Accepted Manuscript

On Fisher Vector Encoding of Binary Features for Video Face Recognition

Yoanna Martínez-Díaz, Noslen Hernández, Rolando J. Biscay, Leonardo Chang, Heydi Méndez-Vázquez, L. Enrique Sucar

PII: S1047-3203(18)30024-5
DOI: https://doi.org/10.1016/j.jvcir.2018.01.017
Reference: YJVCI 2129

To appear in: *J. Vis. Commun. Image R.*

Received Date: 5 July 2017
Revised Date: 4 December 2017
Accepted Date: 27 January 2018

Please cite this article as: Y. Martínez-Díaz, N. Hernández, R.J. Biscay, L. Chang, H. Méndez-Vázquez, L. Enrique Sucar, On Fisher Vector Encoding of Binary Features for Video Face Recognition, *J. Vis. Commun. Image R.* (2018), doi: https://doi.org/10.1016/j.jvcir.2018.01.017

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

ACCEPTED MANUSCRIPT

On Fisher Vector Encoding of Binary Features for Video Face Recognition

Yoanna Martínez-Díaz[a], Noslen Hernández[b,*], Rolando J. Biscay[c], Leonardo Chang[e], Heydi Méndez-Vázquez[a], L. Enrique Sucar[d]

[a]*Advanced Technologies Application Center (CENATAV), Havana, Cuba*
[b]*University of Sao Paulo (USP), Sao Paulo, Brazil*
[c]*Centro de Investigación en Matemáticas (CIMAT), Guanajuato, México*
[d]*Instituto Nacional de Astrofísica, Óptica y Electrónica (INAOE), Puebla, México*
[e]*Tecnológico de Monterrey, Campus Estado de México, Estado de México, México*

Abstract

Several approaches have been proposed for face recognition in videos. Fisher vector (FV) encoding of local Scale-Invariant Feature Transforms (SIFT) is among the best performing ones. Aiming at speed up the computation time of this approach, a method based on FV encoding of binary features was recently introduced. By using Binary Robust Independent Elementary Features (BRIEF), this method gained in efficiency but lost in accuracy. FV representation of binary features demands appropriated mathematical tools, which are not as easy available as for continuous features. This paper introduces a new way for obtaining FV encoding of binary features that is still efficient and also accurate. We show that BRIEF combined with FV are discriminative enough, and provide as good performance as the one obtained by using SIFT features for video face recognition. Besides, we discuss several insights and promising lines of future work in regard to FV encoding of binary features.

Keywords: fisher vector, binary features, face recognition, video

1. Introduction

Video face recognition has attracted the attention from both academy and industry in the last years, mainly due to its relevance for real-world applications (e.g., video-surveillance, human-computer interaction) [1, 2, 3]. For such applications it is necessary to develop algorithms that operate with limited computing resources, maintaining a high accuracy and a processing speed close to real-time.

Recently, state-of-the-art results in video face recognition have been achieved by using convolution neural networks [4, 5, 6]. However, the main limitation of these approaches is the requirement of a massive amount of labeled training data (millions of images) to achieve good generalization, demanding thus a lot of training time and

*Corresponding author: noslenh@gmail.com

Download English Version:

https://daneshyari.com/en/article/6938337

Download Persian Version:

https://daneshyari.com/article/6938337

Daneshyari.com