Pattern Recognition 60 (2016) 734–747

Contents lists available at ScienceDirect

Pattern Recognition

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

Data-driven techniques for smoothing histograms of local binary patterns

Juha Ylioinas [a,*], Norman Poh [b], Jukka Holappa [a], Matti Pietikäinen [a]

[a] Center for Machine Vision and Signal Analysis, Faculty of Information Technology and Electrical Engineering, University of Oulu, P.O. Box 4500, FI-90014, Finland

[b] Department of Computing, Faculty of Engineering and Physical Sciences (FEPS), University of Surrey, GU2 7XH, UK

ARTICLE INFO

Article history:
Received 11 August 2015
Received in revised form
27 June 2016
Accepted 28 June 2016
Available online 29 June 2016

Keywords:
Local binary patterns
Local binarized descriptors
Histogram
Soft-assignment
Kernel density estimation
Histogram smoothing

ABSTRACT

Local binary pattern histograms have proved very successful texture descriptors. Despite this success, the description procedure bears some drawbacks that are still lacking solutions in the literature. One of the problems arises when the number of extractable local patterns reduces while their dimension increases rendering the output histogram descriptions sparse and unstable, finally showing up as a reduced recognition rate. A smoothing method based on kernel density estimation was recently proposed as a means to tackle the aforementioned problem. A constituent part of the method is to determine how much to smooth a histogram. Previously, this was solved via trial-and-error in a problem-specific manner. In this paper, the goal is to present data-driven methods to determine this smoothing automatically. In the end, we present unsupervised and supervised methods for the given task and validate their performance with a representative set of local binary pattern variants in texture analysis problems covering material categorization and face recognition.

1. Introduction

The goal of image appearance description is to effectively describe the salient class information in images and at the same time to cope with within-class variations due to degradations caused by varying imaging conditions. As it plays a key role in almost all computer vision applications concerning detection, recognition and classification, image appearance description is considered one of the most important topics in the field of computer vision.

An important group of image description methods is formed by those that locally describe image textures by binary codes, a methodology better known as local binary patterns (LBP[1]). Usually, these descriptors (like Ojala's variant [1]) are based on first convolving an image patch with a set of filters and then quantizing the filter responses (feature maps) via thresholding. As a result, for every pixel in an input image, there is a binary code (or pattern) that somehow describes the local texture content in the given location. Once the local texture content has been described by these patterns, a higher-level description for larger image regions (or even for the whole image) is usually constructed by populating a histogram of these patterns. Such statistical modeling of local image features for image description or representation purposes belongs to a class of methods known as Bag-of-Visual-Word (BoVW) models. See Fig. 1 for an illustration of the LBP histogram description process given above.

The characteristics and the fashion of using LBP and many of its derivatives give rise to two fundamental problems. The first one arises from scalar quantization that makes the pattern coordinates sensitive to even minimal changes in the input. Such alterations might result from different forms of image noise such as salt-and-pepper noise [2,3]. The second issue, which is the one we are more interested in, is related to the role of the histogram to which the final representation commonly builds on. It is widely acknowledged that the histogram in general is often very unstable due to the very nature of the counting process in it. For LBP it means that the features that are under represented or sparse results in very few and unreliable counts in the histogram. As a result, under-represented patterns render the resultant histogram representation to be statistically unstable that may eventually lead to reduced capability to describe and handle well the discussed salient class information and within-class variations. Fig. 2 illustrates a hypothetical scenario where LBP histograms are not able to perform well due to the limited size of the input texture sample.

In this paper, we focus on the use of LBP histograms in building representations for image textures. We show particular interest in their performance when only small image patches are available to populate the histogram bins. The major aim is to understand

* Corresponding author.
E-mail addresses: juyl@ee.oulu.fi (J. Ylioinas), normanpoh@ieee.org (N. Poh), jukka.holappa@ee.oulu.fi (J. Holappa), mkp@ee.oulu.fi (M. Pietikäinen).

[1] By LBP we mean any variation reported in the literature.

http://dx.doi.org/10.1016/j.patcog.2016.06.029
0031-3203/

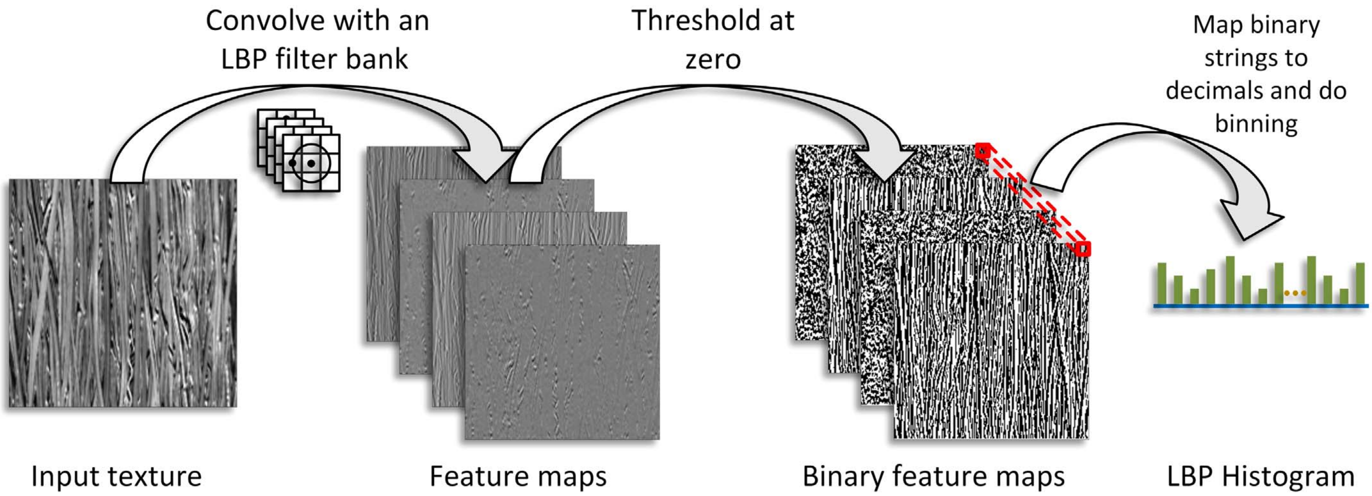

Fig. 1. A picture illustrating the conventional pipeline of image description process via LBP histograms (best viewed in color). The given example is of using $LBP_{n=4,r=1}$ yielding 16-dimensional histograms. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)

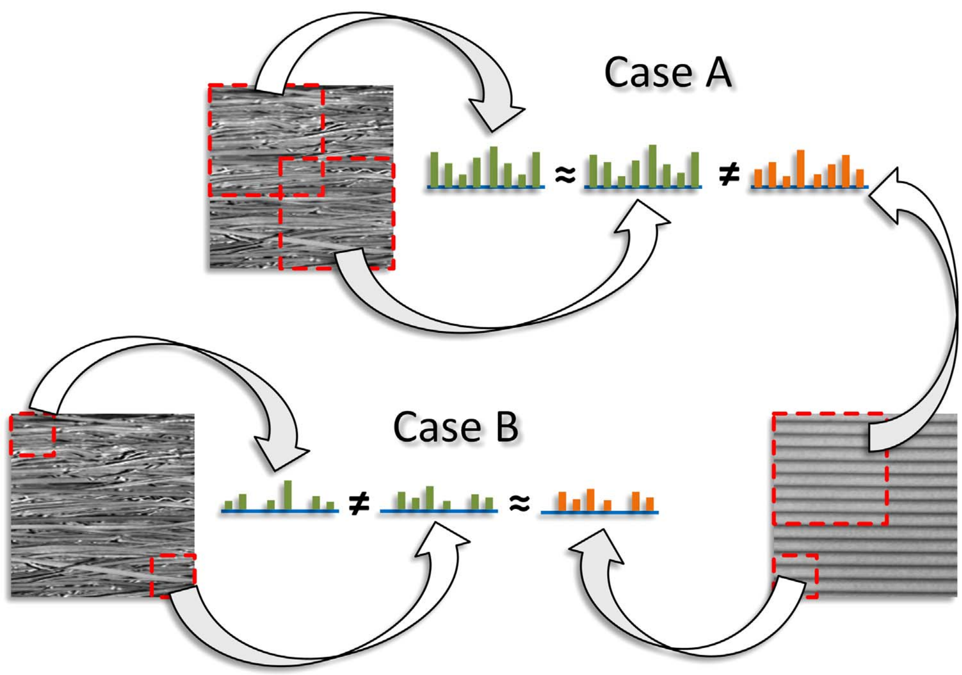

Fig. 2. An illustration of a scenario where LBP histograms are not able to describe and handle well salient class information and within-class variations. When the input sample is large enough (Case A), histograms represent well their corresponding texture classes, but when the input sample is small (Case B), histograms may easily mix up between different texture classes.

whether only a small number of local pattern instances can serve as a good basis for the process of describing images for subsequent recognition. We start by reviewing a method that can to some extent mitigate the above-mentioned issue of instability inherent to LBP histograms. The method is based on kernel density estimation in n-dimensional binary space. In order to be consistent with the previous studies [4–6], we also refer to it as histogram smoothing. Even though the method has already been presented in the literature, our work presented here clearly extends and differs from any of those studies conducted, especially from [4–7]. The main contributions can be summarized as follows: After illustrating how histogram smoothing is effortlessly put into action via a simple matrix-vector product, we commit ourselves to formulate data-driven methods for tuning the smoothing parameter of the kernel function. We also discuss several other aspects that the matrix-vector product formulation facilitates. For example, we show how the proposed smoothing can be effortlessly applied in modeling joint distributions of LBP and how it can be combined with standard dimensionality reduction. In the end, we show how the method can be made computationally less expensive by enforcing the smoothing matrix to contain many zero-entries. In summary, our study offers new insights and tools to the methodology of LBP texture description.

The structure of this study is as follows. We first provide a review of the most relevant studies related to the topic at hand. We start with BoVW (Bag-of-Visual-Words) models and different soft-assignment schemes applied with them. After that, we proceed to a short review of LBP and its variants that we consider closely related to our proposed method. In detail, the variants cover three largely applied methods for producing binary patterns, but also different techniques related to post-processing LBP histograms. Couple of alternative methods to produce smoother LBP