Pseudo-Euclidean Representation for Effective Multi-Class Texture Classification

A. K. C. Wong
H. C. Shen
P. W. Wong

Systems Design Engineering
University of Waterloo

Abstract

This paper describes how the theory of pseudo-Euclidean embedding can be applied for effective multi-class texture classification. The textures are classified according to the nearest neighbour rule based on our recently developed texture metric. Our objective is to significantly reduce the amount of computation from an exhaustive search scheme. For this purpose, a distance preserving vector space representation of the texture data base is constructed. This representation facilitates the selection and organization of a small subset of class prototypes so that the search for the nearest class can be carried out economically. This methodology is demonstrated by experiments on 720 texture samples belonging to eight classes. On average, a reduction of close to 70% is achieved.
Keywords: *pseudo-Euclidean spaces, distance preserving embedding, multi-class texture classification*

1 Introduction

The effectiveness of a supervised classification scheme depends on what distance function is used to discriminate the samples as well as how the search for the correct classes are conducted. For multi-class texture classification, we have recently developed a texture metric which has been demonstrated to be highly effective in discriminating perceptually similar textures of real objects[1,2]. In our original approach, classification is performed by simply computing the distances of the texture to be classified to *all* of the samples in the texture database. This exhaustive method is clearly un-economical when the database is large. In this article we will propose an improvement of the exhaustive method by addressing the issue of effective search. Our methodology employs a distance preserving vector space representation of the texture data. This representation facilitates the selection of a small subset of class prototypes which furnish all essential spatial information for nearest neighbour classification. Accurate and economical classifications can then be achieved by comparing against these prototypes alone.

In order to identify such prototypes, one must first have an *overall* picture of the database. If we could visualize our data geometrically, we would be able to describe them in terms of the relative locations of the classes, as well as their shapes and boundaries. By treating the texture elements as a finite metric set [4], we can *embed* them (i.e. represent them as points) in a pseudo-Euclidean vector space. The geometric configuration arising from the representation space would then provide the basis for the selection of prototypes as well as the development of a hierarchical classification scheme.

Though our present application is specifically on texture classification, it is in fact an illustration of a new approach to a very general class of problems. Provided we have a good metric, our methodology can be generally applied to any type of data which cannot originally be represented as pattern vectors in a feature vector space.

2 Pseudo-Euclidean Representation of Texture Data

For completeness, we will briefly describe the theory leading to our methodology. The reader is referred to the original references [3-6] for complete details.

2.1 Basic Concepts and Definitions

Let our texture metric be denoted by f. It has been shown in [2] that f satisfies the following properties:

reflexivity for all texture samples t_i, t_j, $f(t_i, t_j) \geq 0$ with equality iff $t_i = t_j$,

symmetry for all t_i, t_j, $f(t_i, t_j) = f(t_j, t_i)$,, and

triangle inequality for all t_i, t_j, t_k ,

$$f(t_i, t_k) \leq f(t_i, t_j) + f(t_j, t_k).$$

Let $S = \{t_0, \ldots, t_N\}$ be a set of $N + 1$ texture samples. S, together with f, is referred to as a *finite metric set*.