Real-Time Motion Estimation by Object-Matching for High-Level Video Representation

Aishy Amer Amar Mitiche Eric Dubois

INRS-Télécommunications Université d'Ottawa
Montréal, H5A 1C6 Canada Ottawa, K1N 6N5 Canada

E-mail: amer@inrs-telecom.uquebec.ca

Abstract

Motion estimation plays a key role in many video applications, such as frame-rate video conversion, video retrieval, video surveillance, and video compression. The key issue in these applications is to define appropriate representations that can efficiently support motion estimation with the required accuracy. In this paper, a low-complexity object motion estimation technique is proposed that is designed to fit the needs of high-level video representation such as in video surveillance or retrieval applications. In these applications, a representation of object motion in a way meaningful for high-level interpretation, such as event detection and classification, foregoes precision of estimation. The proposed method relies on the estimation of the displacements of the minimum bounding box (MBB) sides of an object. Two motion estimation steps are proposed: initial coarse estimation to find a single displacement for an object using the four sides of the MBB between two successive images and detection of non-translational motion and its estimation. The result is the detection of the type of object motion and the subsequent estimation of one or more motion values per object depending on the detected motion type. Special consideration is given to object motion in interlaced video and at image margin. Various simulations show that the proposed method provides a response in real-time and gives good estimates to use for object tracking, event detection, and high-level video representation. The proposed object motion estimation method is insensitive to inaccurate segmentation in these applications.

1 Introduction

Objects can be classified into three major categories: rigid, articulated, and non-rigid [7]. The motion of a rigid object is a composition of a translation and a rotation. An articulated object consists of rigid parts linked by joints. Most video applications, such as entertainment, surveillance, or retrieval, assume rigid objects.

An image acquisition system projects a 3-D world scene onto a 2-D image plane. When an object moves its projection is animated by a 2-D motion, to be estimated from the space-time image variation. These variations can be divided into global and local. Global variations can be a result of camera motion or global illumination change. Local variations can be due to object motion, local illumination change and noise. Motion estimation techniques estimate apparent motion which is due to true motion or to various artifacts, such as noise and illumination change. The goal of a motion estimation technique is to assign a motion vector (displacement or velocity) to each pixel in an image. Motion estimation relies on hypotheses about the nature of the image or object motion, and is often tailored to applications needs. Difficulties in motion estimation arise from unwanted camera motion, occlusion, noise, lack of image texture, and illumination changes. Motion estimation is an ill-posed problem which requires regularization. A problem is ill-posed if no unique solution exists or the solution does not continuously depend on the input data.

The choice of a motion estimation approach strongly depends on the application and on the nature of the processes that will interpret the estimated motion. A key issue when designing a motion estimation technique is its degree of efficiency with enough accuracy to serve the purpose of intended application. For instance, in video surveillance and retrieval, a tradeoff is required between computation cost and quality of results. This paper proposes a real-time method to estimate 2-D object motion from a video using segmented object. The method aims at representing object motion in a way meaningful for high-level applications, e.g., event classification, that foregoes precision of estimation.

The paper is organized into five additional sections. Section 2 discusses related work, Section 3 discusses models for object-based motion estimation, Section 4 describes a real-time motion estimation method based on extracted object, Section 5 presents experimental results, and Section 6 concludes the paper.