AN ANALYTICAL FRAMEWORK FOR OVERLAPPED MOTION COMPENSATION

Jiro Katto and Mutsumi Ohta

Information Technology Research Laboratories, NEC Corporation,
1-1, Miyazaki 4-chome, Miyamae-ku, Kawasaki-shi, KANAGAWA 216, JAPAN.
E-Mail: katto@dsp.cl.nec.co.jp

ABSTRACT

This paper presents a novel framework which proves the superiority of overlapped motion compensation. Window design problem is revised by introducing a statistical model of motion estimation process. The result clarifies relationship between the optimum window and image characteristics in an explicit formula and quantifies prediction error reduction achieved by overlapped motion compensation. Experimental results using real image sequences support the proposed theory and demonstrate its superiority. Overlapping in warping prediction is also considered and its effectiveness is shown.

1. INTRODUCTION

Motion compensated interframe prediction is a key technology to accomplish efficient video codecs. In particular, overlapped motion compensation has received much attention because it can reduce blocking distortions and prediction errors simultaneously [1, 2]. This method utilizes motion vectors of neighboring blocks and takes the weighted sum of plural predicted values. Window optimization has been a topic in order to improve its prediction efficiency [3, 4]. On the other hand, there exists an alternative approach known as warping prediction [5, 6, 7]. This method calculates motion vectors for each pixel using vector interpolation from neighboring control grid points. Vector interpolation smoothes motion fields and similar advantages to overlapped motion compensation have been achieved.

This paper firstly revises a window optimization problem for overlapped motion compensation. The result removes heuristics from the previous considerations and proves the superiority of overlapped motion compensation. Combination of overlapping and warping prediction is also considered in a very natural way. Experimental results demonstrates validity of the proposed framework and effectiveness of the proposed algorithms, respectively.

2. FORMULATION OF OVERLAPPED MOTION COMPENSATION

2.1. Revision of Window Optimization Problem

Overlapped motion compensation (Figure 1) is characterized by its prediction structure utilizing plural displacement vectors (i.e. plural reference blocks). It is formulated by

$$\hat{z}(x,y) = \sum_{i} \sum_{m,n} w_i(m,n)\, z_i(x - d_{i,x}, y - d_{i,y}) \qquad (1)$$

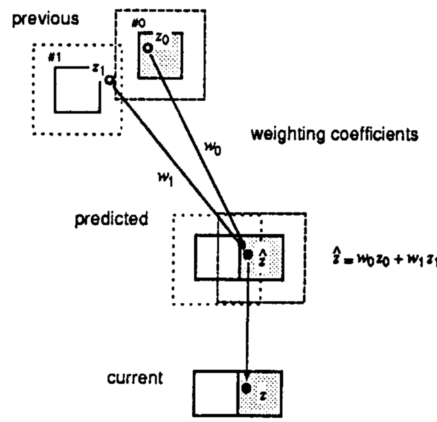

Figure 1: Overlapped motion compensation

where $\hat{z}$ is a predicted pixel, (m,n) is a pixel position in a macroblock, $w_i(m,n)$ is a weighting coefficient (composing a window function), $(d_{i,x}, d_{i,y})$ is an estimated displacement vector and z_i is a candidate pixel, where i indicates the i-th neighboring block (i=0 means itself).

Adequate choice of weighting coefficients brings about drastic reduction in prediction errors. We have already formulated it in a statistical manner [4]. In one-dimensional case, it is given by

$$w_0(m) = 0.5 + 0.5 \cdot \frac{C_0 - C_1}{1 - C_2} \qquad (2)$$

where

$$C_0 = \frac{E[z \cdot z_0]}{E[z^2]},\ C_1 = \frac{E[z \cdot z_1]}{E[z^2]},\ C_2 = \frac{E[z_0 \cdot z_1]}{E[z^2]} \qquad (3)$$

and $w_1 = 1 - w_0$. Measuring these statistical quantities per pixels in a macroblock from real image sequences, the optimum window has been acquired.

In order to provide a fully analytical solution, we place two assumptions on the motion estimation process (Figure 2):

- corresponding pixels do not change from frame to frame (the previous pixel may not be on sampling grids),
- spatial correlation between pixels decreases exponentially.

These assumptions are represented by

$$z(x,y) = z_0(x - d_{0,x} - \Delta d_{0,x}, y - d_{0,y} - \Delta d_{0,y}) \qquad (4)$$

 0-7803-2431-5/95 $4.00

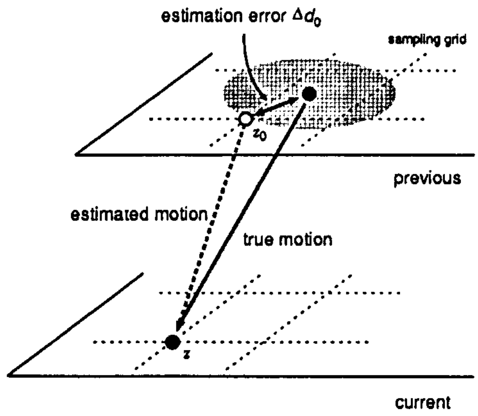

Figure 2: Assumptions on motion estimation process

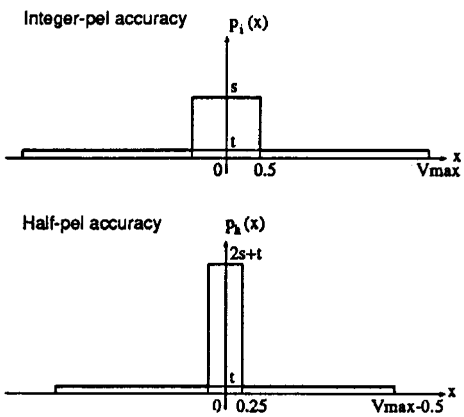

Figure 3: Probability density functions of displacement estimation errors

$$\frac{E\left[z_0(x,y)\, z_0(x-\Delta d_{0,x}, y-\Delta d_{0,y})\right]}{\sigma_z^2} = \rho_S^{E[|\Delta d_0|]} \quad (5)$$

where z is an input pixel of the current frame to be predicted, Δd_i is a displacement estimation error vector, σ_z^2 is input source variance, ρ_S is a spatial correlation factor and $E[\cdot]$ is an expectation. Applying these assumptions to the Eq.(2), the optimum weighting coefficient is given by

$$w_0(m) = 0.5 + 0.5 \cdot \frac{\rho_S^{a_0} - \rho_S^{a_1}}{1 - \rho_S^{(a_0+a_1)/2}}, \quad (6)$$

which is a function of a spatial correlation factor ρ_S and expectations of absolute displacement estimation errors a_i ($=E[|\Delta d_i|]$). Parameter a_0 will increase monotonously from the center of a macroblock to its boundary and approaches to a_1 [4].

2.2. Quantitative Evaluation of Prediction Error Variances

Let e_{blk} be a prediction error of block motion compensation, i.e.

$$e_{blk} = z - z_0. \quad (7)$$

Taking the assumptions in Eq. (4) and (5) into account, $E[e_{blk}^2]$ / $E[z^2]$ is given by

$$2 \cdot (1 - \rho_S^{a_0}) \quad (8)$$

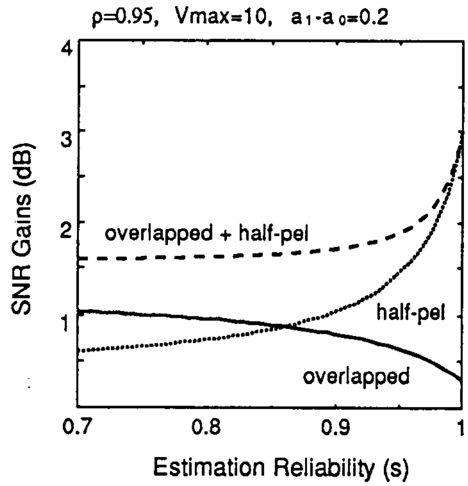

Figure 4: Comparison of prediction error variances

which indicates prediction efficiency of block motion compensation. For overlapped motion compensation, $E[e_{olmc}^2] / E[z^2]$ is provided by

$$2 \cdot \left[1 - \rho_S^{a_0} - \frac{1}{1-\rho_S^{(a_0+a_1)/2}} \left(\frac{1 - \rho_S^{(a_0+a_1)/2} - \rho_S^{a_0} + \rho_S^{a_1}}{2}\right)^2\right]. \quad (9)$$

The third term in the parenthesis represents improvement of prediction efficiency achieved by overlapped motion compensation.

Effect of motion estimation accuracy is also included. For this purpose, probability density functions of displacement estimation errors are assumed to have quantization-error-like shapes with siderobes as shown in Figure 3. This p.d.f. is chosen instead of the Gaussian assumption [8] due to its mathematical tractability. In this figure, parameter s means a probability density of accurate estimation, t does that of inaccurate estimation and V_{max} is a maximum value of displacement estimation errors. These parameters are directly connected to the a_0 by a simple formula. The p.d.f. for half-pel accuracy is defined by halving the range in which accurate estimation is guaranteed. According to this deformation, a_0 for the half-pel accuracy is redefined.

Figure 4 shows an example of theoretical SNR gains obtained by three motion compensation techniques against the integer-pel accuracy block-based one. Horizontal axis (s in Figure 3) represents motion estimation reliability. Contribution of overlapped motion compensation is reflected by parameter a_1 which is statistically larger than a_0 in a macroblock. This result suggests that overlapped motion compensation performs well when the estimation reliability is low. In other words, even if motion estimation is done roughly, overlapped motion compensation brings about prediction gains by utilizing displacement vectors of neighboring blocks. Notice that the effect of half-pel estimation accuracy is decreased in such an environment. It is also pointed out that the overlapped motion compensation with half-pel estimation accuracy provides stable gains whether the motion estimation works well or not. This means that these techniques work so as to compensate for each other's disadvantages.

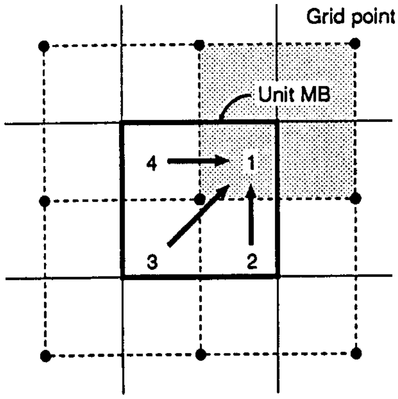

Figure 5: Overlapping in Quadrilateral Warping

3. OVERLAPPING IN WARPING PREDICTION

3.1. Quadrilateral

One macroblock in a predicted image is composed of parts of four non-overlapped blocks created by quadrilateral warping prediction. In Figure 5, these blocks are indicated by 1, 2, 3 and 4. Overlapping procedure can be directly introduced by considering vector extrapolation. For example, a predicted macroblock in the region 1 is created by taking the weighted sum of four different pixels indicated by interpolated or extrapolated motion vectors calculated in each region. As a window function, trapezoid or raised-cosine shapes [3, 4] are candidates.

3.2. Triangle

By dividing each quadrilateral in the control grid into two triangles, triangle warping prediction is defined. In regard to the dividing way, there are two choices: right-angled or left-angled. For this problem, two approaches are known: fixing one side [7] or adaptive selection according to MSE [9]. Overlapping concept, however, encourages us to use both of them. Figure 6 shows this idea. Two shadowed quadrilaterals with 45° sloped are composed of four non-overlapped triangles which are determined according to different arrangement patterns. Each group of quadrilaterals can cover whole images. By adding them, another prediction image is created. As a window function, a simple averaging window or a diamond shape [3] are candidates.

3.3. Discussions

We can show a simple background which supports utilization of both of two triangle structures. Let z be a current pixel. Let z_0 and z_1 be predicted pixels created by different prediction methods. If these prediction methods contribute similarly,

$$\frac{E[z \cdot z_0]}{E[z^2]} = \frac{E[z \cdot z_1]}{E[z^2]} = \rho \tag{10}$$

can be assumed. A prediction error variance using a single prediction method is given by

$$\sigma_s^2 = 2(1-\rho) \cdot E[z^2]\,. \tag{11}$$

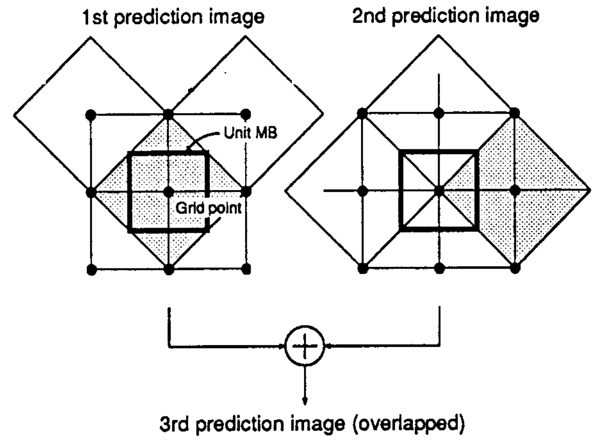

Figure 6: Overlapping in Triangle Warping

When two predicted pixels are averaged, its prediction error variance is shown by

$$\sigma_a^2 = \frac{1}{2}(3 - 4\rho + \rho') \cdot E[z^2] \tag{12}$$

where $\rho' = E[z_0 \cdot z_1] \,/ E[z^2]$. Therefore,

$$\sigma_s^2 - \sigma_a^2 = \frac{1}{2}(1-\rho') \cdot E[z^2] \tag{13}$$

is formed. This means that, when different prediction methods are used ($\rho' < 1$) and they work with similar effects, an averaging procedure always brings prediction gains.

4. EXPERIMENTS

4.1. Premises

Simulations are carried out on the condition that

- Picture format: eight QCIF sequences, 7.5 frames/sec, 90 frames (Miss_America and Susie are 30 frames),
- Motion estimation: conventional block matching, full search, 16×16 macro block, ±15 × ±15 search range.

This paper focuses on motion compensation process only, just to show the overlapping effect. Iterative motion estimation algorithms with overlapping [3] or grid point modification [5, 6] are not considered.

4.2. Verification of the Proposed Framework

Figure 7 shows simulation results using real image sequences. SNR gains obtained by overlapped motion compensation and half-pel accuracy motion estimation are presented against the classical method (block and interger-pel accuracy). A window shape applied to overlapped motion compensation is a trapezoid [4]. These results are mostly connected to the analytical expectations in Figure 4. We can estimate parameter a_i in Eq. (6), which is not directly measurable, by calculating

$$a_i = \frac{\ln\left(1 - \frac{1}{2} \cdot \frac{E[e_{blk}^2]}{E[z^2]}\right)}{\ln \rho_S} \tag{14}$$

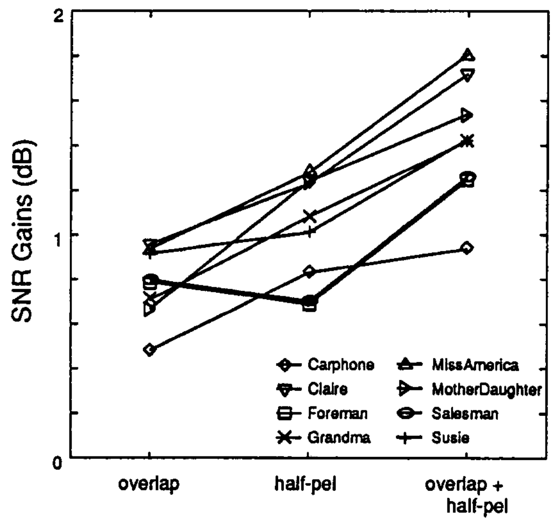

Figure 7: SNR gains achieved by overlapped motion compensation and half-pel motion estimation for eight QCIF sequences

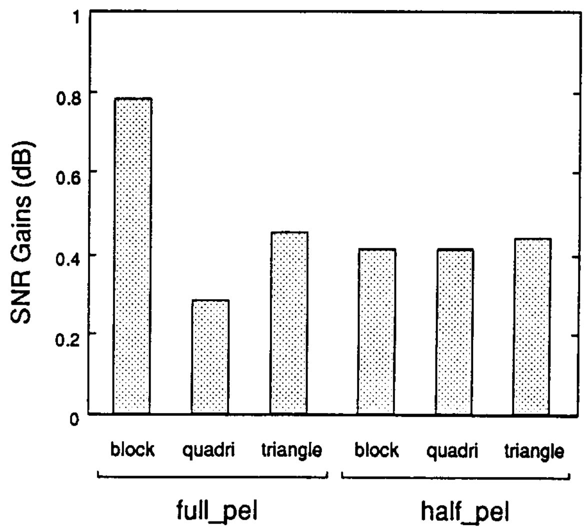

Figure 8: Averaged SNR gains achieved by overlapping in three motion compensation techniques

for each pixel in a macroblock. Although this equation has a problem on its stability now, it presents estimation reliability of each image sequences and provides excellent windows for some sequences.

4.3. Effectiveness of Overlapping Algorithms

Figure 8 shows averaged SNR gains achieved by overlapping in three different motion compensation techniques. *Block* corresponds to the conventional approach without vector interpolation, *quadri* does to warping prediction in Figure 5 and *tri* does to that in Figure 6. Competitors are non-overlapping versions of each strategy. In the triangle case, gains to an averaged SNR of the 1st prediction image and the 2nd one in Figure 6 are shown. Applied windows are a trapezoid to the quadrilateral and an average to the triangle, respectively. This result fully confirms us of the effectiveness of overlapping in any motion compensation methods.

5. CONCLUSIONS

This paper presents a novel framework for overlapped motion compensation and expands its concept to warping prediction. Overlapping means that prediction images are composed of plural reference images indicated by plural motion vectors. This is identical to the case of B-picture used in MPEG [10]. In other words, *interpolation* is a key issue to improve coding efficiency both in spatial and temporal directions. Experimental results verify the superiority of overlapped motion compensation and encourage us to introduce it into a practical video coder [4] or MPEG4.

6. ACKNOWLEDGMENT

The authors thanks to Prof. M. Orchard and Prof. T. G. Campbell for their valuable discussions. They also thanks to Dr. T. Nishitani for his helpful suggestions.

7. REFERENCES

[1] M. Ohta and S. Nogaki: "Hybrid picture coding with wavelet transform and overlapped motion-compensated interframe prediction coding", IEEE Trans. Signal Process., pp.3416-3424, Dec.1993.

[2] H. Watanabe and S. Singhal: "Windowed motion compensation", VCIP'91, pp.582-589, Nov.1991.

[3] M. T. Orchard and G. J. Sullivan: "Overlapped block motion compensation: an estimation-theoretic approach", IEEE Trans. Image Process., pp.693-699, Sep.1994.

[4] J. Katto, J. Ohki, S. Nogaki and M. Ohta: "Wavelet codec with overlapped motion compensation for very low bit-rate environment", IEEE Trans. CSVT, pp.328-338, Jun.1994.

[5] G. J. Sullivan and R. L. Baker: "Motion compensation for video compression using control grid interpolation", ICASSP'91, pp.2713-2716, May.1991.

[6] Y. Nakaya and H. Harashima: "Motion compensation based on spatial transformations", IEEE Trans. CSVT, pp.339-356, Jun.1994.

[7] T. G. Campbell: "Warping prediction for very low bitrate video coding", PCS'94, pp.269-272, Sep.1994.

[8] R. Girod: "Motion-compensating prediction with fractional-pel accuracy", IEEE Trans. Commun., pp.604-612, Apr.1993.

[9] H. Katada, T. Aono and H. Kusao: "Motion compensation scheme in very low bit rate video coding", PCSJ'94, pp.121-122, Oct.1994 (in Japanese).

[10] ISO-IEC/JTC1/SC29/WG11: "Generic coding of moving pictures and associated audio", Committee Draft 13818-2, Nov.1993.