Finding perceptually dominant orientations in natural textures

ROSALIND W. PICARD and MONIKA GORKANI
Perceptual Computing Section, MIT Media Laboratory, 20 Ames Street, Cambridge, MA 02139, USA

Received for publication 22 September 1993

Abstract—An algorithm for detecting orientation in texture is developed and compared with results of humans detecting orientation in the same textures. The algorithm is based on the steerable filters of Freeman and Adelson (IEEE Trans. PAMI 13, 891–906, 1991), orientation-selective filters derived from derivatives of Gaussians. The filters are applied over multiple scales and their outputs non-linearly contrast-normalized. The data for humans were collected from forty subjects who were asked to identify 'the minimum number of dominant orientations' they perceived, and the 'strength' with which they perceived each orientation. Test data consisted of 111 grey-level images of natural textures taken from the Brodatz album, a standard collection used in computer vision and image processing. Results show that the computer and humans chose at least one of the same dominant orientations on 95 of the natural textures. Of these textures, 74 were also in 100% agreement on the location of all the dominant orientations chosen by both humans and computer. Disagreements are analyzed and possible causes are discussed. Some apparent limitations in the current filter shapes and sizes are illustrated, as well as some (surprisingly small) effects believed to be caused by semantic recognition and gestalt grouping.

1. INTRODUCTION

Orientation is one of the most perceptually significant components in texture recognition (Tamura et al., 1978; Rao and Lohse, 1992) and in visual attention (Treisman and Gelade, 1980; Wolfe et al., 1989). Psychophysical evidence exists that humans use orientation as a cue for discriminating textures (Tamura et al., 1978; Phillips and Wilson, 1984; Julesz, 1991; Rao and Lohse, 1992), and physiological experiments suggest the existence of orientation selective mechanisms in the human visual system (Hubel and Wiesel, 1968; Webster and De Valois, 1985). Local orientation information has also been argued to play a critical role in curve detection (Zucker, 1985).

Extraction of orientation over a large scale can be used for rotating images to align them before beginning closer comparison, a process possibly done by humans during pattern recognition (Shepard and Cooper, 1982). An observer can also obtain shape and perspective information from texture (Aloimonos and Shulman, 1989; Choe and Kashyap, 1991) and consequently orientation may play a key part in this process. Orientation at a finer level is also a fundamental component of texture and pattern—wood grain, sand ripples, parquet floors, bookshelves, and parking lots; all these contain dominant orientation information which can be related to their formation or function.

Not only are orientation detection cells at work in the low levels of the visual cortex and in a possibly higher level process of pattern rotation for alignment, but orientation can also be considered a semantic feature of image data. 'Semantic'
here refers to the human use of orientation in natural language when describing visual information. Semantic features are important in rapidly growing new applications such as searching for images by their visual context (Picard and Kabir, 1993).

1.1. Applications of orientation

Imagine in a few years when every home has ‘video on demand’ and all art collections, movies, patent libraries, photo albums, books, journals, and other visual collections are accessible online. There will be exabytes of information, making it impossible to find a particular image or video clip without spending inordinate amounts of time looking for it. One can currently search through text for keywords, but not through video for keyframes. In the next several years it will become tremendously important to have automated tools that can search for visual information whether or not it has a description attached. A likely scenario is one where a user shows the computer a pre-existing pattern, and requests all other patterns ‘like’ this one.

In such a scenario, the search for a similar image should agree with the human’s notion of similarity. To succeed, the computer must know which features attract the human’s attention, and how to combine these features for locating perceptually similar patterns. The goal is to get the computer to identify textures the person would identify if he or she had time to look through them all. As orientation is one of the most significant features for human attention and texture matching, it is important that algorithms which recognize orientation be developed.

The results of this study on orientation apply directly to the image search problem. Furthermore, in the future when search tools become more semantic and there are pre-stored object descriptions to be matched, then one will also be able to associate to keywords such as ‘brick wall’ the feature, ‘usually has two dominant orientations’. The orientation detection algorithm can then do a much faster search for the corresponding visual data than it could do without this information. Even if a portion of an image has no dominant orientations, that is still important information to use in speeding up comparisons. For example, two primary categories of textures have long been recognized: structural and statistical (Haralick, 1979). Orientation is a salient feature that can be used to decide which of these (or other) categories a texture is close to, i.e. the statistical is likely to not have more than one dominant orientation. Given such information, one may select a model more suitable to recognizing that class of pattern. Detailed model features can then be extracted relative to the dominant orientations to provide invariant measures within inhomogeneous data.

1.2. Orientation over scale

How much pattern recognition can be achieved using only orientation information is an open question. A key difficulty is that it appears to be important to gather the information over multiple scales (Bergen and Adelson, 1988). Also, orientation information is complicated by its interactions with effects such as contrast (Heeger, 1991), similarity grouping and gestalt effects (Hamey, 1992), and prior knowledge present when a pattern is recognized (Richards, personal