Accepted Manuscript

Video analytical coding: When video coding meets video analysis

Yuyang Liu, Ce Zhu, Min Mao, Fangliang Song, Frederic Dufaux, Xiang Zhang

PII: S0923-5965(18)30498-3
DOI: https://doi.org/10.1016/j.image.2018.05.012
Reference: IMAGE 15392

To appear in: *Signal Processing: Image Communication*

Received date : 27 October 2017
Revised date : 5 April 2018
Accepted date : 17 May 2018

Please cite this article as: Y. Liu, C. Zhu, M. Mao, F. Song, F. Dufaux, X. Zhang, Video analytical coding: When video coding meets video analysis, *Signal Processing: Image Communication* (2018), https://doi.org/10.1016/j.image.2018.05.012

Video Analytical Coding: When Video Coding Meets Video Analysis

Yuyang Liu[a], Ce Zhu[a], *Fellow, IEEE*, Min Mao[a], Fangliang Song[a],
Frederic Dufaux[b], *Fellow, IEEE*, Xiang Zhang[a]
[a]University of Electronic Science and Technology of China, Chengdu, China
[b]Laboratoire des Signaux et Systèmes, Paris, France

Abstract

Leveraging on the properties of human visual system, most of the well-designed video coding standards utilize rate-distortion optimization techniques by maximizing a fidelity cost function (e.g. peak signal noise ratio, PSNR) under an available bit rate budget constrain. However, a huge amount of video data is consumed by computers rather than by human beings in several application scenarios. In view of this, this paper proposes a new coding framework called video analytical coding (VAC) for video analysis. We use the term "analytical distortion" to denote the difference of video analysis performance when video quality degrades and analytical distortion is estimated by compression distortion. Meanwhile, we develop a new rate-analytical-distortion optimization (RADO) method to trade off the bit rate and the analytical distortion. Specifically, we consider moving object detection as the analysis task and develop a novel rate analytical distortion (RAD) model and a quantization parameter adaptation strategy for video coding, where the analytical distortion is related to the object detection performance represented as F1-measure. Experimental results show that the performance of the video analysis task can be significantly improved (up to 40% reduction of analytical distortion).

Index Terms: Video analysis; video analytical coding; analytical distortion; rate distortion optimization.

1 Introduction

The increasing availability of portable or installed cameras and the introduction of new multimedia applications to fulfill emerging needs, have given rise to new requirements on video compression and communication. As for many multimedia applications, e.g. surveillance, video content is not only presented to human beings but also analyzed by computers for variously applicable purposes, such as object detection, tracking, recognition, and so on. In other words, computers have become as viewers of the videos. Meanwhile, the considerable amount of generated videos need to be efficiently compressed due to the cost-effective storage and bandwidth limitation.

Currently, the main goal of most studies on video coding is to achieve high coding efficiency. Most of the widely deployed video coding standards, such as H.264/MREG-4 AVC (Advanced Video Coding) [1] and HEVC (High Efficiency Video Coding) [2], are designed under the assumption that human beings are the target viewers. Meanwhile, traditional rate distortion optimization (RDO) framework is applied into the video coding standards by optimizing the trade-off between the entropy of the discretized representation (rate) and the error arising from the quantization (distortion). However, applying the traditional RDO framework during video compression may be suboptimal when the video is intended for machine analysis. The critical issue is that resulting compression distortion may bring a