J Real-Time Image Proc
DOI 10.1007/s11554-013-0383-9

SPECIAL ISSUE PAPER

A low energy adaptive motion estimation hardware for H.264 multiview video coding

Yusuf Aksehir · Kamil Erdayandi · Tevfik Zafer Ozcan · Ilker Hamzaoglu

Received: 26 January 2013 / Accepted: 16 November 2013

Abstract Multiview video coding (MVC) is the process of efficiently compressing stereo (two views) or multiview video signals. The improved compression efficiency achieved by H.264 MVC comes with a significant increase in computational complexity. Temporal prediction and inter-view prediction are the most computationally intensive parts of H.264 MVC. Therefore, in this paper, we propose novel techniques for reducing the amount of computations performed by temporal and inter-view predictions in H.264 MVC. The proposed techniques reduce the amount of computations performed by temporal and inter-view predictions significantly with very small PSNR loss and bit rate increase. We also propose a low energy adaptive H.264 MVC motion estimation hardware for implementing the temporal and inter-view predictions including the proposed computation reduction techniques. The proposed hardware is implemented in Verilog HDL and mapped to a Xilinx Virtex-6 FPGA. The FPGA implementation is capable of processing $30 \times 8 = 240$ frames per second (fps) of CIF (352×288) size eight view video sequence or $30 \times 2 = 60$ fps of VGA (640×480) size stereo (two views) video sequence. The proposed techniques reduce the energy consumption of this hardware significantly.

Y. Aksehir · K. Erdayandi · T. Z. Ozcan · I. Hamzaoglu (✉)
Faculty of Engineering and Natural Sciences, Sabanci University, Tuzla, 34956 Istanbul, Turkey
e-mail: hamzaoglu@sabanciuniv.edu

Y. Aksehir
e-mail: yusufaksehir@sabanciuniv.edu

K. Erdayandi
e-mail: kerdayadi@sabanciuniv.edu

T. Z. Ozcan
e-mail: tzaferozcan@sabanciuniv.edu

Keywords H.264 · Multiview video coding · Motion estimation · Hardware implementation · FPGA

1 Introduction

Since the recently developed H.264 video compression standard has significantly better video compression efficiency than previous video compression standards, it is already started to be used in many consumer electronic devices [1, 2]. Motion estimation (ME) is used for compressing a video by removing the temporal redundancy between the video frames. Since it constitutes up to 70 % of the computations performed by a video encoder, it is the most computationally intensive part of a video encoder hardware. The improved compression efficiency achieved by ME in H.264 standard comes with an increase in computational complexity.

Block matching (BM) is used for ME in H.264 standard. BM partitions the current frame into non-overlapping N × N rectangular blocks and finds a MV for each block by finding the block from the reference frame in a given search range that best matches the current block. Sum of Absolute Differences (SAD) is the most preferred BM criterion. The SAD value of a search location defined by the motion vector $d(d_x, d_y)$ is calculated as below where $c(x, y)$ and $r(x, y)$ represent current and reference frames, respectively. The coordinates (i, j) denote the offset locations of current and reference blocks of size N × N.

$$\mathrm{SAD}\left(\vec{d}\right) = \sum_{x=0}^{N-1}\sum_{y=0}^{N-1}\left|c(x+i,y+j) - r(x+i+d_x,y+j+d_y)\right|$$

Multiview video coding (MVC) is the process of efficiently compressing stereo (two views) or multiview

Published online: 11 December 2013