Texture Classification

Textures, Filter Banks and Patch Based Approaches

Student:

Mohammad Doosti Lakhani
98722278

Instructor:

Dr. Soryani

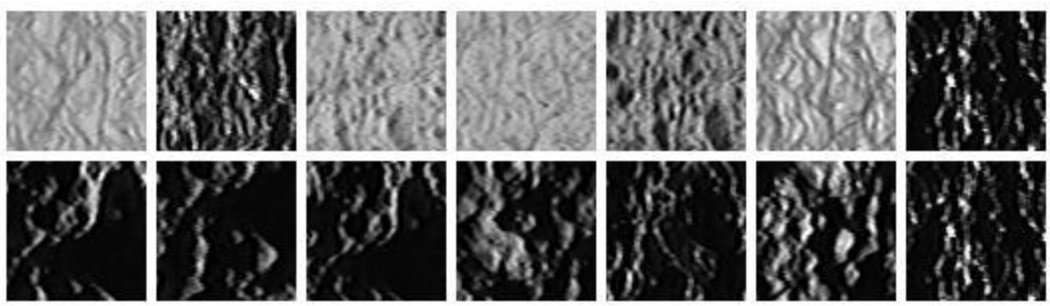
Course:

Computer Vision

Intro

This article is about classifying materials using their image without consideration of illumination or color intensity variations to some predefined classes. So, all input images to the weak classifiers are in grayscale mode and no color info has been preserved.

One of the major challenges in classifying materials using textures is that it is not similar to conventional classification tasks where objects have defined structures and classifier only needs to learn that particular structure. The main reason is that textures vary due to degree of illumination and angle of camera, which produces excessively wide range of differences for identical texture.
For instance, in image below, in the first row, angle is constant but illumination differs and in second row, angle and illumination are varying and constant respectively.



The other dominant problem is the small inter-class variation regarding textural info.

To answer these problems, this article has dichotomized to two statistical approaches consisting of

1. Without Filter
2. Using Filters

CUReT

There are 61 texture classes present in the database and each has been imaged under 205 viewing and illumination conditions. The effects of secularities, inter-reflections, shadowing and other surface normal variations are plainly evident and this makes the database far more challenging for a classifier. The limitations of the database are a lack of significant scale change and limited in-plane rotation

Classification W/O Filters

Definition

It has been demonstrated that textures can be classified using the joint distribution of intensities over a small patch such as 3x3.with similar performance w.r.t using filter banks.

Image Patch Based

This approach is identical to the classification with Filters but the only difference is that in step of convolving training images with bank filters, raw pixel intensities of an N*N patch has been flattened. This is called Joint Classifier.

So, what is does can be summarized in the way that first, after flattening the patch, central pixel of patch will be mapped into a specific number of bins and other pixels of that patch will be mapped to textons as before. This representation can be interpreted as a Markov Random Field distribution of texture. Note that the final frequency is a histogram of bins-textons while in using filter banks, it was histogram of only textons.

The reason that this approach outperform filter bank approaches is that even small patches such as 3x3 can be used to learn the distribution of objects bigger than this patch size by using a collection of small patches.

Also using a small patch size enable the algorithm to capture textures regardless of its scale.

Results

From experimentation these can comprehended:

1. Increasing patch size decreases performance of model
2. small Patch Based outperforms all of scale-orientation-invariant filter bank approaches.
3. Regarding same filter size, Patch Based outperforms filter bank approaches in any arbitrary filter size.

Classification with Filters

Definition

The algorithm consists of learning and classification stages.
In the learning stage, different variation of same image has been fed to model. Then model convolves these images with a bank of predefined filter and the responses(textons) are collected as a dictionary using K-means clustering. Note that for each class of texture we have a cluster with multiple centers.

Now we need to learn a model for each texture class. To achieve this, for an arbitrary training image, first convolving it with filter bank then labeling the result with closest responses from previous step(textons). The model of that particular training image is the frequency of occurrence of textons at the time of labeling. As we have a model for each training image, there are multiple models for each texture class. (These can be reduced, please see section Reducing the Number of Models)

At the time of classification, same procedure as previous step needs to be taken and as the final step closest histogram using K-Nearest-Neighbor using Chi-Square distance measure enable us to select corresponding texture class.

Filter Banks

Filter banks need to be chosen by considering their performance of classification using their clustered textons.

But how to generate filter banks? Filter banks can be generated by including first or second derivatives of Gaussians or Laplacian of Gaussians in different scales and orientations. But to reduce number of response filters and having orientation invariant filters which increases performance, Maximum Response 8 can be derived by considering only the maximum filter response across all orientations to handle oriented patches. For reducing dimensionality, taking maximum regarding scale can also be considered to handle slow and bulky memory operations.

As an instance, Schmid filters are constructed using isotropic Gabor-like filters which are rotationally invariant.

$$F(x,y,\tau,\sigma) = F_0(\tau,\sigma) + \cos\left(\frac{\sqrt{x^2+y^2}\,\pi\,\tau}{\sigma}\right) e^{-\frac{x^2+y^2}{2\sigma^2}}$$

First term added to obtain zero DC which helps filters to be more robust to illumination changes and obtain invariance to intensity translations. 13 filters with scales from 2 to 10 and number of cycles of harmonic function between 1 and 4 has been used. Support size of 21x21 has been used for grayscale filters.

The anisotropic Gaussian filtering method allows fast calculation of edge and ridge maps, with high spatial and angular accuracy. the anisotropic Gaussian can be decomposed by dimension. This appears to be extremely efficient from a computing perspective.

$$g_\theta(x, y; \sigma_u, \sigma_v, \theta) = \frac{1}{\sqrt{2\pi}\,\sigma_x} \exp\left\{-\frac{1}{2}\frac{x^2}{\sigma_x^2}\right\} * \frac{1}{\sqrt{2\pi}\,\sigma_\varphi} \exp\left\{-\frac{1}{2}\frac{t^2}{\sigma_\varphi^2}\right\}$$

Demonstrates separability of anisotropic gaussian filter into two 1D gaussian filters. This can implement by first applying a 1-D Gaussian convolution in the -direction. The resulting image is then convolved with a 1-D Gaussian in the gamma-direction yielding the anisotropic smoothed image.

By applying recursive anisotropic Gauss filters the original image is filtered at different orientations and scales, and the maximum response per pixel over all filters is accumulated. At each pixel, the local orientation and best fitting ellipse is available to be further processed.

Reducing the Number of Models

1. Model Selection: Using machine learning approaches such as K-Medoid and Greedy to choose a subset of models while maximizing accuracy.
2. Geometric Descriptor: Incorporating some of geometric information into model to have illumination-angle invariant descriptor for images.

Results

The CUReT dataset has been used for this experiment.

From the experiments, these points can be extracted:
1. Using invariant filter banks increases accuracy to %35, and 4X to 12X decrease in number of filters provides huge speed and less memory consumption.
2. Using lower number of response filters helps the clustering algorithm so MR algorithms has advantage of using this technique.

Conclusion

Filter Bank vs. Small Patch Based:

1. Filter banks because of their large filter size learn fewer texture samples from training images
2. As the path size is huge in filter bank approaches, the estimation of boundary regarding different textures is not an acceptable segmentation.
3. As we use the response of filters in bank filter approaches, it smooths texture which causes loss in fine details.

The performance of the MR8 filter bank is always worse than any other classifier at the same neighborhood size. The MRF "Best" curve shows results obtained for the best combination of texton dictionary and number of bins for a particular neighborhood size. For neighborhoods up to 11×11, dictionaries of up to 3050 textons and up to 200 bins are tried. For 13×13 and larger neighborhoods, the maximum size of the texton dictionary is restricted to 1220 because of computational expense. The best result achieved by the MRF classifiers is 98.03% using a 7×7 neighborhood with 2440 textons and 90 bins. The best result for MR8 is 97.43% for a 49×49 neighborhood and 2440 textons.

So, in simple words, Small Patch Based approach outperforms Bank Filter Based approach!