Multi-module Image Classification System*

Wonil Kim[1], Sangyoon Oh[2,**], Sanggil Kang[3], and Dongkyun Kim[1]

[1] College of Electronics and Information Engineering at Sejong University, Seoul, Korea
wikim@sejong.ac.kr, kdk0909@paran.com
[2] Computer Science Department at Indiana University, Bloomington, IN, U.S.A.
ohsangy@cs.indiana.edu
[3] Department of Computer Science, The University of Suwon, Gyeonggi-do, Korea
sgkang@suwon.ac.kr

Abstract. In this paper, we propose an image classification system employing multiple modules. The proposed system hierarchically categorizes given sports images into one of the predefined sports classes, eight in this experiment. The image first categorized into one of the two classes in the global module. The corresponding local module is selected accordingly, and then used in the local classification step. By employing multiple modules, the system can specialize each local module properly for the given class feature. The simulation results show that the proposed system successfully classifies images with the correct rate of over 70%.

1 Introduction

The fast developing digital multimedia technology enables us to access much more information through the Internet and TV than any time before. The information is easily accessible, but it brings another side: how we store and manage these rich digital videos and images so that we can collect the proper information whenever and wherever we want. Because of recent hundreds of TV sports news programs broadcasted from all over the world, TV viewers need to choose the most interesting news and watch its highlight channel. Also, in order to manage a digital library for sports videos and images, we need an automatic image classification system.

The main purpose of this paper this paper is applying MPEG-7 to sports images for feature extraction and classification system. By analyzing MPEG-7 descriptors, we create a prototype system that can be used for sports image classification techniques under visual environments, and introduce an effective methodology of image classification via experiments. Our approach we present in this paper employs multi-module neural networks for the sports image classification system. The input value for the neural network is one of the values of visual features extracted by MPEG-7 descriptors.

In the next section, we discuss several related methods of the image classification including neural network and feature extraction using MPEG-7 descriptors. Then, we propose a Multi-Module Image Classification System in Section 3 and then discuss some simulation environments and results in Section 4. We conclude in Section 5.

* This paper is supported by Seoul R&BD Program.
** Author for correspondence : +1-812-856-0751.

H. Legind Larsen et al. (Eds.): FQAS 2006, LNAI 4027, pp. 498 – 506, 2006.

2 Related Works

2.1 Multi-module Approach

Given a complicated input-output mapping problem, it is hard to design a single network structure and solve the problem with this architecture. It usually takes a long time to train a monolithic network and may not produce a good generalization result. Since the mapping may be realized by a local method which captures the underlying local structure [1], there has been considerable research [2, 3, 4, 5, and 6] designed to take advantage of a modular network architecture. In this paradigm, each module is assigned to a specific (local) area and focuses only on its special area. Learning is more efficient when a neural network is organized in this way.

A neural network is said to be modular if the computation performed by the network can be decomposed into two or more modules that operate on distinct inputs without communicating with each other [7]. The outputs of the module are mediated by an integrating unit that may not feed information back to the modules. In particular, the integrating unit decides how the outputs of the modules should be combined to form the final output of the system, and decides which modules should learn which training patterns. Modular networks utilize the principle of divide and conquer, which permits us to solve a complex computational task by dividing it into simpler subtasks and then combining their individual solutions [8].

The use of a local method offers the advantage of fast learning and therefore requires relatively few training iterations to learn the task. Alternatively, an approximation may be realized using a global method that captures the underlying global structure of mapping. The use of global methods offers the potential advantages of smaller storage requirement and better generalization performance. The use of a modular approach may also be justified on neurobiological grounds. Modularity appears to be an important principle in the architecture of vertebrate nervous systems, and there is much that can be gained from the study of learning in modular networks in different parts of the nervous system.

2.2 Image Classification

Image classification is the core process of digital image analysis. It is used in many areas like remote sensing and image retrieval. Remote sensing is the acquisition of meaningful information from an object by a recording device that is not in physical or intimate contact with the object. For example, image classification is applied to a data interpretation process of remotely acquired digital image by a Geographic Information System (GIS).

The image retrieval also uses image classification. A user requests an image by query and it returns an image (or a set of ordered images) from its image database by matching features, like color histogram and textual measures, of a query image with those of the database images. Image classification is also used to create image databases and adding images to it for the image retrieval system. The system extracts semantic description from images and putting them into semantically meaningful categories. We focus our related work survey on this kind of systems.

For content-based image retrieval systems, one of the classic classification problems is city images vs. landscapes. Gorkani and Picard [9] separate urban images and