Information Sciences 361–362 (2016) 1–13

Contents lists available at ScienceDirect

Information Sciences

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

An investigation on the use of local multi-resolution patterns for image classification

Elena González [a,*], Francesco Bianconi [b], Antonio Fernández [a]

[a] School of Industrial Engineering, Universidade de Vigo, Campus Universitario, Vigo 36310, Spain
[b] Department of Engineering, Università degli Studi di Perugia, Via G. Duranti 93, Perugia 06125, Italy

ARTICLE INFO

Article history:
Received 20 July 2015
Revised 11 March 2016
Accepted 28 April 2016
Available online 4 May 2016

Keywords:
Rotation-invariant multi-resolution patterns
Local binary patterns
Texture spectrum
Image classification

ABSTRACT

The subject of this study is the use of local multi-dimensional patterns for image classification. The contribution is both theoretical and experimental: on the one hand the paper introduces a complete and general mathematical model for encoding multi-resolution, rotation-invariant local patterns; on the other experimentally evaluates the use of multi-resolution patterns for image classification both from an information- and performance-based standpoint. The results indicate that the joint multi-resolution model proposed in the paper can actually convey an additional amount of information with respect to the marginal model; but also that the marginal model (i.e. concatenation of features computed at different resolutions) can be a good enough approximation for practical applications.

1. Introduction

Image analysis through local patterns plays a pivotal role in a number of machine vision applications including product inspection, computer-assisted diagnosis, remote sensing, surface grading, face recognition, and crowd monitoring – to cite some. Among the various approaches, the orderless bag-of-features model (BOF) has emerged as a very popular and successful strategy. This approach can be considered the transposition into the context of image analysis of the bag-of-words model (BOW), a procedure to natural language processing which represents a text through the probability of occurrence of a set of words irrespectively of the order in which they appear [5,32]. Possible extensions of the bag-of-features model involve considering pairs of patterns instead of single patterns [21,37,44]. Making again a parallel with text analysis, this solution can be regarded as using multigrams instead of unigrams [4,45,49]. In image processing a common implementation of this scheme is the multi-resolution approach, which consists of considering groups of local patterns having the same center but different radii [38]. Estimating the resulting joint distribution, however, requires a high number of features, which implies significant computational cost. A common work-around to this problem consist of using the marginal distribution instead of the joint distribution, as for instance suggested in Ref. [38]. However, the extent to which this assumption may affect the discrimination of the method has not been investigated deeply.

The objective of this paper is to understand the behaviour of local multi-dimensional patterns for image classification. Specifically, we are concerned with the evaluation of the effectiveness of different approaches for rotation-invariant multi-resolution analysis, and in particular between the joint and marginal distribution of local features. Within this topic the contribution of the present work to the literature is threefold. First, we present a rigorous and general theoretical frame-

* Corresponding author. Tel.: +34 986818602; fax: +34 98812201.
E-mail addresses: elena@uvigo.es (E. González), bianco@ieee.org (F. Bianconi), antfdez@uvigo.es (A. Fernández).

http://dx.doi.org/10.1016/j.ins.2016.04.044

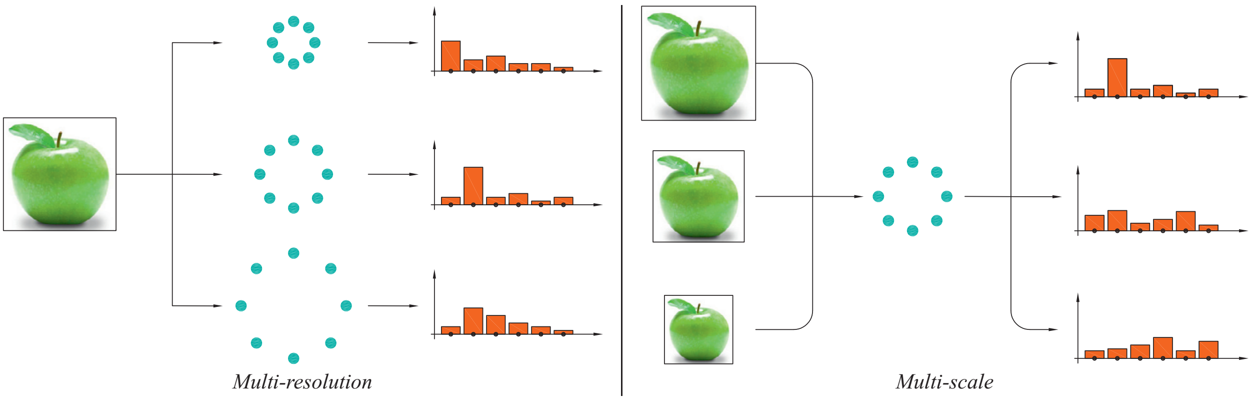

Fig. 1. Multi-resolution and multi-scale image analysis.

work for representing rotation-invariant local multi-resolution patterns. Second, we tested the model on a set of image classification experiments over seven databases of texture images using Local Binary Patterns and Texture Spectrum as local feature models. We found that, on average, the joint distribution performed better than the concatenation of the marginal distributions. Third, of this outcome we provide a theoretical explanation in terms of re-constructibility of the joint distribution from the marginal ones showing that, for the datasets considered, slight additional amount of information is carried out by the true joint distribution versus the joint distribution reconstructed from the marginal ones.

In the remainder of the paper, after briefly surveying the recent literature (Section 2), we introduce a general framework for representing rotation-invariant multi-resolution local patterns (Section 3). Then we describe the experimental set-up in Section 4, discuss the results in Section 5 and conclude the paper with some final considerations (Section 6).

2. Related research

The use of multi-resolution local patterns for image classification has been addressed by various authors in recent years. To avoid confusion we wish to make a preliminary distinction between the two concepts of *multi-resolution* and *multi-scale*. Adelson et al. [1] correctly observed that when it comes to detecting a pattern over many scales one can either (a) construct copies of the target pattern at different scales and use them to scan the original image; or (b) construct copies of the image at reduced scales and scan them with the target pattern. In this paper we refer to the former as the multi-resolution approach and to the second as the multi-scale approach. In practice, with the term multi-resolution herein we mean the analysis of a given image through local, concentric neighbourhoods of different size (Fig. 1 – left). With the term multi-scale we denote a two-step process which works as follows: (a) a set of images is obtained from the original image through suitable transforms (e.g. filtering and subsampling) which somehow alter the ratio (scale) between the real dimension of the object captured and its dimension in the image; (b) the resulting images are processed using one neighbourhood of given size (Fig. 1 – right).

It is worth noting that the literature lacks consensus about the correct way to refer to these two concepts: Jia et al. [26] for instance refer to both of them as multi-scale, Nikisins [36] use the term multi-scale to indicate what is actually a multi-resolution method – according to the definition given above. He and Sang [25] and Doshi and Schaefer [18] respectively use the terms *multi-ring* and *multi-dimensional* to indicate multi-resolution analysis. Moreover, the two concepts are not mutually exclusive: it is indeed possible to carry out a multi-resolution analysis on a multi-scale image representation, as proposed by various authors [23,48]. However, for the sake of simplicity, throughout the paper we will assume that multi-resolution analysis be carried out at the original resolution at which the source images are available.

The common strategy to deal with multi-resolution analysis consists of considering the feature distributions computed at different resolutions as independent phenomena, as was originally proposed by Ojala et al. [38]. Any interaction between the different resolutions is discarded in this model, and the feature vector is just the concatenation of the feature vectors computed at each resolution. The approach, originally developed for Local Binary Patterns, was later on extended to K-ary patterns (with $K \geq 2$) by Zhu and Wang [50], and is now common practice [17,22,31,35].

An interesting attempt to take into account the interaction of features at different resolutions has been recently made by Doshi and Schaefer [18]. Their multi-dimensional Local Binary Patterns (MD-LBP) consider the product of the probabilities (contingency table) of rotation-invariant local binary patterns (LBP^{ri}) at different resolutions. This is however an incomplete model, since the rotation-invariant distribution of multi-resolution features is not simply the product of the rotation-invariant distributions of the single-resolution features as it will be shown later in the paper.

Other authors proposed a number of methods to combine features computed at different resolutions while trying to limit the overall length of the resulting feature vector. Among them Qi et al. [41] described a two-resolution strategy in which the features at the inner neighbourhood are computed through rotation-invariant, uniform Local Binary Patterns (LBP^{riu}),