AN ADAPTIVE BLOCK-MATCHING ALGORITHM FOR MOTION ESTIMATION

Vasily G. Moshnyaga

Department of Electronics and Computer Science, Fukuoka University
8-19-1 Nanakuma, Jonan-ku, Fukuoka 814-0180, JAPAN

ABSTRACT

A new adaptive algorithm for the block matching motion estimation is presented. The algorithm works in the full-search fashion but unlike the FSBMA it adjusts the number of computations dynamically to picture variation. Due to incorporated mechanism of data-driven thresholding, the proposed approach performs as four times as less operations comparing to the FSBMA while maintaining the same quality of results. Its hardware implementation is simple and compact. A supportive hardware design as well as simulation results on benchmarks are outlined.

1. INTRODUCTION

1.1. Motivation

Motion estimation is a basic bandwidth compression task utilized in video-coding systems. Among various computation methods[1], the *Full Search Block Matching Algorithm* (FSBMA) is most popular. Having successive video frames divided into blocks of $(N \times N)$ pixels, the FSBMA determines a *motion vector* (v) for every *reference block* (X) of the current image by comparing it with all *candidate blocks* $(Y(0,0), Y(0,1), ...,)$ within the search area surrounding the position of the reference block in the previous frame. Let $x(i,j)$ be the luminance value of the reference block pixel, $y(i,j)$ the luminance value of the candidate block pixel, p the maximum displacement allowed in both vertical and horizontal directions. Then, the position (m,n) of a candidate block $Y(m,n)$ that results in the minimum distortion D denotes the motion vector v:

$$D(m,n) = \sum_{i=0}^{N-1}\sum_{j=0}^{N-1} |y(i+m, j+n) - x(i,j)|, \quad -p \le m,n \le p-1 \qquad (1)$$

$$v = \arg \min_{-p \le m,n \le p-1} D(m,n) \qquad (2)$$

The FSBMA provides optimal precision, regular data flow as well as higher parallelism, a characteristic that is advantageous for VLSI implementation. However, it is extremely time consuming because $N^2 \times (2 \times p)^2$ computations of the distortion (D) have to be performed per each reference block. If a frame has 720×480 pixels, $p = 16$, $N = 16$ (MPEG2, MP@ML complexity level), the FSBMA requires over 11 GOPS.

Due to such an enormous computational rate, existing hardware implementations of the FSBMA[2] are extremely power hungry: over a half of energy dissipated in a modern encoder is burned in the motion estimation hardware! As result, algorithms and architectures which ensure low power operation have become very important, especially for portable video application.

1.2. Related research

Over the years, a large variety of fast and computationally inexpensive block-matching algorithms have been proposed. Examples include the 2D-logarithm search[3], the three-step search[4] and its modification[5], the conjugate direction search [6], the cross-search[7], etc. These algorithms search only a small subset of available candidate blocks and consequently execute less computations. However, they lack in terms of Peak Signal to Noise Ratio (PSNR), i.e. subjective picture resolution. Therefore, recent research attempts have been put on decreasing the operational complexity of the FSBMA.

An approach to reduce the FBMA complexity is to transform the 8-bit gray-scale data into binary numbers and then use binary level distortion metric instead of multi-bit arithmetic, as proposed in [8],[9]. This approach, as well as the LSB-bit truncation scheme[10], minimizes the amount of energy dissipating switches during the distortion calculation. However, it does not affect the amount of operations involved in the FSBMA. To lower the number of operations in the FSBMA, work [11] suggest the dynamic search range adjustment to the picture content variation. The idea is to run the FSBMA using a large p over a number of frames and then shrink p, if possible, to accommodate 95% of motion vectors. The technique achieves a quadratic reduction in operations but restricts itself to the highly correlated video sequences. Moreover, it challenges the FSBMA's optimality which is not acceptable. Up to our knowledge there have been reported only one method[12] capable of reducing the FSBMA complexity without sacrificing its accuracy. Before computing the exact distortion, the method performs a conservative estimation of the distortion value. If the estimate is larger than the minimum distortion found so far, the exact distortion value is not calculated and the corresponding computations not performed. The approach can halve the computational count, but at the cost of extra large hardware overhead.

1.3. Contribution

In this paper, we propose a novel algorithm for adaptive elimination of unnecessary computations in the full-search block matching. In contrast to FSBMA, we dynamically