Pattern Recognition 47 (2014) 1375–1382

Contents lists available at ScienceDirect

Pattern Recognition

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

Constructing and applying higher order textons: Estimating breast cancer risk

Xi-Zhao Li*, Simon Williams [1], Murk J. Bottema [2]

School of Computer Science, Engineering and Mathematics, Flinders University, Bedford Park 5042, SA, Australia

ARTICLE INFO

Article history:
Received 17 June 2013
Received in revised form
18 September 2013
Accepted 30 September 2013
Available online 10 October 2013

Keywords:
Texture analysis
Textons
Higher order textons
Breast cancer
Risk assessment

ABSTRACT

Texture analysis based on textons is extended by introducing a method for computing textons of arbitrary order. First-, second- and third-order textons are applied to classify screening mammograms as to indicate a low or high risk of breast cancer. First-order textons are found to provide better estimates of breast cancer risk than other orders on their own but the combination of first- and second-order textons outperforms first-order textons alone and other combinations of two orders. Combining all three orders of textons does not improve classification. This example indicates that including higher-order textons has the potential to improve classification performance.

1. Introduction

The word texton was coined by Julesz to refer to local features that allow human perception to distinguish between iso-second-order textures [13]. The word texton was later re-invented to refer to co-occurrences of filter outputs [5,17,27]. A common realization of this idea is to create a feature vector comprising the outputs at that pixel of a filter bank for each pixel in an image and to search for clusters in the resulting feature space. The clusters are called textons. Each pixel in the image may then be mapped to the texton closest to the representation of the pixel in the feature space. Thus the image is replaced by a texton map. The histogram of the texton classes over the full image can be used to represent or classify the image. In order to incorporate the spatial distribution of the texton map, a natural extension is to study the spatial co-occurrence of textons over the image. Schmid [27] computed "generic descriptors" (textons) based on a "Gabor-like" filter bank and considered spatial frequency clusters. This second-order texton analysis (though not referred to as such) was found to improve image retrieval. However, Varma and Zisserman [30] found that orientation co-occurrence statistics did not improve texture classification.

The motivation and theoretical basis for standard textons (first-order textons) appears in Leung and Malik [17]. These foundations are built on models for human perception of texture in images. Combinations of groves, spots, ridges, and hollows are thought to be perceived as a finite number of textures up to equivalence under changes of scale, orientation and lighting. This motivated the idea of representing pixels by vectors of texture primitives and then clustering these vectors to determine a finite number of representative patterns – textons. The fact that local information ($N \times N$ neighborhoods with small N) is able to distinguish texture patterns on a larger scale is demonstrated in Varma and Zisserman [29]. By taking the image as a discretization of a differentiable surface, the first and second partial derivatives at a point suffice to classify all quadratic surfaces, for example. Since three points allow estimates of both first and second partial derivatives, a 3×3 neighborhood encompasses all the information needed to assign the best quadratic approximation of the image at the central point, regardless of the scale at which the quadratic surface varies with respect to the pixel size. By Taylor's theorem, higher order polynomial approximations require higher-order derivatives, which in the discrete setting of image analysis translates to larger neighborhoods to allow numerical estimates of the required derivatives. Theoretically, there is no limit as to how well the surface may be approximated by considering ever higher-order derivatives and hence ever larger neighborhoods. However, such computations are not practical. First, numerical estimates of high-order derivatives are notoriously unstable. Second, a neighborhood of diameter N results in a feature space of the order of N^2 and so becomes cumbersome for large N. Third, the assumption that the image intensity surface is well approximated by a highly differentiable function is often not valid in images where texture

* Corresponding author. Tel.: +61 882017927.fax No: +61 8 82012904
E-mail addresses: li0691@flinders.edu.au (X.-Z. Li), s.williams@flinders.edu.au (S. Williams), murk.bottema@flinders.edu.au (M.J. Bottema).
[1] Tel: +61 882012037.
[2] Tel: +61 882013652.

0031-3203/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.patcog.2013.09.033

is important. (As an aside, projection images such as the mammograms considered in this paper are inherently represented by discontinuous intensity surfaces.) For these reasons, a more practical approach to capture texture beyond simple low order approximations is to consider patterns of these local low order approximations. Second-order textons capture patterns of first-order textons. As an example, the intersection of two ridges would require a fourth-order polynomial approximation and thus requires a single neighborhood of size 5×5. However, the pattern of quadratic approximations on the nine 3×3 neighborhoods within the 5×5 patch also determines this structure as a combination of local quadratic approximations. The labels associated with these local quadratic structures form the second-order feature vector on which the second-order texton is based. The advantages are that only low-order derivatives (which are numerically more stable) are used, low dimensional feature spaces are considered (9 dimensional instead of 25 dimensional) and the model assumes only a twice differentiable function instead of a four-times differentiable function.

Although the theoretical basis for the method may be explained in terms of differentiable models of the intensity surface, derivatives are not computed explicitly and polynomial models are not constructed. The implementation relies solely on the patterns of local intensity values.

Here a general notion of higher-order textons is introduced. Second-order textons are textons defined on texton maps and third-order textons are textons defined on second-order texton maps and so on. In general, applying filters to texton maps is meaningless since the values comprising the texton map are labels and so carry no rank information. However, higher-order textons do make sense if the process of extracting the features used to construct textons does not involve arithmetic. The $N \times N$ neighborhood intensity features considered by Varma and Zisserman [29] do not require arithmetic, for example. In this method, each pixel in an image is represented by the vector of image intensity values in the $N \times N$ neighborhood of the pixel. The resulting feature vector may be viewed as the output of applying N^2 filters, each comprising an $N \times N$ patch of zeros with a single entry of 1. Thus this $N \times N$ neighborhood method is an example of the general filter bank approach to textons described above. But since this filtering step involves no arithmetic, this version of texton analysis may be applied to the texton map and, iteratively, to higher-order texton maps. Other examples of texture features that do not require arithmetic include features based on gray scale dependence matrices also known as coocurrence matrices [8] and run length statistics [6]. The restriction against the use of arithmetic only applies to second- and higher-order textons. Any method for constructing textons may be used to arrive at the first-order texton map.

Our primary interest in texture analysis lies in computer-aided interpretation of digital mammograms, and in particular, determining the risk of breast cancer based on screening mammograms. Accordingly, the use of second- and third-order textons is tested in the context of automatically classifying screening mammograms as to indicate a high or low risk of breast cancer. This classification is important as strategies for early detection of breast cancer depend on accurate assessment of risk.

The main mammographic indicator of breast cancer is the amount and distribution of the dense tissue. In addition to the density (or its surrogate, intensity), texture is thought to provide information relevant to risk assessment [32]. Several studies have appeared on the use of texture for classifying risk [25,7,24]. One problem in studying texture in screening mammograms is that the relationship between total attenuation of the x-ray beam and image intensity is non-linear. Hence the contribution to intensity of small components (such as ducts) results in an intensity signal that varies according to the local intensity of the background. Thus results reporting a positive contribution to risk assessment based on texture may be due to the indirect measurement of density. To avoid this problem, processing steps reported previously [20] were used in this study to remove the local intensity variation as well as the local intensity mean.

Any study assessing cancer risk suffers from the problem of identifying those “at risk”. A definitive statement is not possible as subjects free of cancer at the end of a study may still be at risk of developing cancer at a later time. In the absence of a gold standard, various surrogates have been developed by researchers: Wolfe [31,32] used parenchymal pattern classes to quantify risk, the American College or Radiology [1] introduced BI-RADS classes shifting the focus from structure to density patterns. In addition, genetic markers such as mutations in BRAC1/2 have been used in this context [11,12,18]. All of these criteria are reasonable but none measure risk directly.

2. Higher-order textons

A general framework for higher-order textons is as follows. Let $X^0 = \{X_1^0, X_2^0, \ldots, X_q^0\}$ denote a collection of images or a single image ($q=1$) and let $p_{i,j}$ denote pixel j in image X_i^0. Let $f^1(i,j)$ denote the feature vector of length L_1 obtained by computing L_1 features associated with pixel $p_{i,j}$. The components of $f^1(i,j)$ may be outputs from linear filters or other descriptors of local phenomena. There is no restriction to the method of feature extraction used in this step. The collection of $f^1(i,j)$ over i and j is viewed as a set of points in an L_1-dimensional feature space. A clustering method is applied to the feature space to identify a set of clusters $T_1^1, T_2^1, \ldots, T_{n_1}^1$. These clusters are the first-order textons. For each i, a new image X_i^1 is formed by assigning label $s \in 1, 2, \ldots, n_1$ to pixel $p_{i,j}$ where s is the index of the cluster closest to $f^1(i,j)$ in the feature space using an appropriate norm (usually the Euclidean distance). The images X_i^1 are called the first-order texton maps.

Second-order textons are obtained by constructing local feature vectors $f^2(i,j)$ of length L_2 on $X^1 = \{X_1^1, X_2^1, \ldots, X_q^1\}$. The features comprising the components of $f^2(i,j)$ must not involve arithmetic operations. Except for this key point, the remaining steps are the same. Thus, a clustering algorithm (not necessarily the same one as used for first-order textons) is applied to the L_2-dimensional feature space to form the second-order textons $T_1^2, T_2^2, \ldots, T_{n_2}^2$ and so on (Fig. 1). The number of textons at each level (texton order) is not necessarily the same. Final representation or classification can be based on the full collection of textons over all levels or a sub-collection.

The following toy example shows that two images (in this case strings) may be indistinguishable by first- and second-order textons but distinguishable by third-order textons. Consider two 1-dimensional binary images X^0 and Y^0, each comprising m entries labeled 1 and the rest labeled 0. Specifically, the distributions of 1s in X^0 is random but in Y^0 all the 1s appear separated by exactly two 0s (except the first and last 1). Thus $Y = (\ldots, 0, 0, 0, 0, 1, 0, 0, 1, 0, 0, 1, 0, 0, 1, \ldots, 0,\ 0, 1, 0, 0, 1, 0, 0, 1, 0, 0, 0, 0 \ldots)$. The feature vector at position i is $f(i) = (f_1(i), f_2(i))$, where $f_1(i) = Y^0(i-1)$ and $f_2(i) = Y^0(i+1)$. For string Y^0, the feature space obtained by plotting $f(i)$ for all i appears in Table 1(a). Here A is a large value that depends on the length of the string. Since the resolution in the feature space is low, clustering is not quite meaningful, but a reasonable analog is to accept three clusters $T_1^1 = (0,0)$, $T_2^1 = (1,0)$ and $T_3^1 = (0,1)$. The clusters T_j^1, $j = 1, 2, 3$ are first-order textons.