Pattern Recognition Letters 52 (2015) 40–47

Contents lists available at ScienceDirect

Pattern Recognition Letters

journal homepage: www.elsevier.com/locate/patrec

# Background suppressing Gabor energy filtering ☆

Albert C. Cruz*, Bir Bhanu, Ninad S. Thakoor

*Center for Research in Intelligent Systems, Winston Chung Hall 216, University of California, Riverside, Riverside, CA 92521, USA*

ARTICLE INFO

*Article history:*
Received 14 May 2014
Available online 13 October 2014

*Keywords:*
Gabor filter
Background texture
Facial emotion recognition
Bioimaging

ABSTRACT

In the field of facial emotion recognition, early research advanced with the use of Gabor filters. However, these filters lack generalization and result in undesirably large feature vector size. In recent work, more attention has been given to other local appearance features. Two desired characteristics in a facial appearance feature are generalization capability, and the compactness of representation. In this paper, we propose a novel texture feature inspired by Gabor energy filters, called background suppressing Gabor energy filtering. The feature has a generalization component that removes background texture. It has a reduced feature vector size due to maximal representation and soft orientation histograms, and it is a white box representation. We demonstrate improved performance on the non-trivial Audio/Visual Emotion Challenge 2012 grand-challenge dataset by a factor of 7.17 over the Gabor filter on the development set. We also demonstrate applicability of our approach beyond facial emotion recognition which yields improved classification rate over the Gabor filter for four bioimaging datasets by an average of 8.22%.

© 2014 Elsevier B.V. All rights reserved.

## 1. Introduction

Features extracted from images are at the core of computer vision and pattern recognition and their applicability can vary depending on the type of images. In this paper, we focus on local appearance features because they are easily adaptable to different application areas. Specifically, we are inspired by the Gabor filter which was originally introduced in 1946 [1,2]. Since then, it has seen extensive use in many fields of pattern recognition. Gabor filtering is the process of representing an image in terms of gratings that approximate the low-level behavior of the human visual system. However, current methods prefer to use features other than the Gabor energy filter. In this paper, we explore Gabor-based features' two areas where other features are more frequently used than the Gabor filter, namely facial emotion recognition and bioimaging.

State-of-the-art features have two properties: (1) the ability to generalize to external and intrinsic factors, such as registration errors, illumination variations, blur or noise. For example, local binary pattern (LBP) features can be rotation invariant and are robust to monotonic grayscale transformation from shadows [3]. The scale invariant feature transform (SIFT) is scale invariant [4]. (2) Compactness of feature representation. LBP uses histograms to reduce feature vector size. The original formulation [5] of the Gabor energy filter does not have either of these properties.

We propose background suppressing Gabor energy filtering which removes background texture with a generalization step, and reduces feature vector size with a computational efficiency step. We improve performance over other frontal face feature representations used for facial emotion recognition on the Audio/Visual Emotion Challenge (AVEC) 2012 grand-challenge dataset [6] and the Cohn–Kanade+ (CK+) dataset [7]. We also provide results on four bioimaging datasets.

## 2. Related Work, motivation and contributions

The focus of this work is local appearance features, the most commonly used of which are LBP [3]. Though the features are often referred to as LBP features, they are actually histograms of a LBP coded image. LBP quantifies textures at a pixel level by encoding the microtexture of a pixel and its neighborhood with an eight-bit code. Ref. [26] conducted a survey of LBP features for use with bio-imaging data and investigated: elongated quinary patterns (EQP), local ternary patterns, improved local binary patterns and center-symmetric local binary patterns. It was found that EQP was the best performer. Ref. [27] detected mTBI from MRI images with LBP in a context based system. In facial emotion recognition, current methods often divide the frontal face into sub-regions and compute the histogram of LBP codes for each sub-region. For example, in Ref. [6], the face was divided evenly into $10 \times 10$ sub-regions, or grids, and the outer regions were discarded because they corresponded to the regions of a face where there were no facial expressions. Uniform LBP features have been used as the baseline for recent facial emotion recognition grand challenges

☆ This paper has been recommended for acceptance by S. Sarkar.

* Corresponding author. Tel.: +1 951 827 3954.
*E-mail address:* acruz@ee.ucr.edu (A.C. Cruz).

http://dx.doi.org/10.1016/j.patrec.2014.10.001
0167-8655/© 2014 Elsevier B.V. All rights reserved.

[6]. There have been many improvements to the original LBP feature. Ref. [11] proposed three-patch and four-patch local binary patterns (TPLBP, FPLBP). Whereas LBP encodes a microtexture of a single pixel, TPLBP and FPLBP encode larger patterns and homogeneity of a region by comparing a pixel's microtexture to the microtextures of neighboring pixels. Ref. [13] proposed extending LBP to a spatiotemporal feature with the use of three orthogonal planes (LBP-TOP). Ref. [28] extended LBP to the spatiotemporal domain with monogenic signals analysis and phase-quadrant encoding and a local XOR operator in three orthogonal planes (STLMMBP).

Not all facial emotion recognition and bio-imaging methods use LBP as a local appearance feature. Ref. [29] detected myopia from retinal fundus images with a bag-of-features including SIFT. Ref. [30] clustered pigmented skin lesions from a dermoscope with LPQ and other features. Ref. [31] classified states of hESC from phase contrast images with Gabor statistics. The top approach for the facial emotion recognition and analysis challenge for discrete emotions used local phase quantization (LPQ) [16]. In LPQ, the phase of a per-pixel discrete Fourier transform (DFT) quantifies the texture. It was found that the phase of DFT of a local neighborhood is invariant to centrally symmetric blur. Sub-region histograms give LPQ a compact representation. Ref. [21] used a difference image to quantify facial motion, and a discrete cosine transform (DCT) to compress the feature vector size. Ref. [4] proposed the scale-invariant feature transform (SIFT), which quantifies local features with the maxima and minima of a difference-of-Gaussians. Recently, it was used by Ref. [23], where the SIFT features were computed at 83 fiducial feature points. A summary of related work is given in Table 1.

### 2.1. Motivation

We focus our work on improving the Gabor energy filter because it has been, and still is an important feature for computer vision [32], though it has no generalization or computational efficiency steps. There are approaches that have improved upon the original Gabor energy filter [5]. Ref. [33] used the imaginary part of the Gabor filter for cerebrovascular images. Ref. [34] applied a spatiotemporal Gabor filter to emotion recognition on the Cohn–Kanade dataset. Ref. [20] represented the output of Gabor energy filters with sub-region histograms. These two methods improve the Gabor energy filter, but do not address both generalization and compactness of the representation. Out of the top six approaches for AVEC 2011, only one approach used a Gabor energy filter [17]. Approaches preferred LPQ, LBP or active appearance models. We assert that the Gabor filter can still be effectively applied to facial emotion if the following technical challenges are addressed:

(1) *Generalization*: Gabor energy filters do not generalize well in unconstrained settings because a Gabor energy filter captures edge magnitudes at almost all orientations, including edges from noise due to background texture. Current local appearance features have additional steps in an effort to be more generalizable and robust. Ref. [34] addressed this by extending the Gabor filter to temporal domain with Gabor motion energy features. However, the feature vector size was increased by the number of temporal scales over the original Gabor energy filter, which already has a large feature vector size. For example, the feature vector was increased by a factor of 3.72 between Refs. [5] and [34]. We address this technical challenge with background suppressing Gabor energy filtering, which removes the edges due to background noise but retains the significant edges that correspond to the edges of the objects in a scene. We also compute texture at a pixel, microtexture level, so the method is invariant to monotonic grayscale transformations. We prefer Gabor filters over LBP because the background suppression pipeline emulates the human visual system (HVS). It requires tuned filters which can be approximated by the Gabor filter. LBP does not approximate the HVS in this way.

**Table 1**
Summary of related work. Size: feature vector size.

| Method | Feature | Generalization | Computational efficiency | Size | Recent usage |
|---|---|---|---|---|---|
| [3] | Local binary patterns (LBP) | Rotation invariance, robust to monotonic grayscale transformations | Uniform patterns reduce number of codes, sub-region histograms | 5900 | Baseline features for facial emotion recognition grand-challenges [6,8]; dynamic sampling approach [9]; survey with varying classifiers [10] |
| [11] | Three-patch and four-patch local binary patterns (TPLBP, FPLBP) | Robust to monotonic grayscale transformations | Sub-region histograms | Not stated | Used with prototype hyperplane learning on labeled faces in the wild dataset [12] |
| [13] | Local binary patterns from three orthogonal planes (LBP-TOP) | Robust to monotonic grayscale transformations, temporal information | Sub-region histograms | 12 | Action unit detection on the man machine interaction dataset [14] |
| [15] | Local phase quantization (LPQ) | Robust to blur | Sub-region histograms | 25 600 | Top approach in facial emotion recognition and analysis sub-challenge for discrete facial emotions [16] |
| [5] | Gabor energy filter | – | – | 595 353[a] | Entry to AVEC grand-challenge [17] |
| [18] | Local Gabor binary pattern histogram sequence (LGBPHS) | Illumination invariance | Sub-region histograms | 151 040 | Survey with other LBP features [19] |
| [20] | Gabor energy filter histograms | – | Sub-region histograms | 2400 | Entry to AVEC grand-challenge [20] |
| [21] | Discrete cosine transform (DCT) | Difference image, accounts for motion only | DCT to compress difference image | $<$10 000[a,b] | Applied to AVEC grand-challenge [22] |
| [4] | Scale invariant feature transform (SIFT) | Scale, rotation and affine invariance | Histograms | 10 624 | Used with regional covariance matrix for multi-view face representation on BU-3DFE [23] |
| Proposed method | Background suppressing Gabor energy filtering | Removes background noise, robust to monotonic grayscale transformations | Maximal response determines significant edges, sub-region histograms | 6400 | Detection of microtubules in bioimaging data [24,25] |

[a] Assuming the image was a square image of 100 pixels width.
[b] Varies based on how much energy is retained by the DCT.

Download English Version:

https://daneshyari.com/en/article/536329

Download Persian Version:

https://daneshyari.com/article/536329

Daneshyari.com