J Real-Time Image Proc
DOI 10.1007/s11554-013-0362-1

SPECIAL ISSUE PAPER

Fast mode decision algorithm for H.264/SVC enhancement layer

Amina Kessentini · Amine Samet · Mohamed Ali Ben Ayed · Nouri Masmoudi

Received: 29 November 2012 / Accepted: 7 June 2013

Abstract Scalable video coding (SVC) has been standardized to extend the capabilities of the H.264 advanced video coding (AVC). The SVC can compress several video sequences of various resolutions as a single bit-stream. In the SVC enhancement layer, for Joint Scalable Video Model (JSVM) software implementation, an exhaustive mode decision process based on the base layer mode predictions is performed to obtain the best mode for each macroblock (MB). This technique may achieve a higher coding efficiency; however, it induces a significant computational complexity in the encoding engine. In order to speedup the SVC encoder, a fast mode decision algorithm was proposed in this paper. In other words, our aim was to decrease the number of candidate modes to reduce the computational complexity and maintain the same level of coding efficiency, this approach used the spatial and temporal correlation between MB situated at the enhancement layer and its co-located MB at the base layer. Our statistical analyses were made using several HD sequences with different motion characteristics. Experimental results show a significant improvement in terms of time encoding which is a major constraint for any real-time implementation. However, this gain is accompanied with an acceptable loss in video quality and a tolerable bit rate increase for most media supports. In fact, our proposed algorithm permits a major improvement that can reach up to 64.9 % in terms of computational effort. This gain will induce an average loss yield to 10.5 or 13.87 % that is comparable to the 13.12 % of the He Li's algorithm with an acceptable loss in terms of subjective video quality.

Keywords H.264/SVC encoder · Mode decision · Spatial correlation · Temporal correlation

A. Kessentini (✉) · A. Samet · M. A. Ben Ayed · N. Masmoudi
Electronics and Information Technology Laboratory,
National Engineering School of Sfax, University of Sfax,
Sfax, Tunisia
e-mail: ksontini.amina@gmail.com

N. Masmoudi
e-mail: Nouri.Masmoudi@enis.rnu.tn

A. Samet · M. A. Ben Ayed
High Institute of Electronics and Communication Sfax,
University of Sfax, Sfax, Tunisia

1 Introduction

H.264/SVC is developed by the joint video team (JVT) as an extension of H.264/advanced video coding (AVC) [1–4]. In fact, many features or tools are introduced to get a better coding efficiency which causes a major increase in terms of video encoder computational complexity [5, 6]. In comparison with previous video coding standards, the purpose of the scalable video coding is to encode the signal once, but enables its decoding from partial streams depending on the specific rate and resolution required by certain applications [7]. Figure 1 shows the H.264/SVC encoder with three spatial layers. The base layer (BL) holds a lower resolution or a reduced quality version of each coded frame. For the enhancement layers (ELs), Layer 1 and Layer 2, the input is in a higher resolution and will be coded as an ordinary H.264/AVC along with inter-layer predictions that will provide additional coding choices, such as inter-layer motion vectors, intra-prediction, and residual information deduced from the BL [8]. SVC presents three types of scalability [3, 4, 9, 10]: temporal, quality, and spatial scalability. Temporal scalability in the SVC is achieved using a structure of hierarchical B pictures inherited from H.264/AVC standard without any additional

Published online: 25 June 2013