

Chapter 9

Generative Group Activity Analysis with Quaternion Descriptor

Guangyu Zhu

National University of Singapore, Singapore

Shuicheng Yan

National University of Singapore, Singapore

Tony X. Han

University of Missouri, USA

Changsheng Xu

Chinese Academy of Sciences, China

ABSTRACT

Activity understanding plays an essential role in video content analysis and remains a challenging open problem. Most of previous research is limited due to the use of excessively localized features without sufficiently encapsulating the interaction context or focus on simply discriminative models but totally ignoring the interaction patterns. In this chapter, a new approach is proposed to recognize human group activities. Firstly, the authors designed a new quaternion descriptor to describe the interactive insight of activities regarding the appearance, dynamic, causality, and feedback, respectively. The designed descriptor along with the conventional velocity and position are capable of delineating the individual and pairwise interactions in the activities. Secondly, considering both activity category and interaction variety, the authors propose an extended pLSA (probabilistic Latent Semantic Analysis) model with two hidden variables. This extended probabilistic graphic paradigm constructed on the quaternion descriptors facilitates the effective inference of activity categories as well as the exploration of activity interaction patterns. The extensive experiments on realistic movie and human group activity datasets validate that the multilevel features are effective for activity interaction representation and demonstrate that the graphic model is a promising paradigm for activity recognition.

DOI: 10.4018/978-1-4666-1891-6.ch009

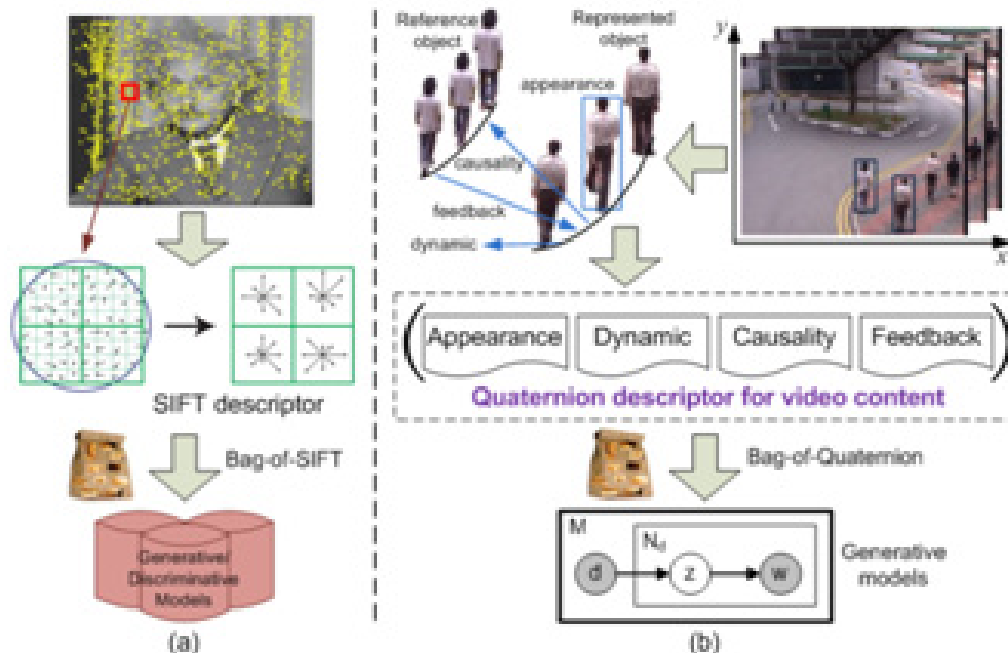
INTRODUCTION

Video-based human activity analysis is one of the most promising applications of computer vision and pattern recognition. Turaga *et al.* (2008) presented a recent survey of the major approaches pursued over the last two decades. Large amount of the existing work on this problem mainly focused on the relatively simple activities of single person (Laptev, 2003; Liu, 2009; Niebles, 2008; Schuldt, 2004; Wang, 2009), e.g., sitting, walking and hand-waving, which has achieved particular success. In recent years, recognition of group activity with multiple participators (e.g., fighting and gathering) is gaining increasing amount of interests (Marszalek, 2009; Ni, 2009; Ryoo, 2007; Zhou, 2008) from both academia and industry.

Upon the definition given by Turaga (2008), where an activity is referred to a complex sequence of actions performed by several objects who could be interacting with each other, the interactions

among the participants reflect the elementary characteristics of different activities. The effective interaction descriptor is therefore essential for developing sophisticated approaches of activity recognition. Most previous research stems from the local representation in image processing. As shown in Figure 1(a), the common sense of constructing local representation is to extract the pattern descriptors (e.g., SIFT (Lowe, 2004)) from spatial salient points and generate the feature representation (e.g., bag-of-SIFT accordingly) using bag-of-words strategy. Such successful scenario has been naturally extended to the video processing by extracting the pattern descriptors based on spatio-temporal salient points (Laptev, 2003; Liu, 2009; Marszalek, 2009; Niebles, 2008; Schuldt, 2004). Although the widely used local descriptors are demonstrated to allow for the recognition of activities in the scenes with occlusions and dynamic cluttered backgrounds, they are solely representations of appearance and motion patterns.

Figure 1. Comparison of representation and modeling for image and video analysis: (a) Image representation and modeling; (b) Video representation and modeling



14 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/generative-group-activity-analysis-quaternion/69076

Related Content

Towards the Notion of Typical Documents in Large Collections of Documents

Mieczyslaw Klopotek, Slawomir Wierzchon, Krzysztof Ciesielski, Michal Draminski and Dariusz Czernski (2011). *Visual Analytics and Interactive Technologies: Data, Text and Web Mining Applications* (pp. 1-18).

www.irma-international.org/chapter/towards-notion-typical-documents-large/48387

Analysis and Comparison of Clustering Techniques for Chronic Kidney Disease With Genetic Algorithm

Sanat Kumar Sahu and A. K. Shrivastava (2018). *International Journal of Computer Vision and Image Processing* (pp. 16-25).

www.irma-international.org/article/analysis-and-comparison-of-clustering-techniques-for-chronic-kidney-disease-with-genetic-algorithm/214071

A Deep Learning Approach for Hepatocellular Carcinoma Grading

Vitoantonio Bevilacqua, Antonio Brunetti, Gianpaolo Francesco Trotta, Leonarda Carnimeo, Francescomaria Marino, Vito Alberotanza and Arnaldo Scardapane (2017). *International Journal of Computer Vision and Image Processing* (pp. 1-18).

www.irma-international.org/article/a-deep-learning-approach-for-hepatocellular-carcinoma-grading/183657

An Efficient System for Blocking Pornography Websites

Tarek M. Mahmoud, Tarek Abd-El-Hafeez and Ahmed Omar (2014). *Computer Vision and Image Processing in Intelligent Systems and Multimedia Technologies* (pp. 161-176).

www.irma-international.org/chapter/an-efficient-system-for-blocking-pornography-websites/108414

Intelligent Control of Vehicles' Number Plates on Toll Gates in Developing Nations

Tawanda Mushiri, Charles Mbohwa and Simbarashe Sarupinda (2018). *Handbook of Research on Advanced Concepts in Real-Time Image and Video Processing* (pp. 232-280).

www.irma-international.org/chapter/intelligent-control-of-vehicles-number-plates-on-toll-gates-in-developing-nations/186280