Adaptive Long-term Reference Selection for Efficient Scalable Surveillance Video Coding

Le Dao Thi Hue, Giap PhamVan, Xiem HoangVan
VNU – University of Engineering and Technology
hueledao94@gmail.com, giap_pham@outlook.com, xiemhoang@vnu.edu.vn

Abstract

The exponential growth of video surveillance has been asking for a more powerful video coding solution, which is characterized by not only the high compression efficiency but also the adaptive video streaming capability. The surveillance video content, however, usually contains a large number of background areas and having high temporal correlation between frames. In this context, we propose a novel adaptive long-term reference mechanism for scalable surveillance video coding, which provides the quality and temporal scalabilities while achieving the high compression performance. The proposed long – term reference is mainly selected based on the content analysis of video sequence. The long-term reference selection solution is integrated into the most recent Scalable High Efficiency Video Coding (SHVC) standard. Experiments conducted for a rich set of surveillance videos show that the proposed scalable video coding solution can achieve around 5.38% bitrate saving when compared to the traditional SHVC video coding benchmark.

Keywords: Surveillance scalable video coding, SHVC standard, long – term reference, bitrate saving

1. Introduction

In recent years, there has been an accelerated expansion of surveillance systems to cope with security and safety's threats. Considerable numbers of surveillance cameras have been mounted in public and private areas. The emergence of large video surveillance infrastructures leads to a massive amount of content that must be stored, analyzed and managed by security teams with limited resources. Furthermore, the heterogeneity of networks, display devices, and transmission environments has been rising as a critical issue in modern video communication era [1]. To fulfill these challenges, it is necessary to have an efficient and adaptable surveillance video compression system, which provides not only the compression efficiency but also the adaptability to the network and transmission variation.

The recent achievements of video coding technology have resulted in a new video coding solution, namely High Efficiency Video Coding (HEVC) [2]. As reported, HEVC significantly outperforms the well-known H.264/AVC standard [3]. For adaptive video streaming, the HEVC scalable extension, namely SHVC, has been introduced in 2014 [4]. SHVC is mainly designed based on a layered coding structure in which one base layer (BL) is used to compress the video sequence with low and basic quality / resolution fidelities and one or several enhancement layers (EL) is used to provide enhanced quality/ resolution fidelities.

Though SHVC is the latest scalable video coding standard, its compression performance is still an emerging topic for research and development. The work in [5] proposed an improved EL merge mode while the work in [6] proposed a novel joint layer prediction solution. As reported [5, 6], the proposed EL merge mode and joint layer prediction significantly improve the SHVC compression performance, far beyond the SHVC standard. However, none of these proposals is designed for visual surveillance systems as it is mainly created for the generic video content. In a visual surveillance system, cameras are usually placed at a certain position or moved with a very narrow angle. Therefore, the surveillance video content usually contains a large area of background as well as having a high temporal correlation between frames.

To exploit these characteristics, we propose in this paper an improved SHVC compression solution, which is designed for surveillance video content. The proposed surveillance scalable video coding (SSVC) is created based on the use of an adaptive long – term reference selection and updating mechanism. The video content is carefully analyzed before using for selecting and updating the reference picture. Experimental results have shown that the proposed SSVC solution significantly outperforms the relevant SHVC standard, notably with around 5.38 % bitrate saving while still providing a similar perceptual decoded frame quality.

The rest of this paper is organized as follows. Section 2 briefly discusses the related and