Optics and Lasers in Engineering 49 (2011) 1177–1184

Contents lists available at ScienceDirect

Optics and Lasers in Engineering

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

Image classification with binary gradient contours

Antonio Fernández [a,]*, Marcos X. Álvarez [a], Francesco Bianconi [b]

[a] Department of Engineering Design, Universidade de Vigo, Escola de Enxeñaría Industrial, Campus Universitario, 36310 Vigo, Spain
[b] Department of Industrial Engineering, Università degli Study di Perugia, Via G. Duranti 67, 06125 Perugia, Italy

ARTICLE INFO

Article history:
Received 22 March 2011
Accepted 3 May 2011
Available online 13 May 2011

Keywords:
Texture features
BGC
LBP

ABSTRACT

In this work we present a new family of computationally simple texture descriptors, referred to as binary gradient contours (BGC). The BGC methodology relies on computing a set of eight binary gradients between pairs of pixels all along a closed path around the central pixel of a 3×3 grayscale image patch. We developed three different versions of BGC features, namely single-loop, double-loop and triple-loop. To quantitatively assess the effectiveness of the proposed approach we performed an ensemble of texture classification experiments over 10 different datasets. The obtained results make it apparent that the single-loop version is the best performer of the BGC family. Experiments also show that the single-loop BGC texture operator outperforms the well-known LBP. Statistical significance of the achieved accuracy improvement has been demonstrated through the Wilkoxon signed rank test.

1. Introduction

Texture analysis plays an important role in computer vision and pattern recognition. Tumor detection in medical imaging for computer-aided diagnosis, automated surface inspection for industrial quality control and terrain classification through the analysis of remote sensed imagery are just some of the applications in which textural information can be successfully exploited. Texture analysis techniques have been recently extended to study dynamic events such as recognition of facial expression [1] or monitoring of paint drying process [2]. Texture descriptors are traditionally classified into four categories: statistical, model-based, geometrical and signal processing methods [3,4]. Among these categories, statistical methods have become very popular, mainly because they provide good accuracy at an affordable computational cost. The rationale behind statistical texture description is that texture can be represented through the joint distribution of pixel intensities in a local neighborhood. Based on this assumption, a stationary texture image (i.e., an image that contains a single type of texture) could be ideally characterized by means of the probability distribution of the possible grayscale patterns. This probability can be estimated by a histogram that measures the occurrence frequency of the different grayscale patterns throughout the image. To compute such a histogram, the image is scanned by one-pixel steps with a sliding window, and at each window position the bin corresponding to the detected pattern is incremented by one unit. Although this approach results attractive for its conceptual simplicity, a straightforward application of the method is impractical, since the number of entries in the histogram is overwhelmingly large even for small neighborhoods. To cope with multidimensional histograms it is useful to partition the feature space into a discrete vocabulary of local features [5]. It has been recently proposed to reduce the dimension of the histogram through unsupervised clustering of grayscale patterns into a dictionary of *textons* [6]. Reported results show that this method achieves high success rates in texture classification experiments. However, clustering has a number of drawbacks: dependency of the texton dictionary upon the texture samples used to train the classifier, influence of parameter tuning on classification accuracy, and large computational overhead (especially when large neighborhoods are considered). An alternate approach to partition the feature space is through a closed-form mapping [7]. Mapping-based histogram reduction does not have the drawbacks of clustering, since these schemes define a universal vocabulary of textural features, are parameter-free and compute fast. Several mappings have been proposed by diverse research groups [8–14]. Despite all of these implementations share the same underlying principle, to the best of our knowledge they have not been yet integrated into a general framework. In this paper we present such a unifying framework. Our claim is that these apparently diverging dimensionality reduction schemes can be interpreted as a mapping from the set of grayscale patterns to a set of integer indexes. This mapping induces a partition of the set of grayscale patterns into groups of *equivalent patterns*. Dimensionality reduction is achieved by merging the occurrence frequencies of equivalent

* Corresponding author.
E-mail addresses: antfdez@uvigo.es (A. Fernández), marcos@uvigo.es (M.X. Álvarez), bianco@unipg.it (F. Bianconi).
URLS: http://webs.uvigo.es/antfdez/ (A. Fernández), http://dismac.dii.unipg.it/bianco/ (F. Bianconi).

0143-8166/$ - see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.optlaseng.2011.05.003

patterns into a single histogram bin. We used this mapping-based framework to describe a novel family of texture descriptors, called binary gradient contours (BGC), which consider the binary gradient of the grayscale values along the eight peripheral pixels of a 3×3 window. In this class of models a texture is described through the occurrence frequency of the resulting binary 8-tuples. The effectiveness of BGC features has been experimentally demonstrated through an ensemble of texture classification experiments. We have found that one out of the three proposed BGC models is more efficient in discriminating texture than the well-known LBP model.

The remaining of the paper is organized as follows. In Section 2, we present a general framework for texture description based on pattern mapping. Section 3 is devoted to describe the novel family of texture descriptors. The purpose of Section 4 is threefold. The first is to compare the proposed features with the closely related local binary pattern (LBP) concept. The second is to introduce some theoretical considerations to justify the efficiency of our approach. The third is to comparatively analyze the characteristics of the texture descriptors considered in this work. Experimental results are shown in Sections 5, and Section 6 summarizes the main conclusions that can be drawn from our work.

2. Framework for texture description based on pattern mapping

To describe the proposed framework, we shall begin by defining the notation to be used henceforth. Let $\mathbf{I}$ be a matrix of M rows and N columns representing the raw pixel intensities of an image quantized to G gray-levels, and $I_{m,n}\in\{0,1,\ldots,G-1\}$ the pixel intensity corresponding to the m-th row and n-th column. We denote by $\mathbf{S}_{m,n}$ a square crop of 3×3 pixels centered at pixel (m,n) of image $\mathbf{I}$:

$$\mathbf{S}_{m,n}=\begin{bmatrix} I_{m-1,n-1} & I_{m-1,n} & I_{m-1,n+1} \\ I_{m,n-1} & I_{m,n} & I_{m,n+1} \\ I_{m+1,n-1} & I_{m+1,n} & I_{m+1,n+1}\end{bmatrix} \tag{1}$$

Without loss of generality we can rename the terms of the equation above in order to remove the dependance on (m,n). Thus, let $\mathbf{S}$ be a matrix representing the pixel intensities of a generic square neighborhood with support 3×3. Let I_c be the gray-level of the central pixel and I_j the gray-levels of the peripheral pixels ($j\in\{0,1,\ldots,7\}$), which are arranged as follows (see Fig. 1(a)):

$$\mathbf{S}=\begin{bmatrix} I_7 & I_6 & I_5 \\ I_0 & I_c & I_4 \\ I_1 & I_2 & I_3\end{bmatrix} \tag{2}$$

Let us denote by $\mathcal{M}_{3\times3,G}$ the set of all the possible instances defined by Eq. (2). A typical value for G is 2^8 (i.e., pixel intensity is quantized in 256 levels) since the depth of digitization of most commercial imaging devices is 8 bits. It readily follows that in this case the number of different 3×3 grayscale patterns is given by

$$\#\mathcal{M}_{3\times3,256}=2^{72} \tag{3}$$

where # stands for "cardinality of". It emerges from Eq. (3) that the texture description through the joint distribution of pixel intensity over a 3×3 neighborhood involves a huge feature vector of roughly 4.7×10^{21} components. Suppose that one intends to describe a texture image through the occurrence frequency of 3×3 grayscale patterns. Provided that the number of possible patterns is several orders of magnitude greater than the number of image pixels, even for high resolution imagery, the vast majority of histogram bins would remain empty. It is well-known that such extremely sparse, ultra high dimensional histograms provide an unreliable estimation of the underlying distribution and have negligible discriminant power in image description [15]. Moreover, the memory required to store one of such histograms would largely exceed the capacity of the currently available computers. The simplest way to reduce the joint histogram dimensionality would be by decreasing G. However, as the neighborhood size increases, the number of bins grows exponentially and soon far outweighs the number of datapoints available in a single image with which to populate the histogram. To tackle such ultra high dimensional feature space we propose to partition $\mathcal{M}_{3\times3,G}$ into groups of patterns. Dimensionality reduction is straightforwardly attained by merging the histogram bins corresponding to patterns belonging to the same group into a single bin. The partition can be adequately formalized through a mapping that assigns each pattern a non-negative integer index that uniquely identifies the group the pattern belongs to

$$f:\mathcal{M}_{3\times3,G}\longrightarrow\mathbb{N}$$
$$\mathbf{S}\mapsto k=f(\mathbf{S}) \tag{4}$$

The function above establishes an equivalence relation in $\mathcal{M}_{3\times3,G}$, denoted by $\sim$:

$$\mathbf{S}_1\sim\mathbf{S}_2\Leftrightarrow f(\mathbf{S}_1)=f(\mathbf{S}_2)\quad\forall\mathbf{S}_1,\mathbf{S}_2\in\mathcal{M}_{3\times3,G} \tag{5}$$

Let $\mathcal{Q}$ be the range of f [16]

$$\mathcal{Q}=f(\mathcal{M}_{3\times3,G}) \tag{6}$$

and q the number of different groups of patterns, i.e., the number of equivalence classes:

$$q=\#\mathcal{Q} \tag{7}$$

The partition can be therefore expressed as

$$\mathcal{M}_{3\times3,G}=\bigcup_{k\in\mathcal{Q}}\mathcal{M}_{f,k} \tag{8}$$

where the family of subsets $\{\mathcal{M}_{f,k}|k\in\mathcal{Q}\}$ is pairwise disjoint, and each subset is defined by

$$\mathcal{M}_{f,k}=\{\mathbf{S}\in\mathcal{M}_{3\times3,G}|f(\mathbf{S})=k\} \tag{9}$$

In the proposed framework, the mapping f makes it possible to represent a texture image $\mathbf{I}$ by a q-dimensional vector $\mathbf{h}_f(\mathbf{I})$ in which the k-th component is given by

$$h_{f,k}(\mathbf{I})=\frac{\#\{(m,n)|f(\mathbf{S}_{m,n})=k\}}{(M-2)\times(N-2)} \tag{10}$$

It is useful to note that in order for sub-image $\mathbf{S}_{m,n}$ to be fully enclosed into $\mathbf{I}$, the crop center cannot be located at the one-pixel width periphery of the image, and therefore Eq. (10) must satisfy that $2\le m\le M-1$ and $2\le n\le N-1$.

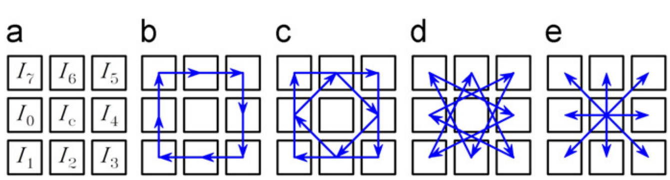

Fig. 1. (a) Spatial arrangement of a 3×3 grayscale pattern and schematic representation of the texture models considered in this paper, (b) single-loop, (c) double-loop, and (d) triple-loop versions of the binary gradient contour concept, and (e) layout of the well-known local binary pattern.

3. Binary gradient contours

We define the binary gradient contour of a 3×3 grayscale image patch as the binary 8-tuple that results of a two-step