

Chapter 79

Application of Computer Vision Techniques for Exploiting New Video Coding Mechanisms

Artur Miguel Arsenio

Nokia Siemens Networks SA, Portugal & Universidade Tecnica de Lisboa, Portugal

ABSTRACT

One of the main concerns for current multimedia platforms is the provisioning of content that provides a good Quality of Experience to end-users. This can be achieved through new interactive, personalized content applications, as well by improving the image quality delivered to the end-user. This chapter addresses these issues by describing mechanisms for changing content consumption. The aim is to give Application Service Providers (ASPs) new ways to allow users to configure contents according to their personal tastes while also improving their Quality of Experience, and to possibly charge users for such functionalities. The authors propose to employ computer vision techniques to produce extra object information, which further expands the range of video personalization possibilities on the presence of new video coding mechanisms¹.

INTRODUCTION

Telecommunication operators need to deliver their clients not only new profitable services, but also good quality, personalized and interactive content. This chapter addresses mechanisms for transforming image content on videos (video clips, video stream), in the form of objects, into other objects as selected by end-users using the à priori availability of 3D models of objects (which we denote content impersonation). The chapter presents possible objects representations, according to a set of transmission parameters, so that it is transmitted not only information concerning image content, but also information concerning how objects in such image (within a video stream) can be transformed. We also address the efficient transmission of such content over a multimedia distribution network, describing a methodology that exploits new object-oriented video coding mechanisms.

DOI: 10.4018/978-1-5225-5204-8.ch079

Another important factor driving users' quality of experience, besides multimedia content personalization and interactivity, is the transparent reception of multimedia content over fixed-mobile convergent networks. Nowadays, terminal devices have different processing capabilities – often, home devices have considerable processing power and higher display resolution, while mobile terminals, such as mobile phones, have smaller displays. For content recording, mobile phones often produce video streams of lower resolution than the desirable ones to see on a large home set. So, it is desirable to increase resolution to specific image objects in order to better view the later in larger displays with higher resolution. We will further show how computer vision techniques can be employed to improve video consumers' Quality of Experience on a multimedia distribution network, while simultaneously expanding the content generation possibilities from users' mobile devices, and the correspondent content consumption. Such techniques support the convergence of mobile and fixed content, increasing objects' resolution, and recovering from transmission errors in order to provide the end-user a better service. Therefore, we not only exploit object based coding on new video coding mechanisms, but also scalable video coding allowing the end-device to select among a set of scales the resolution that best fits its capabilities.

Current solutions are appropriate for merging video segments from different sources (in slices), or by removing segments (such as advertisement segments), or even replacing some segments by others. The other main application of current solutions is to respond to events on a video stream (e.g. upon appearance of a certain object, to display information concerning this object).

The closest solutions already in the market include Content Management Systems, used either for simply managing content provided by external entities or to provide some removal operations such as advertisement removal for videos. These solutions have the disadvantages that they are not able to adapt video content to users, by merging other data to create new content or transforming existing one – their power rely mostly on the removal of features and on adding descriptive texts (e.g. Biography of a soccer player on soccer games, which may be event triggered, such as the push of a button by a user).

This chapter describes solutions that allow ASPs to provide users automatic ways to adapt and configure the (online, live) content to their tastes—and even more—to manipulate the content of live (or offline) video streams (in a way that PhotoShop did for images or AdobePremiere, into a certain extent, to offline videos).

Indeed, a telecommunication provider or operator delivers to its client various services according to different quality of Service guarantees, sometimes dependent on the bandwidth available in the respective communication network. The objective of the approach provided in this chapter is in particular to enable an operator or application service provider to offer additional applications and/or services to a user that are of high interest for this user and allows a high degree of personalization to adjust the content to his or her individual taste. In order to overcome this problem, a method for processing a data stream is provided comprising the steps of: 1) identifying at least one object within a data stream that is tagged as transformable; 2) transformation of object within such stream. This new approach allows an Application Service Provider (ASP) to provide users with the possibility to adapt and/or configure a data stream, e.g., a TV-broadcast, a video on demand, an online or live content to their particular preferences and in particular to further manipulate the content thereof.

This approach in particular relates to content manipulation of data streams comprising in audio and/or video content. The data stream may be provided in or via a network (such network may comprise for instance an IP network). The user is connected to such IP network via any access network, comprising, e.g., a Digital Subscriber Line (DSL), a Radio Access Network (RAN), cable network, or the like.

27 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:

www.igi-global.com/chapter/application-of-computer-vision-techniques-for-exploiting-new-video-coding-mechanisms/197029

Related Content

Patent Infringement Risk Analysis Using Rough Set Theory

Chun-Che Huang and Hao-Syuan Lin (2011). *Visual Analytics and Interactive Technologies: Data, Text and Web Mining Applications* (pp. 123-150).

www.irma-international.org/chapter/patent-infringement-risk-analysis-using/48394

A Novel Approach of Restoration of Digital Images Degraded by Impulse Noise

Rashmi Kumari, Anupriya Asthana and Vikas Kumar (2014). *International Journal of Computer Vision and Image Processing* (pp. 1-17).

www.irma-international.org/article/a-novel-approach-of-restoration-of-digital-images-degraded-by-impulse-noise/115836

Accelerating Industry 4.0 Using Computer Vision Algorithms and Analysis: Analysing the Significance and Characteristics of Machine Vision in the Advancing Industry Culture

Kunal Dhibar, Prasenjit Maji, Hemanta Kumar Mondal and Swadhin Kumar Mondal (2023). *Investigations in Pattern Recognition and Computer Vision for Industry 4.0* (pp. 37-52).

www.irma-international.org/chapter/accelerating-industry-40-using-computer-vision-algorithms-and-analysis/330232

Region-Based Graph Learning towards Large Scale Image Annotation

Bao Bing-Kun and Yan Shuicheng (2013). *Graph-Based Methods in Computer Vision: Developments and Applications* (pp. 244-260).

www.irma-international.org/chapter/region-based-graph-learning-towards/69080

Classifying Behaviours in Videos with Recurrent Neural Networks

Javier Abellan-Abenza, Alberto Garcia-Garcia, Sergiu Oprea, David Ivorra-Piqueres and Jose Garcia-Rodriguez (2017). *International Journal of Computer Vision and Image Processing* (pp. 1-15).

www.irma-international.org/article/classifying-behaviours-in-videos-with-recurrent-neural-networks/195006