^{(n+1)} = (B^T B + \frac{1}{\gamma} I)^{-1} \left\{ B^T + \frac{1}{\gamma}(w^{(n)} - d_1^{(n)}) \right\}$

  calculate $u_2^{(n+1)}$ with equation (23)

  calculate $w^{(n+1)}$ with equation (24), (25), (26)

  $d^{(n+1)} = d^{(n)} + u^{(n+1)} - Lw^{(n+1)}$

  $\gamma \leftarrow \eta\gamma \quad n \leftarrow n + 1$

end while

V. Experiments

We compare the proposed method and LIC with AIHT for grayscale images of test sequences ”foreman” and ”coastguard” as shown in Figure 1 and 2. The resolution is down sampled to 128×128 pixels. GOP structure consists of one I-frame and three consecutive P-frames, and a P-frame refers only previous I- or P-frame. We used only 16 frames, namely 4GOPs due to time consumption. The block size is 16×16, and the other parameters are set as: $\lambda = 0.001$, $\epsilon = 10$, $\eta = 0.97$. The reference blocks in a reference frame, and the initial value of $\gamma = 3$. The number of reference blocks, namely K, is varied from 1 through 10, and we evalutate the compensated image quality with the mean values of SSIM (Structural Similarity [4]) to the original images. Only to estimate LIC performance, SSIM values are obtained for LIC images: then, DCT coefficients coding for motion compensated errors is not added to reconstructed images.

Figure 3 and 4 depict the average SSIM with varying K, which tell us that the reconstructed image quality of our method is superior to that of the conventional method. Figure 5 through 10 show the average values of SSIM of first, second, and third P-frames in GOPs. These figures also indicate the similar superiority of our method. Reconstructed images with the proposed and the conventional methods are illustrated in Figure 11 through 14, which furthermore impressed us that our method accurately predicts inter frames.

Fig. 1. foreman

Fig. 2. coastguard

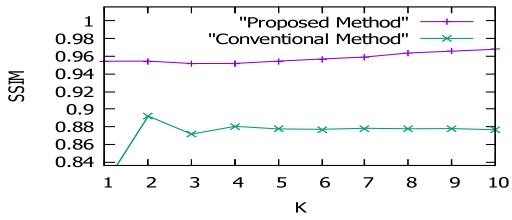

Fig. 3. foreman's SSIM

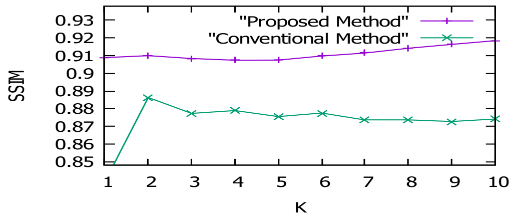

Fig. 4. coastguard's SSIM

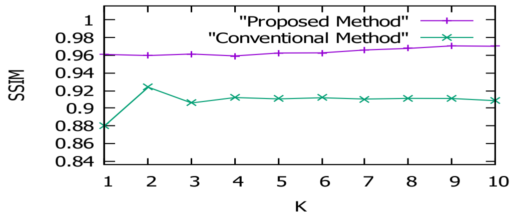

Fig. 5. foreman's SSIM (P=1)

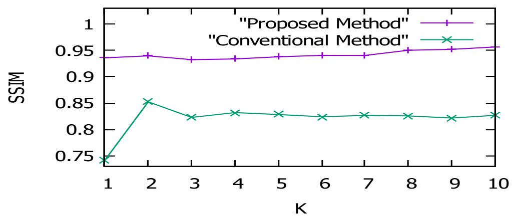

Fig. 6. foreman's SSIM (P=2)

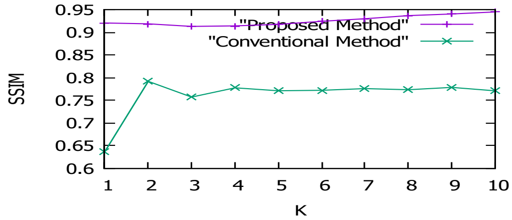

Fig. 7. foreman's SSIM (P=3)

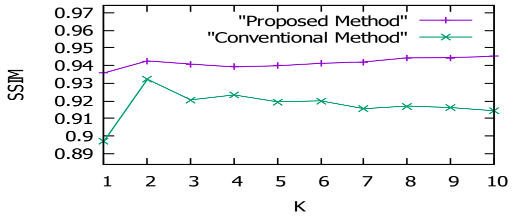

Fig. 8. coastguard's SSIM (P=1)

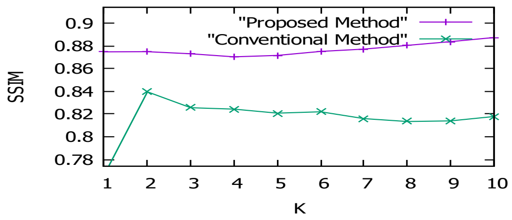

Fig. 9. coastguard's SSIM (P=2)

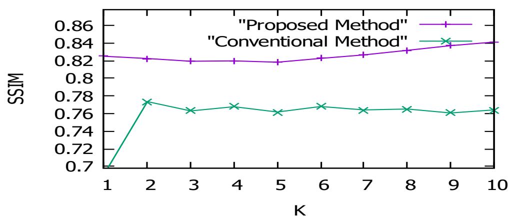

Fig. 10. coastguard's SSIM (P=3)

Fig. 11. Proposed Method, 4-th frame (SSIM = 0.8831)

Fig. 12. Conventional Method,4-th frame (SSIM=0.3973)

Fig. 13. Proposed Method, 4-th frame (SSIM=0.8317)

Fig. 14. Conventional Method, 4-th frame (SSIM = 0.6923)

VI. Conclusions

This paper has proposed a L_0 norm restricted LIC for inter frames prediction. To apply ADMM to the LIC, we have involved the convex distance function of reference blocks with non-zero coefficients and the indicator function to restrict K-Sparse coefficient vector. The experimental results have demonstrated that our method generate more accurate prediction images that the LIC using AIHT. In future, we would like to investigate encoding of motion vectors and quantized coefficients to involve this method to encoding systems.

Acknowledgment

This study is funded from the JSPS Grants-in-Aid for Scientific Research 15K00257.

References

[1] N. Sprljan, S. Paschalakis, and P. Wu, “Local intensity compensation for inter prediction in HEVC,” Doc. JCTVC-B096, JCT-VC of ITU-T VCEGand ISO/IEC MPEG, Geneva, Switzerland, July 2010.
[2] T.Blumensath, ”Accelerated iterative hard thresholding,” Signal Processing, vol.92, no.3, pp.752-756, Mar.2012.
[3] D. Gabay and B. Mercier, ”A dual algorithm for the solution of non linear variational problems via finite-element approximations,” Computers & Mathmatics with Applications, vol.2, pp.17-40, Jan. 1976.
[4] Z. Wang, A.C. Bovik, H.R. Sheikh and E.P. Simoncelli, ”Image qualicty assessment: From error visibility to structural similarity,” IEEE Transactions on Image Processing, vol 13, no. 4, pp. 600-612, Apr. 2004.