Chapter 3
Performance Scalable Motion Estimation for Video Coding: An Overview of Current Status and a Promising Approach

Golam Sorwar
Southern Cross University, Australia

Manzur Murshed
Monash University, Australia

ABSTRACT

Motion estimation is one of the major bottlenecks in real-time performance scalable video coding applications due to high computational complexity of exhaustive search. To address this, researchers so far focused on low-complexity motion estimation and rate-distortion optimization in isolation. Proliferation of power-constrained handheld devices with image capturing capability has created demand for much smarter approach where motion estimation is integrated with rate control such that rate-distortion-complexity optimization can be effectively achieved. It is indeed crucial to provide such performance scalability in motion estimation to facilitate complexity management in such devices. This chapter presents an overview of motion estimation. Beginning with an introduction to the importance of motion estimation, it systematically examines various motion estimation techniques and their strengths and weaknesses, focussing primarily on block-based motion search. It then examines the limitation of the existing techniques in accommodating performance scalability, introduces a promising approach, Distance-dependent Thresholding Search (DTS) motion search, to fill in this gap, and concludes with future research directions in the field. The authors suggest that the content of the chapter will make a significant contribution and serve as a reference for multimedia signal processing research at postgraduate level.

DOI: 10.4018/978-1-4666-2660-7.ch003

INTRODUCTION

Digital video technology has been facing by an exponential growth in the last few years driven by new and fundamental applications such as videoconferencing, net-meeting, video e-mail, video over mobile phone (3G), wireless multimedia applications and video streaming over personal digital assistants. These new and emerging multimedia applications have driven the need for more advanced video coding standards. Currently, there are several video standards, such as H.261/3/4/ AVC established for a variety of multimedia applications. Video encoding, however, involves a huge amount of computations and hence, there is significant interest and need to speed up the processing of video encoders. Video coding exploits temporal redundancy in order to reduce the bandwidth while preserving the quality of the receiver-reconstructed images. This has resulted in many motion based video compression strategies which have become an integral part of multimedia applications for both communication and entertainment purposes. Most work in this area has been mainly focused on the fast motion estimation algorithms and attempted tooptimise only *rate-distortion* performance on scalable video coding. However, it is indeed crucial to provide performance scalability in motion estimation in facilitating complexity management in video coding, especially in real-time software-only low bit rate video CODECs (Coder and Decoder) or low-power video CODECs for mobile or hand-held computing platforms which particularly require a more flexible trade-off between complexity and quality (Richardson, 2002).

In this chapter, we present an overview of motion estimation techniques, providing a snapshot of its current status, focussing primarily on block-based motion estimation algorithms (BMAs). We identify the limitations of BMAs in accommodating performance scalability in real-time video coding applications and then present a promising approach, *Distance-dependent Thresholding Search (DTS) motion search,* to fill in this gap. Finally the chapter concludes with future research directions in the field.

BACKGROUND

A video sequence is a much richer source of visual information than a still image, due to the capture of motion. While a single image provides a snapshot of a scene, a sequence of images (widely termed as frames) register the dynamics within it. The registered motion is a very strong cue for human vision to recognise objects as soon as they move even if they are inconspicuous when still. Motion is, therefore, the most obvious and effective feature to provide global and local understanding as well as describing the dynamic content within a video sequence. The extraction of motion information from sequences of time-varying images has numerous applications in a wide range of areas especially computer vision and image processing. Perhaps the most important application of motion is in video data compression. In the evolving digital technology era, video compression has become an integral part of multimedia applications for both communication and entertainment purposes. As the diversity of these applications indicate, *Motion estimation* (ME) has been the focus of extensive research over many years, and this is reflected in the plethora of motion estimation and analysis techniques that have been proposed. Existing motion estimation techniques may be broadly classified into three distinct classes which will be discussed briefly below.

Gradient-Based Motion Estimation

Pixel-based motion estimation is a gradient-based method (Aggarwal & Nandhakumar, 1988; Lucas & Kanade, 1981; Barron et al., 1994) which focuses on estimating the apparent motion of intensity patterns in a video sequence, known as *optical flow,* based on two assumptions. Firstly,