Abstract processes in texture discrimination

J. M. H. du Buf

Signal Processing Laboratory, Swiss Federal Institute of Technology, EPFL-Ecublens, CH-1015 Lausanne, Switzerland

Received 13 July 1990; revised 28 October 1991; accepted 4 November 1991

Abstract—In this study some experiments on texture segmentation are reported using the local Gabor power spectrum. The techniques applied are: (1) supervised pixel classification; (2) boundary detection by spectral dissimilarity estimation; (3) region-based segmentation based on Gaussian spectral estimation; and (4) the same as (3) but based on central moments of the local spectrum. It is shown that very-acceptable-to-excellent results can be obtained. It is argued, however, that the shortcomings of region-based and boundary-based approaches require that both processes should act in parallel, not only in digital image processing but also in the modelling of visual perception.

1. INTRODUCTION

The purpose of this article is to report some experiments in digital image processing, on texture segmentation to be precise, and to argue that the problems encountered here may have implications for the modelling of visual perception. It is stressed that the techniques to be considered are not meant to provide definitive solutions to either texture segmentation in digital image processing applications or the modelling of human texture discrimination, but rather an indication of current approaches which when refined might further synchronize the two research areas.

In digital image processing, a number of qualifications are applied to address and distinguish specific texture-segmentation processes. The same terminology is used in this article, which starts with a brief but necessary overview, and then turns to a discussion of some aspects related to human perception and the simplifications which may lead to good working models.

Digital image processing

Normally, texture segmentation is obtained by the (independent) processes of feature extraction, feature reduction or selection if the feature dimensionality is too large, followed finally by segmentation. The goal of feature extraction is to map differences in spatial structure to differences in amplitude. As a result, different textures should have different feature values, and in the n-dimensional case different feature vectors, while the feature values (vectors) should be constant within the same texture. Of course, if the textures are not extremely homogeneous, this automatically leads to ambiguities: what is homogeneous and what is not, or, translated into segmentation performance, which result is acceptable and which is not. In any case, feature extraction always implies the application of a texture model, the model parameters being used as features. Apart from the fact that a unified texture model does not yet exist, there is the problem that good models have, in general, many parameters. Hence, for the computational efficiency of a segmentation algorithm, a feature-selection algorithm may be required. This can be done in a supervised way, by visual
inspection and selection of feature images, or in an unsupervised way, for example by applying a Karhunen–Loève transform. The reduced feature vectors can then be analysed by a segmentation algorithm to detect the homogeneous regions.

Furthermore, it is usual to make a distinction between region-based and boundary-based segmentation algorithms. Region-based methods try to extract homogeneous regions, whereas boundary-based methods try to detect the inhomogeneities. This distinction may appear somewhat artificial, because detecting inhomogeneities and deciding to include some pixels in a region are very similar processes. But the requirements are the same: accurate and closed boundaries. In general, the performance of a segmentation algorithm depends strongly on the quality (signal-to-noise ratio) of the feature set, and therefore on the texture homogeneity within regions and the differences between regions. A complication in judging the performance of a segmentation algorithm is therefore the fact that the segmentation quality, that is the combination of texture classes correctly identified and the boundary accuracy, also depends on the feature-extraction and selection processes.

Yet another distinction that is often made is between supervised and unsupervised segmentation. In this case, supervised segmentation means some form of pixel classification, and implies that training sets are available. These may consist of prototype textures in a database, or may be defined as parts of homogeneous regions in an input image to be segmented. In other words, one knows already what to look for. Unsupervised segmentation, on the other hand, means that no prior information is used, neither texture classes nor training areas in the input image. Such methods seek the best solution by analysing the feature space, without any prior information or human intervention. In the next sections examples are given of supervised, unsupervised, boundary-based, and region-based approaches.

Local spectral decompositions

Innumerable texture feature extraction methods have been proposed in the past (see e.g., du Buf et al. 1990). Recent studies have shown that a Gabor decomposition, that is a local spectral decomposition by means of a bank of frequency- and orientation-selective filters, is applicable as well (e.g., Turner, 1986; Beck et al., 1987; Porat and Zeevi, 1989; Bovik et al., 1990). This new approach is motivated by its resemblance to a process in low-level human vision, i.e. the coding of images by means of frequency and orientation-selective simple cells in the striate cortex (grouped in hypercolumns). The Gabor decomposition being the first mathematical model proposed (Marčelja, 1980), this is not necessarily the best and only model (Koenderink and van Doorn, 1990; Stork and Wilson, 1990). Alternatives are the pseudo-Wigner distribution (Reed and Wechsler, 1991), wavelets (Mallat, 1987), finite prolate spheroidal sequences (Wilson and Spann, 1988), and Hermite polynomials (Koenderink and van Doorn, 1990). The real problem is probably not which scheme to choose, but what to do with the image representation obtained in order to extract geometrically and semantically meaningful image attributes. In any case, texture models which are based on an image representation obtained by a bank of for example Gabor filters might unify or replace existing texture models. Worth mentioning are the texton theory (Julesz and Bergen, 1983) and the raw primal sketch (Marr and Hildreth, 1980), which were proposed for the early stages in vision (so-called preattentive vision with effortless texture segregation).

A major difficulty which arises in applying a Gabor image representation is the