COLOR TEXTURE RETRIEVAL THROUGH CONTOURLET-BASED HIDDEN MARKOV MODEL

Zhihua He and Maja Bystrom

ECE Dept. Boston University, 8 St. Mary’s St. Boston, MA 02215, {zhihua,bystrom}@bu.edu

ABSTRACT

Two statistical models for color texture retrieval based on a hidden Markov model (HMM) in the contourlet domain are described in this paper. Through a contourlet transformation, each component of an image is decomposed into a set of directional subbands with texture details captured in different orientations. By exploiting inter-scale dependencies and in-band spatial dependencies, the distribution of the coefficients in each subband triplet (subbands of three color components at the same scale with the same orientation) can be estimated using a vector hidden Markov model. The Kullback-Leibler distance (KLD) is used to measure the difference between the distributions of query texture images and those of images in the database. The experimental results show the proposed retrieval systems yield high retrieval rates and better visual quality as compared with previous methods employing hidden Markov models for luminance component alone.

1. INTRODUCTION

Efficient retrieval of information from massive databases is an important challenge in current communication and storage applications. While images may be retrieved on the basis of color and texture independently, we study texture-based retrieval in color components in this work. We propose two contourlet-based statistical models, which exploit cross-scale and in-band dependencies for color textures, respectively.

When texture images are passed through a wavelet transform, texture information with different orientations at various scales is captured in different directional subbands. Statistical models can characterize structure repetitions in a small number of parameters. In particular, HMM-based multi-scale statistical models are naturally good candidates for texture features. Different multi-scale decomposition schemes may be employed, while HMM-based texture retrieval systems can exploit different Markov dependencies to capture texture features. In the framework of pyramid-based multi-scale analysis, three main sources of Markovity are typically considered, namely, cross-scale, cross-band and spatial in-band. The first two dependencies have been investigated in [1, 2, 3, 4], and some authors have suggested the exploitation of in-band dependencies. In [1], Crouse, et al., exploited cross-scale hidden Markov dependencies in the estimation of statistical model parameters for feature extraction in the wavelet domain, and suggested possible further extensions utilizing cross-band and in-band (spatial) dependencies. In [2], Do and Vetterli adopted a vector HMM model for characterization of textures in the steerable pyramid wavelet domain, investigating both cross-scale and cross-band dependencies. Fan and Xia [3] have proposed a HMM model for texture retrieval and synthesis also considering cross-band correlation. Recently, Po and Do [4] have studied the statistical modelings in contourlet domain, and exploited cross-scale dependencies for their HMM approach to the texture retrieval problem. A HMM exploiting both cross-scale and in-band dependencies in contourlet domain is described in [5] for texture retrieval.

Many texture retrieval systems focus on the luminance components of the image; however, color has always been an important feature for image retrieval. In this paper, we describe a system where color information is fused into the contourlet-based HMM for texture retrieval, which can be viewed as a color extension of previous work [4, 5]. Two classes of dependencies are considered, namely, cross-scale dependencies and in-band spatial dependencies. Given an image, each color component is passed through a contourlet decomposition. Subbands from different color components with the same orientation at the same scale are grouped to form a set of subband triplets. The distribution of the coefficients in each triplet is modeled as a 3D Gaussian mixture. When in-band spatial dependencies are considered, we construct a HMM which links coefficient within a subband triplet where the bands are scanned according to their orientation, and the parameters of the Gaussian mixtures for each triplet are estimated and adopted as texture features. When cross-scale dependencies are considered, each triplet at scale $L$ is associated with a parent triplet at scale $L + 1$, and the parameters of the distribution for each triplet are estimated as in [4]. Brief introductions to the contourlet decomposition and the HMM are given in Section 2.1, while details of the two models are discussed in...
2.2. Comparisons of the performance between the proposed two models and related methods are given in Section 3.

2. CONTOURLET-BASED COLOR HIDDEN MARKOV MODELS

2.1. Contourlet and Hidden Markov Model

Designed by Do and Vetterli [6], the contourlet transform provides an efficient multi-scale directional representation of an image. In a contourlet decomposition, the images are first passed through a pyramid Laplacian decomposition, then the high frequency subband images from each scale are passed through directional filters with prescribed orientation resolution. Fig. 1(a) shows a block diagram of a two-level contourlet decomposition, and Fig. 1(b) shows the subband images of test image pepper [7], where significant cross-scale and in-band information is observed [4]. Compared with other directional wavelet transform, e.g., the steerable pyramid [8], complex wavelet [9], and curvelet [10], the contourlet provides more flexible directional filtering and an efficient representation.

Hidden Markov models capture dependencies between the states of observed data. Recently, HMMs have been widely used in wavelet-based signal analysis. In particular, for contourlet-based image analysis, given a subband with orientation $\theta$, the majority of the coefficients exhibit small variance (denoted as being in state 0), this corresponds to smooth regions in the image. Only the coefficients corresponding to edges with direction near $\theta$ yield high variance (denoted as being in state 1). Experimental results support the Gaussian assumption for the distribution of coefficients conditioned on their states [4]. Thus, the distribution of each directional subband can be modeled by a Gaussian mixture whose parameters can be estimated by a HMM. For example, the parameters of a $k$-state Gaussian mixture model include $p_i, i = 0, \ldots, k-1$, the prior probability for each state, $m_i, C_i, i = 0, \ldots, k-1$, the means and variances/covariance of the Gaussian distributions conditioned on states, and $a_{ij}, i, j = 0, \ldots, k-1$, the transition probabilities from state $i$ to state $j$. The Gaussian mixture model distribution is then given by $\sum_{i=0}^{k-1} p_i \frac{1}{(2\pi)^{d/2}|C_i|^{1/2}} \exp(-x - m_i)'C_i^{-1}(x - m_i)/2$ given $x$ in $d$-dimensional space. A typical choice of $k$ in a wavelet-based HMM is two; a 3D color space is considered in this work. The HMM parameters can be estimated through an efficient algorithm based on expectation maximization (EM) [1,11]. The EM algorithm will only yield a local minimum solution, thus, the construction of the HMM is important. There are three classes of state dependencies in contourlet-based HMM, namely, cross-scale, spatial in-band and cross-band within the scale. We focus on the first two classes of dependency, which are more significant than the third [4].

2.2. Two Contourlet-Based Color HMMs

Consider the contourlet decomposition of a color image, which yields three sets of directional subband images, corresponding to the three color components. Subbands with orientation $\theta$ at scale $L$ are grouped together to form a triplet, denoted as $T^L_\theta = [r^L_\theta, g^L_\theta, b^L_\theta]$, which is the child of a parent triplet with same/similar orientation $\theta$ at scale $L + 1$. Each triplet coefficient cimentor $T^L_\theta(i)$ can be associated with a hidden state variable $S^L_i$, which varies between state $s^0$ and state $s^1$. In the contourlet decomposition, each coefficient cimentor in a parent triplet has four children cimentors in a child triplet. The parent and children coefficients exhibit cross-scale dependencies, while the coefficients within each triplet exhibit directional spatial in-band dependencies.

When in-band spatial dependencies are considered, each triplet is scanned along its direction, i.e., vertical triplets are scanned vertically and horizontal triplets are scanned horizontally. The distribution of $T^L_\theta(i)$ is modeled as a 3D Gaussian conditioned on $S^L_i$. The transition probability $a_{mn}; m, n = 0, 1$ between $S^L_i = s^m$ and $S^L_{i+1} = s^n$ captures the spatial directional Markov dependency. Similar to [1], we tie triplet coefficient cimentors spatially to obtain robust training, i.e., every vector pair $\{T^L_\theta(i), T^L_\theta(i + 1)\}$ is assumed to be independent, identically distributed. Using the EM algorithm we can solve for the distribution of the subband pair with parameters $p_i, m_i, C_i, a_{mn}; i, m, n = 0, 1$. Estimation of transition matrix $a_{mn}$ exploits mainly directional spatial correlation and estimation of covariance matrix $C_i$ and mean $m_i$ for each state exploits color information. Due to the down-sampling structure in contourlet [12], only horizontal and vertical spatial dependencies for (near) horizontal triplets and (near) vertical triplets are considered in this work; experimental results show that this is sufficient for retrieval.

When cross-scale dependencies are considered, the distribution of $T^L_\theta(i)$ is also modeled as 3D Gaussian conditioned on $S^L_i$. Now the transition probability $a_{mn}, m, n = 0, 1$ is defined as between $S^L_i = s^m$ and $S^{L+1}_i = s^n$ which captures the cross-scale Markov dependency, and $\rho(i)$ is the parent coefficient cimentor of coefficient cimentor $i$. Similar to [4], a hidden Markov tree can be constructed, and EM algorithm can be used to solve for the parameters of the Gaussian mixtures of each subband. Note that for the cross-scale case, all the directional subbands are considered.

The parameters for the distributions of all subbands are used as the texture features, and the Kullback-Liebler distance (KLD) is adopted as similarity measure for distance between different distributions.
3. EXPERIMENTAL RESULTS

Using a texture database generated from MIT VisTex database [13], we evaluate the two proposed color texture retrieval models against the approaches of [4, 5], which are based solely on image intensity using cross-scale dependencies and in-band dependencies. A group of 163 texture images from VisTex is selected to generate 652 images, by dividing each image into four sub-images. In our experiments, a three-level contourlet decomposition is adopted, where 16 directional subbands are generated with resolution $\frac{2}{8}$ on both the first and second levels while 8 directional subbands are generated on the last level. Texture features estimated through directional HMM are compared through the KLD, estimated by a Monte-Carlo method $D(P_q(x)P_d(x)) \approx \frac{1}{N} \sum_{n=1}^{N} \left[ \log P_q(x^n) - \log P_d(x^n) \right]$, where $x^n$’s are randomly generated data following query image distributions $P_q(\cdot)$. In our experiment, every image in the database is used as query image for retrieval through the database. The retrieval rate is defined as $R_M = N_M / 652$, where $N_M$ is the number of successful retrievals when only $M$ candidate images are preserved. The retrieval results using proposed two models and those of [4, 5] are compared in Table 1, where the number of texture features (number of Gaussian mixture parameters) for each image is also included in the last column (as an estimation of storage cost). The proposed two retrieval methods give similar retrieval rates, but the method exploiting in-band dependencies requires fewer features. A direct combination of the proposed two methods gives higher retrieval rates, at the expenses of more features. As compared to the two luminance component-based schemes [4, 5], both proposed methods give higher retrieval rates. In particular when only top 5 and top 10 candidate images are considered, over 10% higher rates are observed. When both retrieval rate and number of features are considered, the proposed in-band color HMM model turns out to be the best choice. Fig.’s 2 (a)-(d) show four sets of sample retrieval results based on two proposed models and those of [4, 5], respectively. Note that the proposed models return ordered retrieved images that are more visually similar than those of [4, 5].
Fig. 2. Query image bark on the left followed by the top 6 retrieval results using the proposed color HMMs and luminance-based schemes of [4, 5]. (a) Retrieval results using proposed color HMM exploiting cross-scale dependencies. (b) Retrieval results using proposed color HMM exploiting in-band dependencies. (c) Retrieval results using the method of [5] (d) Retrieval results using the method of [4].

4. REFERENCES


