A Modified Parabolic Prediction Based Fractional Pixel Motion Estimation Using Preset Corrector

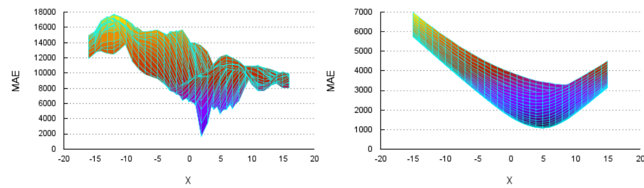
Chang-Uk Jeong and Hiroshi Watanabe
Graduate School of Global Information and Telecommunication Studies
Waseda University
1-3-10 Nishi-Waseda, Shinjuku-ku, Tokyo 169-0051, Japan

***Abstract*—In general, a motion estimator in today's video coding includes fractional pixel motion estimation (FME) as well as integer pixel motion estimation (IME). FME can get better quality performance at the cost of higher computational complexity than IME alone. In this paper, a modified parabolic prediction based FME for H.264 video coding is proposed. The proposed method uses specific correction coefficients to improve the PSNR performance of the existing prediction based algorithm. In the simulation results, compared with the original algorithm, our technique shows more accurate predictive power, meanwhile, it does not require the interpolation process.**

Keywords-Video coding; fractional pixel motion estimation

I. INTRODUCTION

Thanks to the development of digital video compression technology, especially digital broadcasting has become more and more popular, such as standard definition television (SDTV) and high definition television (HDTV). Ultra high definition television (UHDTV) proposed by NHK is the latest digital video format, which has 16 times the resolution of HDTV [1]. On the other hand, as digital video contents such as HDTV sequence include a massive amount of the information, they are becoming a major cause of today's network traffic, especially on the Internet. Moreover, the sum of all video types is forecasted to occupy most of global consumer traffic for years ahead. As a result, the increasing popularity of digital video and broadcasting has led to the evolution of video compression standards, such as MPEG-1,2,4, H.261, H.263 [2], and H.264/AVC [3]. As video compression is the most significant attempt to reduce video data, it is a process by which digital signals are simplified by eliminating redundancy. Generally, a motion estimation (ME) module in a video encoder has the highest computational complexity. To reduce the computational complexity, therefore, many researches on ME have been tried up to now. Particularly, quite a number of techniques for integer pixel ME (IME), including the diamond search (DS) [4], the hexagon-based search (HEXBS) [5], and the unsymmetrical-cross multi-hexagon-grid search (UMHexagonS) [6], have been proposed. Although fractional pixel ME (FME) has a strong impact on PSNR, it increases the total encoding time of the video encoder because of the process based on the interpolation. The center biased fractional pixel search (CBFPS) [6] is one of the most well-known interpolation based FME algorithms. In this paper, our research focuses on an interpolation-free method in order not to use any search points.



(a) (b)

Figure 1. (a) Error surface of IME. (b) Error surface of FME.

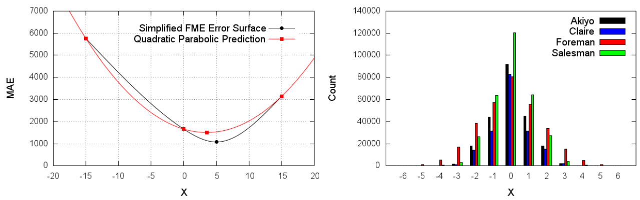
II. PARABOLIC PREDICTION BASED FME

Two error surfaces are illustrated in Figure 1. They were simulated for CCIR601 "Garden" sequence with the search range of ±16. 1/16 pixel motion vector (MV) resolution is used for FME. The error surface of FME is undoubtedly unimodal but that of IME is irregular. The fractional pixel search points within the FME search area are generated by performing the interpolation process using the IME error costs. Accordingly, the FME error cost increases monotonically as the search point moves away from the best point with the minimum error cost. The side of a FME error surface is also shaped like parabola, as shown in Figure 1 (b). To plot a parabolic model, a degenerate quadratic parabolic prediction function is introduced in Equation 1, as already mentioned in [7].

$$F(x,y) = c_4 x^2 + c_3 x + c_2 y^2 + c_1 y + c_0 \tag{1}$$

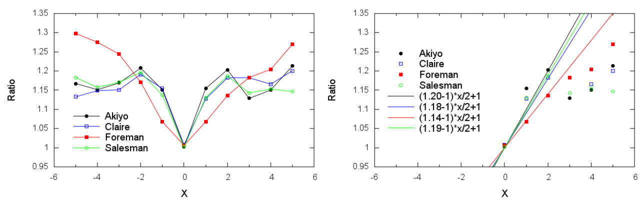
where x and y denote fractional pixel position and c_4, c_3, c_2, c_1, and c_0 are the five coefficients of the function F. If the five integer pixel search points $P_4(0,1)$, $P_3(0,-1)$, $P_2(1,0)$, $P_1(-1,0)$, and $P_0(0,0)$ are known by the IME process, the five coefficients can be computed by substituting the coordinates and error costs of the five integer pixel search points for the variables of Equation 1. As shown in Equation 2, the predictive minimum error cost is obtained by partially differentiating F with respect to x and y, respectively. When $\partial_x F = \partial_y F = 0$, the x and y coordinates are regarded as the best prediction position.

$$\frac{\partial F}{\partial x}(x,y) = 2c_4 x + c_3 = 0, \quad \frac{\partial F}{\partial y}(x,y) = 2c_2 y + c_1 = 0 \tag{2}$$
$$x = -\frac{c_3}{2c_4}, \quad y = -\frac{c_1}{2c_2}$$



(a) (b)

Figure 2. (a) An Example of the difference between real-world FME error surface and parabolic prediction model. (b) The matching counting by the distance between the best position of FFPS and that of PPBF.



(a) (b)

Figure 3. (a) The ratios P_1/P_2 and P_2/P_1 by the distance between the best position of FFPS and that of PPBF. (b) The line fitting for [0,6] of (a).

III. Proposed FME

Figure 2 (a) represents the difference between real-world FME error surface and parabolic prediction model. The above mentioned parabolic model has perfectly bilateral symmetry by an axis but real-world FME error surface is a little bit uneven. Thus, we introduce a solution to narrow the gap. A simulation has been tried for QCIF "Akiyo", "Claire", "Foreman", and "Salesman" sequences. The simulation focuses on how the average FME error costs will change according to the distance between the full fractional pixel search (FFPS) and the parabolic prediction based FME (PPBF) at quarter pixel MV resolution. In this simulation, we suppose that the FFPS matching performance is the best of all FME algorithms. Figure 2 (b) shows that the maximum range is -6 to 6 and the number of the matching of the distance zero is the largest whereas that of the ranges [-6,-3] and [3,6] is very few. As shown in Figure 3 (a), particularly, we take note that the higher the ratio of P_1/P_2 or P_2/P_1 is, the farther the distance tends to be. Based on this analysis, the linear function $L(x)$ passing by both two points (0,1) and (2,n) is derived as follows:

$$L(x) = \frac{n-1}{2}x + 1, \quad L^{-1}(x) = \frac{2x-2}{n-1} \tag{3}$$

where n indicates the preset correction coefficient, which is an alternative name for the provisional ratio P_1/P_2 or P_2/P_1 corresponding to the distance 2, $L(x)$ is the ratio of the current P_1 and P_2, x of $L(x)$ is the distance. The distance $L^{-1}(x)$, which is also called the adjusting factor and must be divided by 4 at quarter pixel MV resolution, can be found if we know the preset correction coefficient n as listed in Table I. In the final step of the proposed FME, the adjusting factor added to the best prediction position obtained by PPBF is determined as the best modified prediction position, and then the quantization operation presented by [7] is carried out.

TABLE I. The Preset Correction Coefficients for Sequences

Location	x-coordinate		y-coordinate	
Distance	-2 (P_1/P_2)	2 (P_2/P_1)	-2 (P_3/P_4)	2 (P_4/P_3)
Akiyo	1.207	1.203	1.238	1.270
Claire	1.191	1.182	1.168	1.173
Foreman	1.171	1.136	1.179	1.212
Salesman	1.197	1.187	1.434	1.455

TABLE II. Performance Comparison Between FME Algorithms

Sequence	Method	PSNR (dB)	Bit-rate (kbps)
Akiyo	CBFPS	38.274	25.718
	PPBF	-0.002	26.620
	Proposed	+0.020	26.635
Claire	CBFPS	39.777	32.032
	PPBF	+0.006	32.968
	Proposed	+0.016	32.852
Foreman	CBFPS	35.948	142.010
	PPBF	-0.008	150.846
	Proposed	-0.004	150.973
Salesman	CBFPS	35.814	55.182
	PPBF	-0.037	59.908
	Proposed	-0.029	59.930

IV. Experimental Results and Future Work

The four typical QCIF sequences are used to evaluate the proposed FME. 200 frames are encoded in each video sequence. The simulations were conducted using H.264/AVC reference software JM 12.4 [8] with baseline profile. Quantization parameter is set to 28, search range is 16, and rate distortion optimized mode is on. After UMHexagonS for IME, FME is processed. As shown in Table II, the PSNR performance of the proposed FME is better than that of PPBF and very near that of CBFPS base on the interpolation. Furthermore, the proposed FME and the existing PPBF do not use any search points but CBFPS definitely needs them. The preset correction coefficient applied to the proposed FME, however, is necessary to be determined adaptively based on content.

References

[1] M. Kanazawa, K. Mitani, K. Hamasaki, M. Sugawara, F. Okano, K. Doi, and M. Seino, "Ultrahigh-definition video system with 4000 scanning lines," in *Proc. the Internationa Broadcasting Convention (IBC)*, pp. 321-329, Amsterdam, Netherlands, Sep. 2003.

[2] K. R. Rao and J. J Hwang, *Techniques and Standards for Image, Video and Audio Coding*, Englewood Cliffs, NJ: Prentice Hall, 1996.

[3] *Draft ITU-T Rec. and Final Draft International Standard of Joint Video Specification (ITU-T Rec. H.264-ISO/IEC 14 496-10 AVC)*, Joint Video Team (JVT) of ISO/IEC MPEG and ITU-T VCEG, JVT-G050r1, Geneva, Switzerland, Mar. 2003.

[4] S. Zhu and K. K. Ma, "A new diamond search algorithm for fast block matching motion estimation," *IEEE Trans. on Image Process.*, vol. 9, no. 2, pp. 287-290, Feb. 2000.

[5] C. Zhu, X. Lin, and L. P. Chau, "Hexagon-based search pattern for fast block motion estimation," *IEEE Trans. on Circuits Syst. Video Technol.*, vol. 12, pp. 349-355, May 2002.

[6] Z. Chen, P. Zhou, and Y. He, "Fast integer pel and fractional pel motion estimation for JVT," JVT-F017, 6th meeting, Awaji Island, Japan, Dec. 2002.

[7] J. F. Chang and J. J. Leou, "A quadratic prediction based fractional-pixel motion estimation algorithm for H.264," in *Proc. Seventh IEEE Int. Symp. on Multimedia.*, pp. 491-498, Dec. 2005.

[8] JVT H.264/AVC Reference Software Joint Model (JM), http://iphome.hhi.de/suehring/tml/