Effects of Reduced Precision on Video Motion Estimation

M Panovic and A Demosthenous

Department of Electronic and Electrical Engineering, University College London

Abstract: This paper investigates the effects of estimating motion between video frames at reduced precision (< 8-bits). Simulation results show that for common video sequences, the mean square distance of 6-bits compares very favourably to the mean absolute distance of 8-bits used in digital realisations. Since the mean square distance lends itself particularly well to implementation using analogue CMOS circuits, the reduced precision of 6-bits offers potential for the implementation of small, low power, analogue motion estimators for portable video systems.

1 Introduction

Image compression is an essential technology in the realisation of the digital information highway because it enables bit rate reduction. Several international standards have been established in recent years to deal with various aspects of digital video encoding. Examples include H.261 for video conferencing, MPEG-1 for VHS quality video, MPEG-2 for broadcast video quality, and MPEG-4 for very low bit rate audiovisual coding [1], [2].

All these standards employ both spatial and temporal correlations to achieve data compression. The first technique addresses individual images and seeks to reduce the similarity of pixels within a frame by employing transform domain coding. The second technique seeks to remove similarities between successive frames, by coding their differences. While interframe differencing yields good compression in static parts of the image sequence, it can cause data expansion in those parts which change between frames. Thus, interframe coding is usually augmented by motion compensation. The basic mechanism is that before the differencing stage one of the frames is pre-distorted according to the motion that occurs between the two frames. The distortion is expressed as a set of motion vectors which are calculated by a process called *motion estimation*. The optimum distortion measure (cost function) is the mean squared distance (MSD). However, to ease implementation complexity all digital designs of motion estimators employ the *non-optimum* mean absolute distance (MAD) of 8-bits [3]. Nevertheless, the motion estimator is by far the most expensive component of a digital video codec in terms of size, complexity and power dissipation.

This paper investigates the effects on picture quality for two common video sequences with different levels of motion, when motion is estimated at less than 8-bits using the MSD. In both cases the 6-bit MSD performs similarly to the 8-bit MAD. However, such reduction in precision offers the potential for the use of low power, low complexity, analogue CMOS circuits by exploiting the square-law characteristic of MOS transistors in saturation [4].

2. Principles of Motion Estimation

Motion estimation creates a model of the current frame based on available data in one or more previously encoded frames. The technique used in most video codecs is block matching motion estimation as depicted in Fig. 1. Each of the two frames is divided into template blocks (TBs) of $N \times N$ pixels and for a maximum motion displacement of w pixels per frame, the current TB is matched against a corresponding block at the same co-ordinates but in the previous frame within the search window (SW) of $N + 2w$. The best match on the basis of minimising a cost function yields the displacement (motion vector). The cost functions of the type MSD and MAD evaluated in this paper are defined as follows:

for MSD:

$$MSD(i, j) = \frac{1}{N^2} \sum_{m=1}^{N} \sum_{n=1}^{N} \left(f(m, n) - g(m+i, n+j) \right)^2, \quad -w \le i, j \le w \tag{1}$$

and for MAD:

$$MAD(i, j) = \frac{1}{N^2} \sum_{m=1}^{N} \sum_{n=1}^{N} |f(m, n) - g(m+i, n+j)|, \quad -w \le i, j \le w \tag{2}$$

where $f(m,n)$ represents the current block of N^2 pixels at co-ordinates (m,n) and $g(m+1, n+j)$ represents the corresponding block in the previous frame at new co-ordinates $(m+i, n+i)$.

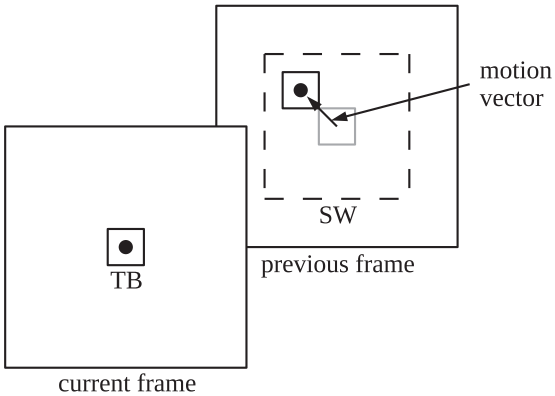
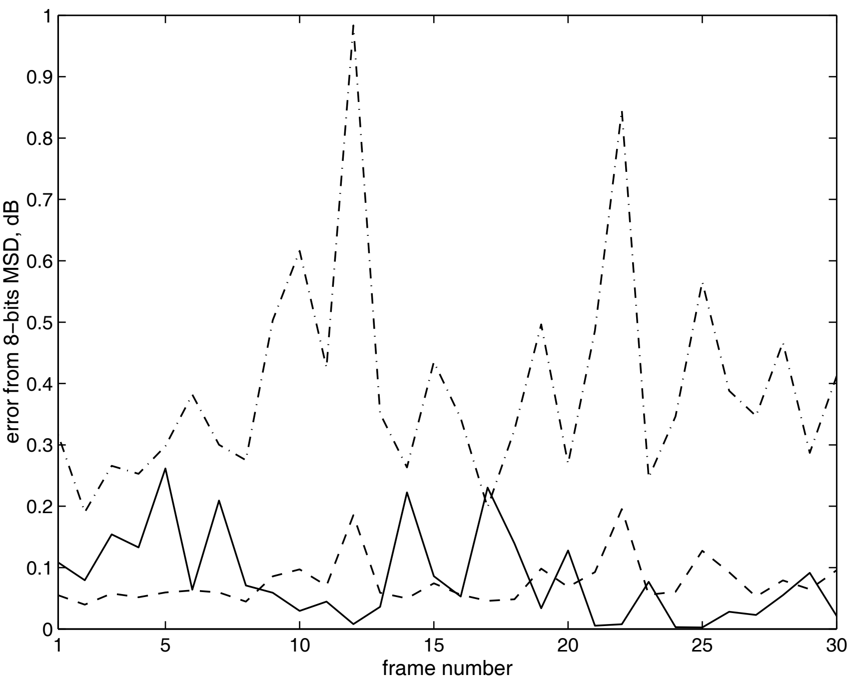


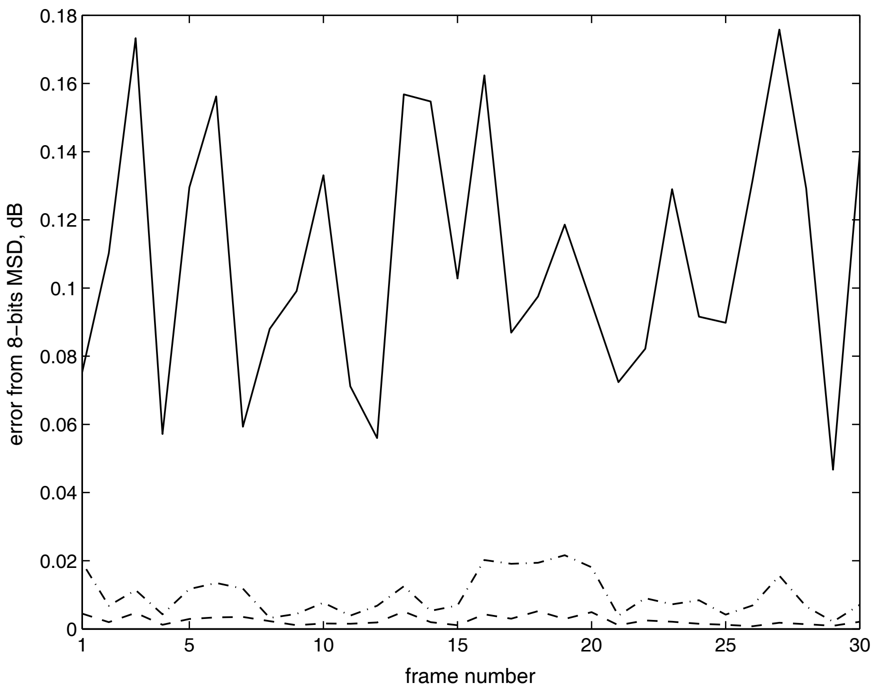
Fig. 1. Block matching motion estimation.

3. Simulation Results

Two standard video sequences of resolution 352 by 288 pixels were considered: a) "Claire", videoconference sequence with movement restricted to the face area of the speaker and with a fixed background, and b) "Foreman", sequence with large motion in all directions and with some camera motion. The simulations were carried out using the MSD (1) and MAD (2) cost functions at various precision levels, and the peak-to-peak signal-to-noise ratio (PSNR) was used to evaluate the objective quality of the motion estimated frames. For both sequences, full-search motion estimation was performed using a TB of 16 by 16 pixels and a SW of 32 by 32 pixels.



(a)



(b)

Fig. 2. Simulation results; (a) "Claire" video sequence, (b) "Foreman" video sequence.

——— MDA 8-bits — — — — — MSD 6-bits —·—·—· MSD 5-bits

Fig. 2a shows the deviation (error) in PSNR from 8-bits MSD precision for the first 30 frames of the "Claire" sequence for the MSD of 6-bits and 5-bits and also for the MAD of 8-bits. As can be seen from the plot, the MSD of 6-bits performs very similarly to the MAD of 8-bits, however, below 6-bits the MSD degrades rapidly. This is because of the low intensity between adjacent areas in the frames of this sequence. On the other hand, the results of the "Foreman" sequence in Fig. 2b, show that the MSD even at 5-bits precision outperforms the MAD of 8-bits. This is because of the high intensity between adjacent areas in the frames of this sequence.

4. Conclusion

It has been shown by computer simulations that for common video sequences, the MSD at much reduced precision compares very favourably to the MAD at full precision. This offers an advantage for the design of an analogue motion estimation processor by exploiting the square-law properties of MOS transistors in saturation. Such a motion estimator would be much smaller and less power hungry than existing digital realisations, making it ideal for low power portable video systems.

Acknowledgments

This work was supported by EPSRC under grant number GR/R64926.

References.

[1] M. Ghanbari, *Video Coding: An Introduction to Standard Codecs*. The Institution of Electrical Engineers, London, 1999.

[2] P. Kuhn, Algorithms, *Complexity Analysis and VLSI Architectures for MPEG-4 Motion Estimation*. Kluwer Academic Publishers, The Netherlands, 1999.

[3] M. Mizuno et al., "A 1.5W single chip MPEG-2 MP@ML video encoder with low power motion estimation and clocking," *IEEE J. Solid-State Circuits*, vol. 32, pp. 1807-1816, Nov. 1997.

[4] K. Bult and H. Wallinga, "A class of analog CMOS circuits based on the square-law characteristic of an MOS transistor in saturation," *IEEE J. Solid-State Circuits*, vol. SC-22, pp.357-365, June 1987.