SELECTION OF THE LAGRANGE MULTIPLIER FOR BLOCK-BASED MOTION ESTIMATION CRITERIA

Pekka Sangi, Janne Heikkilä, Olli Silvén

Machine Vision Group, Infotech Oulu
P.O.Box 4500, FIN-90014, University of Oulu, Finland
Email: {psangi, jth, olli}@ee.oulu.fi

ABSTRACT

In hybrid video coding, motion vectors used for motion compensation constitute an important set of decisions. Cost functions for block motion estimation that take the smoothness of the resulting motion vector field into account, in addition to the motion compensated prediction error, have been proposed. Computationally simple derivatives of SAD and SSD-based criteria are studied in this paper. Cost functions are based on Lagrangian rate-distortion formulation, and the basic question is how the Lagrangian multiplier involved should be selected. Assumptions behind these cost functions are discussed, and a new method is derived for determining the multiplier. Comparisons with other strategies are made with experiments. The results show that the selection of the multiplier is not critical.

1. INTRODUCTION

Motion-compensated transform video coding, also called hybrid video coding, provides an efficient scheme for video compression. Various standards, such as H.263 and MPEG-4, define bitstream syntaxes, and the main problem that remains is how various encoding decisions should be made. One important decision is what motion vector (MV) (or a set of block MVs) should be used for macroblock motion compensation in the inter-mode. In most implementations, the choice is based on the sum of absolute differences (SAD) or the sum of squared differences (SSD) criteria, which correlate with the resulting residual bit rate and distortion. However, these criteria do not consider the bit rate required by encoding of MVs. As this is done in a predictive manner, more regular MV fields can give a lower bit rate with insignificant change in distortion, especially at lower bit rates.

It is therefore interesting to consider other MV estimation criteria which also take the MV bit rate into account. A common way of formulating such criterion is based on Lagrangian optimization, where the original constrained optimization problem is converted to an unconstrained one. A reformulated problem involves coefficients called Lagrange multipliers, and a particular choice of them corresponds to the given constraint.

However, a closed-form or computationally feasible solution for the coefficient cannot be found for the problem in video coding due to dependencies between various encoding decisions. The lack of methods with respect to selection of Lagrange multipliers has been one obstacle to considering the use of a Lagrangian coder control [1]. In this paper, we address this problem from the view point of motion estimation, and consider various strategies for selecting the multiplier. After presenting the related background, an analysis of criteria is used to derive a new method for selecting multipliers, which is then experimentally compared to other methods.

The work described here has been financially supported by TEKES, the National Technology Agency of Finland.

2. BACKGROUND

Due to dependencies between various encoder decisions, obtaining a joint rate-distortion (R-D) optimal solution to a video coding problem is not possible in practice. Therefore, various sequential schemes are used for seeking suboptimal solutions [2]. For example, selection of the quantizer parameter (QP) and MV for motion compensation are done independently.

Considering the choice of MV $\vec{v}$, when the quantizer parameter q has been fixed, the cost function to be minimized is

$$J(\vec{v}, q) = D(\vec{v}, q) + \lambda'(R_{res}(\vec{v}, q) + R_{mv}(\vec{v})) \quad (1)$$

where $D(\vec{v}, q)$ corresponds to the distortion, measured usually as the mean-squared error (MSE) of the luminance channel, $R_{res}(\vec{v}, q)$ is the number of bits needed for encoding the quantized DCT coefficients of the residual, $R_{mv}(\vec{v})$ is the number of bits required by predictive coding of the MV, and λ' is the Lagrange multiplier. Direct use of this criterion requires performing DCT and quantization operations on an MV candidate basis, which is computationally demanding.

In order to avoid the problem, model-based approaches has been proposed (see e.g. [3]), where the criteria used involve parametric approximations of distortion and rate, that is,

$$D(\vec{v}, q) \approx D'(C(\vec{v}), q) \quad (2)$$

$$R_{res}(\vec{v}, q) \approx R'_{res}(C(\vec{v}), q) \quad (3)$$

where $C(\vec{v})$ is the value of the SAD or SSD criterion. Computationally even more appealing is to neglect the effect of residual coding, and use criterion

$$J_C(\vec{v}, q) = C(\vec{v}, q) + \lambda_{mv} R_{mv}(\vec{v}) \quad (4)$$

where λ_{mv} is the Lagrange multiplier [2]. Due to its simplicity, this cost function is very attractive when considering for example VLSI implementation.

One problem related to the criterion is how λ_{mv} should be chosen. Chen and Willson [2] used fixed values in their experiments, $\lambda_{mv} = 15$ or 25 for SAD and $\lambda_{mv} = 150$ or 200 for SSD. They note that the value of the parameter is not critical. In [1][4], the choice is tied to the selection of the quantizer parameter q. It is shown both theoretically, using high-rate approximations, and experimentally tha the selection of the Lagrange multiplier for encoder mode selection can be based on

$$\lambda_{mode}(q) = c\, q^2 \quad (5)$$

where c is a constant determined by the functional relationship between the rate and distortion. In practice, $c = 0.85$ is used (it is noted in [4] that approximation is good up to $q = 25$ for H.263). Then, based on experimental results, it is argued that a good choice is to use

$$\lambda_{mv}(q) = \sqrt{\lambda_{mode}(q)} \quad (6)$$

with SAD criterion. In the case of SSD,

$$\lambda_{mv}(q) = \lambda_{mode}(q) \quad (7)$$

is used [1]. For SAD and SSD, these choices give similar results.

However, there seems to be a lack of analysis about why choices should be made using these formulas. In the following, we address this issue and derive another method for determining Lagrange multipliers. The method is based on linear approximations of the encoder inter-mode performance.

3. BASIC FORMULA FOR MULTIPLIER

Assume that the distortion and rate can be approximated in the model-based criterion, $J_M(\vec{v}, q)$, linearly as

$$D'(C, q) = k_D(q)\, C \quad (8)$$

$$R'_{res}(C, q) = k_R(q)\, C \quad (9)$$

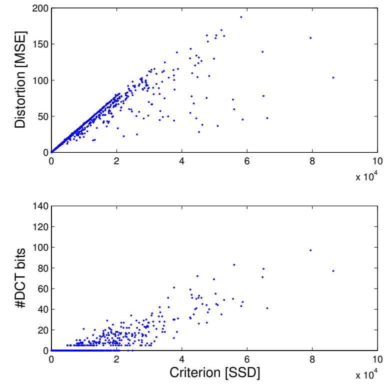

Fig. 1. Subset of data samples for QP = 15.

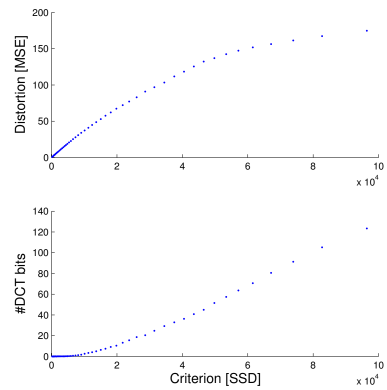

Fig. 2. Average distortion and rate as a function of SSD criterion (QP = 15).

In order to see whether these approximations may be used, (C, D, R_{res}) triplets were gathered using an MPEG-4 encoder[1]. Data and average values are illustrated in Figs 1 and 2 for one QP value. The latter figure shows slightly curved relationships between average values and the criterion. However, as the former figure shows, variation in the rate and distortion for similar prediction errors is high. Based on these observations, and the fact that it is more likely for the encoder to choose intra-mode when the prediction error is higher, the linear approximations may be considered good enough. More accurate approximations may not give significant improvements: according to Chen and Willson [2], the improvement is about 0.1-0.4 dB.

One important note with these approximations is that the model-based approach and residual-neglecting method

[1]Data was collected and also other experiments performed using an MPEG-4 encoder implementation (based on OpenDivX 4.0 alpha 50 release) with an exhaustive search algorithm for full-pel motion estimation and half-pel refinement. Motion was estimated for 16×16 blocks.

(based on (4)) give equivalent results if we choose

$$\lambda_{mv} = \frac{\lambda'}{k_D + \lambda' k_R} \tag{10}$$

as then $J_M(\vec{v}, q) \propto J_C(\vec{v}, q)$.

The primary purpose of rate control is to adjust the quantizer parameter in order to have a lower residual bit rate (or distortion) at the expense of increasing distortion (or residual bit rate). The trade-off is related to the Lagrange multiplier associated with mode selection [1], and we can write

$$\lambda_{mode} = -\frac{\mathrm{d}D}{\mathrm{d}R_{res}} = -\frac{\mathrm{d}D/\mathrm{d}q}{\mathrm{d}R_{res}/\mathrm{d}q} \tag{11}$$

If it were computationally feasible, we would estimate the motion using criterion (1) with the choice

$$\lambda' = \lambda_{mode} \tag{12}$$

and compare the minimum value to the result with other mode choices. In order to use the cost function (4), we combine this choice with linear approximations, which gives us the derivation

$$\begin{aligned} \lambda' &\overset{(12)}{=} \lambda_{mode} \overset{(11)}{=} -\frac{\mathrm{d}D/\mathrm{d}q}{\mathrm{d}R_{res}/\mathrm{d}q} \\ &\overset{(8),(9)}{=} -\frac{\mathrm{d}(k_D C)/\mathrm{d}q}{\mathrm{d}(k_R C)/\mathrm{d}q} = -\frac{\mathrm{d}k_D/\mathrm{d}q}{\mathrm{d}k_R/\mathrm{d}q} \end{aligned}$$

Substituting the result for λ' in (10) gives

$$\lambda_{mv} = \frac{\mathrm{d}k_D/\mathrm{d}q}{k_R\ \mathrm{d}k_D/\mathrm{d}q - k_D\ \mathrm{d}k_R/\mathrm{d}q} \tag{13}$$

which we consider to be the formula for determining the multiplier. As in (6) and (7), the selection is tied to the specific QP value used, but the justifications and the end result are different.

4. SLOPE APPROXIMATIONS

Using equation (13) for determining λ_{mv} involves first an estimation of the functions $k_D(q)$ and $k_R(q)$. In order to do this, macroblock data triples (C, D, R_{res}) are gathered with an encoder using standard SAD or SSD-based motion estimation with different values of q. Then, the slopes are estimated in each case. Here, we calculated the slopes as ratios of sample means.

Specific parametric functions are then fitted to the slope data. For the SSD-based criterion, experiments with the various parametric functions suggested that sufficient approximations are of the form

$$k_D(q) = \sum_{i=0}^{4} a_i q^{i/2} \tag{14}$$

$$k_R(q) = \exp\left(\sum_{i=0}^{4} b_i q^{i/4}\right) \tag{15}$$

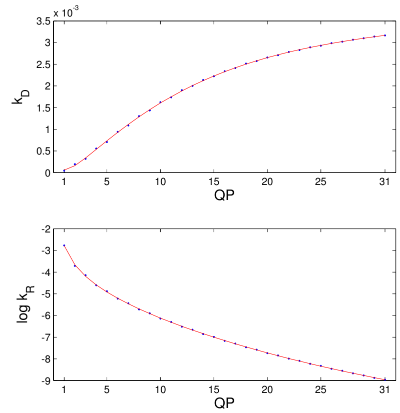

Fig. 3. Slope approximations $k_D(q)$ and $k_R(q)$ for SSD.

where coefficients a_i and b_i are estimated using the slope data. Similar formulae were also used for SAD. Coefficients were found by minimizing an objective function based on the L2 norm, which gave an approximation in Fig. 3 for SSD.

Once the slope functions are available, (13) can be evaluated. The result is illustrated in Fig. 4, where choices of λ_{mv} according to (6) and (7) are shown for comparison. For SAD, the multipliers differ more as the QP value gets higher. For SSD, the curves are similar for high QPs; for lower values, there is some difference.

5. EXPERIMENTS

Experiments were conducted in order to evaluate the encoding performance with different multiplier selection strategies. No R-D optimized mode selection was used.

The results with the *Foreman* sequence are shown in Fig. 5. It can be seen that with the R-D criteria, the results are improved compared to the basic SAD and SSD. With a higher bit rate, the improvement is not significant, about 0.2 dB for 384 kbps. With a lower rate (192 kbps), the gain is about 0.75 dB; in this case, the average QP value is reduced from about 25 to 18.5. In general, the results depend on the content of the sequence as it determines the relative importance of the MV encoding performance.

No multiplier selection method gives a clear advantage over others in Fig. 5, and other tests were made to investigate this issue. For SAD, Fig. 4 suggested especially comparison of our strategy to (6), when the QP value is larger. Similarly, the choice of a fixed $\lambda_{mv} = 175$ was used as a reference method for SSD. In experiments with various sequences, the target bit rates were chosen so that the resulting average QP value were greater than 16. In each case, the average QP value turned out to be lower with our method: in tests with SAD the magnitude of the difference was 0.0-3.1

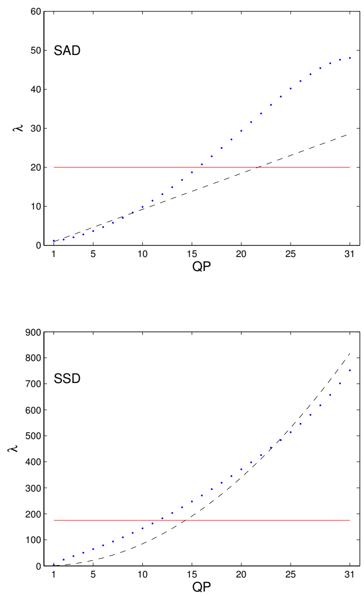

Fig. 4. Values of the Lagrange multiplier with different choice strategies: solid lines are averages of the values used in [2] ($\lambda_{mv} = 20,\ 175$), dashed curves correspond to (6), (7) and dotted ones to the proposed strategy.

units, and with SSD, differences were in the range of 0.1-5.5 units. Also the illustration in Fig. 6 shows small bias in favor of our method.

6. CONCLUSIONS

In this paper, we have studied the selection of the Lagrange multiplier for R-D based block motion estimation criteria. Using the assumption of linear dependence between rate, distortion and prediction error, simple cost functions for the motion estimation were justified, and an approach for finding the multiplier was derived. Analysis gives additional insight to these cost functions, which are attractive from the point of view of applications.

Experimental results indicate that an improvement in terms of PSNR can be obtained using R-D criteria; however, it depends on the content of the video sequence and the target bit rate. Tests with different Lagrange multiplier selection strategies did not show any significant differences between the methods. One can therefore in practice safely select values from a large range.

7. REFERENCES

[1] T. Wiegand and B. Girod, "Lagrange multiplier selection in hybrid video coder control," in *Proc. of Int. Conf. on Image Processing*, 2001, pp. 542–545.

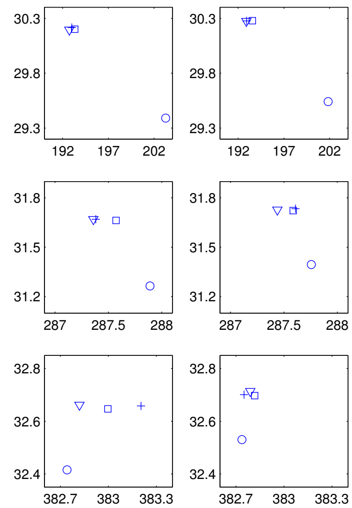

Fig. 5. Result with the *Foreman* sequence (CIF size), for SAD (left) and SSD (right) with different target rates (30 fps). X-axes correspond to actual bit rates and Y-axes to PSNRs. Results with basic SAD/SSD, fixed λ_{mv}, (6)/(7) and our strategy are shown with o, □, + and ▽, respectively.

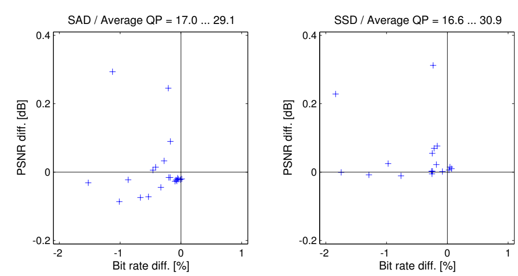

Fig. 6. Change in encoder performance with SAD, when proposed strategy is substituted for method (6), and SSD, when usage of the fixed $\lambda_{mv} = 175$ is replaced by our method. In the figures, the origin corresponds to the result with the reference method.

[2] M.C. Chen and A.N. Willson, "Rate-distortion optimal motion estimation algorithms for motion-compensated transform coding.," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 8, pp. 147–158, 1998.

[3] W.C. Chung, F. Kossentini, and M.J.T. Smith, "An efficient motion estimation technique based on a rate-distortion criterion," in *Proc. of IEEE Int. Conf. on Acoustics, Speech, and Signal Processing*, 1996, pp. 1926–1929.

[4] G.J. Sullivan and T. Wiegand, "Rate-distortion optimization for video compression," *IEEE Signal Processing Magazine*, vol. 15, pp. 74–90, 1998.