Computer Vision and Image Understanding 139 (2015) 161–177

Contents lists available at ScienceDirect

Computer Vision and Image Understanding

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

Local binary patterns on triangular meshes: Concept and applications

Naoufel Werghi[a], Claudio Tortorici[a], Stefano Berretti[b,*], Alberto del Bimbo[b]

[a] Khalifa University of Science Technology and Research, Abu Dhabi, United Arab Emirates
[b] Information Engineering, University of Florence, 50139, Florence, Italy

ARTICLE INFO

Article history:
Received 17 June 2014
Accepted 29 March 2015

Keywords:
3D local binary patterns
Mesh manifold
Ordered ring facets
3D texture classification
3D texture retrieval
3D face matching

ABSTRACT

In this paper, we introduce an original framework for computing local binary like-patterns on 2D mesh manifolds (i.e., surfaces in the 3D space). This framework, dubbed mesh-LBP, preservers the simplicity and the adaptability of the 2D LBP and has the capacity of handling both open and close mesh surfaces without requiring normalization as compared to its 2D counterpart. We describe the foundations and the construction of mesh-LBP and showcase the different LBP patterns that can be generated on the mesh. In the experimentation, we provide evidence of the uniform patterns in the mesh-LBP, the repeatability of its descriptors, and its robustness to moderate shape deformations. Then, we show how the mesh-LBP descriptors can be adapted to a number of surface local and global analysis including 3D texture classification and retrieval, and 3D face matching. We also compare the performance of the mesh-LBP descriptors with a bunch of state of the art surface descriptors.

1. Introduction

The rapid advancement of the 3D imaging technologies resulted in a new generation of acquisition devices capable of capturing the 3D objects geometry in the physical three-dimensional space. High-resolution 3D static scanners as well as devices with 3D dynamic acquisition capabilities that provide a continuous flow of the 3D geometry of a scene are now available. In addition to this, the geometric and RGB information are often captured in a synchronized way. In this respect, the interest and the widespread use of Kinect-like cameras in consumer and research applications is one of the most evident advancement in the field.

The geometric information captured by such 3D acquisition devices is typically in the form of a cloud of points, which represent the three-dimensional coordinates of a set of samples of the 3D object surface. However, the direct processing of these point clouds is not convenient or even possible, so that other representation formats have been established. Depth images are one of the most commonly used imaging modality, since they permitted a direct extension to the depth dimension of many computer vision and pattern recognition solutions developed for analyzing the photometric information in 2D images. Though the possibility of a straightforward extension of 2D techniques is attractive, this modality loses the full 3D geometry, by reducing it to a 2.5D projection. The full 3D shape information is instead preserved and encoded in a simple, compact and flexible format by the triangular mesh manifold modality. This is widely used in many fields, such as animation, medical imaging, computer-aided design and many others. The recent advances in shape scanning and modeling have also allowed the integration of both photometric and geometric information into a single support defined over a 2D mesh-manifold. However, despite the abundance and the richness of the mesh manifold modality, the number of solutions for shape representation that exploit the full 3D geometry of the objects is still limited, and not comparable with the large variety of methods available in 2D. Indeed, many effective solutions capable of capturing discriminative information have been developed for 2D still images and videos, but their 3D counterpart is often not available.

An example of such successful 2D descriptors is represented by the local binary pattern (LBP). Since its first formal definition by Ojala et al. [1,2], the LBP has established itself as one of the most effective local shape descriptors for image representation. It has been originally introduced for representing 2D textures in still images, but its computational simplicity and discriminative power attracted the attention of the image processing and pattern recognition community for other different tasks. Rapidly, LBP has found applications in visual inspection [3,4], remote sensing [5–7], face recognition [8–11], facial expression recognition [12], and motion analysis [13,14]. However, the LBP-based methods developed so far operate either on photometric information provided by 2D color images or on geometric information in 2D depth images. The few solutions that extract surface features directly in 3D (typically in the form of surface normals), resort to the 2D case by converting the 3D extracted

* Corresponding author.
E-mail address: stefano.berretti@unifi.it, berretti@dsi.unifi.it (S. Berretti).

http://dx.doi.org/10.1016/j.cviu.2015.03.016

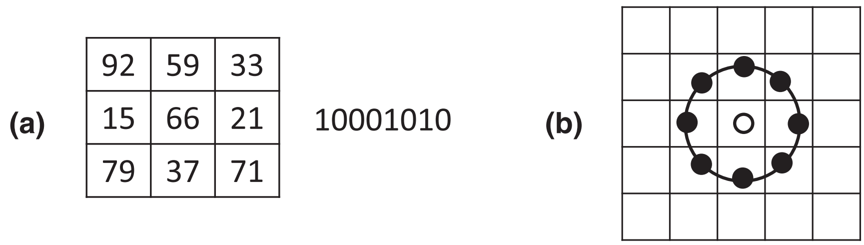

Fig. 1. (a) Computation of the basic LBP code from the 3 × 3 neighborhood of a central pixel. Each pixel, starting from the upper-left corner is compared with the central pixel to produce 1 if its value is greater or equal, 0 otherwise. The result is an 8-bit binary code. (b) Example of a central pixel with a circular neighborhood of given radius.

features to depth values, and then use ordinary LBP processing on 2D images [15–17]. However, to the best of our knowledge, there is no a framework that allows the computation of LBP on a mesh manifold. Since LBP requires an ordered support for its computation, the major factor that contributed in this lack is the absence of an intrinsic order in the triangular mesh manifold. On the contrary, computation of LBP on 2D images, either photometric or depth, benefits from the implicit ordering of the pixels in the 2D image array.

Motivated by these facts, in this paper we address the challenge of computing LBP on a mesh manifold by proposing an original framework that we call mesh-LBP, which allows the extraction of LBP-like patterns directly from a triangular mesh manifold, without the need of any intermediate representation in the form of depth images. With this framework, we can therefore build on the current 2D-LBP analysis methods, extending them to mesh manifolds as well as to the modality that also embeds photometric information into mesh models. To motivate our solution and to relate it to the state of the art approaches, next we provide an overview of the LBP literature.

1.1. Related work

According to its original definition [1], the LBP operator assigns labels to image pixels performing the following steps: given a pixel (central pixel), first its 3 × 3 neighborhood is considered, and the gray value of each pixel in the neighborhood is thresholded with the value of the central pixel (i.e., each pixel in the neighborhood is regarded as 1 if its value is greater or equal to the central value, 0 otherwise); then, the sequence of 0/1 in the neighborhood of the central pixel is regarded as a binary number according to a positional coding convention and considered as the LBP value of the central pixel. This is shown in Fig. 1(a), where the upper left pixel in the neighborhood is regarded as the most significant bit in the final code. This eight bits number encodes the mutual relationship between the gray levels of the central pixel and of its neighboring pixels. The histogram of the numbers obtained in such a way can then be used as a texture descriptor of a region or of an entire image. This operator is distinguished by its simplicity, efficient computation, and invariance to monotonic gray-level transformations. Later, the same authors of the basic LBP proposed an extended version that can operate on circular neighborhood of different radii [2], also allowing sub-pixel alterations (see Fig. 1(b)).

These initial formulations led subsequently to the definition of other neighborhood variants, like the oriented elliptic neighborhood LBP (elongated LBP) proposed by Liao and Chung [18], which accounts for anisotropic information, and the multi-block LBP (MB-LBP) that compares the averages of the gray level intensity of neighboring pixels rather than the value of individual pixels, in order to capture macrostructural features in the image [19]. Other versions have been proposed to improve the discriminative power of the descriptor, such as the improved LBP (ILBP) [20], in which pixel values are compared with the average of the neighborhood, and the extended LBP (ELBP) [21] which encodes, in addition to the binary comparison between pixels values, the amplitude of their difference using additive binary digits. To improve the robustness of LBP, Tan and Triggs [22] introduced the so-called local ternary pattern (LTP), which substitutes the original binary code by a three-values code (1, 0 and -1) by means of a user-defined threshold. This new operator addressed the sensitivity to noise, though at the cost of losing the invariance to monotonic gray-level transformations. A fuzzy-logic version of the LTP was proposed later in [23], where a fuzzy membership function substituted the crisp three-states association used in [22]. A more complete list and discussion on the many LBP variants appeared in the literature can be found in [24].

Considering the case of 3D shape analysis, most if not all the LBP-based approaches have been developed for face recognition applications. Many of the techniques developed in this context operate on standard depth images, where the z-coordinate is mapped to a gray-level value. This format allowed a straightforward application of the 2D-LBP operator as it was demonstrated in the pioneering work of Li et al. [25]. Later, Huang et al. [26] proposed a 3D-LBP operator that also encodes depth differences of neighboring pixels. More recently, Huang et al. [27] extended the 3D-LBP to a multiscale extended LBP (eLBP), which consists of several LBP codes in multiple layers accounting for the exact gray value differences between the central pixel and its neighbors. Sandbach et al. [15] proposed a local normal binary pattern (LNBP), which used the angle between normals at two points, rather than the depth value to obtain the local binary code. Similar to this, in [16] the surface normals are extracted in 3D, then the values of the normal components along the direction of the three coordinate axes are interpreted as depth values, and LBP is computed on these depth maps reporting the values of the normal components. The idea of exploiting surface normals is further extended in [17], where azimuthal projection distance images are constructed. The azimuthal equidistant projection is able to project normals onto points in an Euclidean space according to the direction. Though the projected information is not the depth, depending on the normals of the 3D surface, 2D LBP are still computed on the projection images. Fehr and Burkhardt [28] attempted a 3D-LBP definition specifically tailored for volumetric data by sampling a sphere of a given radius around a central voxel. The approach is computationally expensive in that the rotation-invariance had to be addressed with complex techniques involving spherical correlation in the frequency domain.

1.2. Paper contribution and organization

The analysis above evidences that the LBP descriptor has attracted great interest for the analysis of 2D images, mainly for its simple and efficient computation and for the effective results that can be achieved relying on the LBP theory. Recently, various attempts have been made for extending the LBP framework to the case of 3D meshes, but none of them succeeded in addressing all the issues posed by the need for a simple and effective processing directly performed on a mesh-manifold. Indeed, existing solutions address the LBP extraction on 3D meshes by resorting to the easier 2D case, through the projection of 3D meshes on 2D depth images or by computing normals or