Texture classification based on texton histograms

Edgar A. Margffoy-Tuay
Universidad de los Andes
201412566
ea.margffoy10@uniandes.edu.co

Abstract

Texture detection represents historically one of the main approaches to contour and figure segmentation based on the search and detection of repetitive patterns which contrast with other textures present on the image. This problem has been previously addressed by representing textures as a set of feature activations obtained after applying a filter bank over a set of training images, those features, named textons can be used to represent an image and can be used to classify and detect objects and contours. This report pretends to introduce and analyse this framework by implementing a texton-based classifier over the Ponce's group Texture Database.

1. Introduction

A texton is defined as a the minimal unit of human perception and image composition (Analogue to a phoneme in linguistics), this concept proposed by Béla Julesz in 1981, is based on the decomposition of perception as the response to stimulii induced by linear patterns such as intersection segments and crossings. Based on this notion, Malik *et al.* [3] proposed a method in which an image can be represented by the activation responses of each pixel to a set of linear and circular filters, designed to take in account essential visual features such as rotation, scaling, contrast difference and border detection. The geometrical differences between the filters, enables us to discriminate and differentiate different textures based on the patterns that are present, for instance, circular and linear filters allow to distinguish a polka-dot texture from a brick-like one. After grouping each filter activation pixel-wise, it is possible to cluster and quantize each one of these activations (Namely textons) and define a dictionary of textons over a set of images, the texton representation then can be used to describe an image as an histogram that describes the proportion of each class of textons present on the image.

From this approach, it is possible to appreciate that visual objects of the same category may present similar texton representation, due to the geometric and topological similarity between them, this includes similar texture activations in both contour and innings of the objects subject this representation. From this perspective, a classification framework can be proposed in which, images that present a texture label can be represented by its texton histogram that can be related to an specifc category.

2. Materials and Methods

About the Dataset

The Ponce's group texture dataset[1] [2] is a comprehensive image set that contain one texture category (Out of 25 available categories) per sample. Each grayscale image is of size 640×480px, and

[1]Available at http://www-cvr.ai.uiuc.edu/ponce_grp/data/

compasses textures like wood, glass, marble, brick, among others. To generate the texton dictionary that describes the images of the dataset, a subset of seven images per class were chosen, also a subset of the original images was randomly selected to conform the image testing set of the texton-based texture classifier.

About the Implementation

To implement the framework proposed on [3], it was necessary to define a set of linear filters, based on Gaussian composition and differentiation, each of the filters (Contrast, edge and circular detectors), was defined on six possible scales and twenty rotations with increasing variances departing from the value 0.1, the scaling was done according to the silver ratio rule $(1 : \sqrt{2})$, this parameter selection allows to capture and select minor local patterns that differentiate similar but different textures categories (*i.e.,* different wood classes). Also increasing the number of filters can be of benefit to adquire more data that can be essential to capture more valid features that can improve the Statistic/Machine Learning classifier accuracy. However, increasing the number of filters may imply an increase on the computational power requierd to process the texton clustering, and therefore the total computing time.

To describe an image, a texton dictionary of size 128 was used, this result was computed by vector quantization employing K-Means. To improve the accuracy of the clustering result, the KMC^2 centroid seeding algorithm [1] was employed, this routine enables to select the best initial centroids based on a hidden Markov model. After computing the texton dictionary, it is possible to process all the dataset images as vectors defined on the Texton representation space

References

[1] O. Bachem, M. Lucic, S. H. Hassani, and A. Krause. Approximate k-means++ in sublinear time. In *Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence*, AAAI'16, pages 1459–1467. AAAI Press, 2016.

[2] S. Lazebnik, C. Schmid, and J. Ponce. A sparse texture representation using local affine regions. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(8):1265–1278, 2005.

[3] J. Malik, S. Belongie, T. Leung, and J. Shi. Contour and texture analysis for image segmentation. *International Journal of Computer Vision*, 43(1):7–27, 2001.