内99-27

早稲田大学大学院理工学研究科

博 士 論 文 概 要

論 文 題 目

Studies on Video Coding Methods

(ビデオ符号化に関する研究)
(方式)

申 請 者

趙 庭栄

TINGRONG ZHAO

電子・情報通信 システムVLSI

1999年 11月

Digital image and video compression has become an increasingly important and active field. Progress in compression algorithms, VLSI technology, and coding standards has made digital video an enabling and penetrating technology for many applications. First of all, many new compression algorithms have been developed, which allow transmission or storage of digital video with good quality at reasonable data rate. Secondly, advances in VLSI technology make it possible to implement sophisticated compression algorithms for real-time applications in a cost-effective manner. Finally, standards in image and video compression are emerging rapidly to provide common communication platforms. As a result, new products are being marketed, new trials are being conducted, and new services are being provided, with digital video compression being the core technology.

Generally speaking, video signal contains a significant amount of redundancy within and between frames. The ultimate goal of video coding is the bit-rate reduction for storage and transmission by exploring redundancy existing within and between frames and to encode the video signal into a "minimum set" of information. Motion estimation and compensation technology is recognized as the most powerful tool for exploiting the redundancy existing between video frames (we call this kind of redundancy as temporal redundancy). Nevertheless, motion estimation and compensation requires a large amount of computations. In particular, it has often become a bottleneck problem in real-time applications if the traditional full search method is used.

We have proposed three fast motion estimation schemes that their superior performances have been proved by experimental results. The first scheme uses feature vector instead of individual pixel to conduct motion estimation. This scheme not only reduce search complexity, but also have many advantages based on hardware consideration such as data fetching, basic matching computation, and memory management. In the second scheme, we proposed novel multiresolution motion estimation algorithms by exploiting the correlation of motion vectors (MV) existing in spatially and temporally adjacent as well as hierarchically related blocks. The third compression scheme, in contrast to the second scheme, use discrete wavelet transform (DWT) to decompose a video frame into a set of subframes with different resolutions corresponding to different frequency bands. The main idea of this scheme is that, for a particular subframe, the motion activities at different resolutions are different but highly correlated since they actually specify the same motion structure at different scales. These three schemes can be used separately or jointly depending on the applications and requirements.

In the vast majority of system which employ some form of image or video coding, the final recipient of the reconstructed image signal is the human eye. It is thus natural that attempts should be made to incorporate a measure of the properties of human visual system (HVS) in the coding system to ensure that coding bits are preferentially allocated to those components of output signal which will correspond to structure in the picture to which the eye is most sensitive. Using a small set of visual patterns which are localized subimages containing visually important information to code image has been proposed by researchers. The main merit of visual pattern image coding is its extremely low computational complexity while produce excellent image quality in accordance with perception. In this field, we have done some research on how to use visual patterns in video coding while maintain its advantages such as extremely low computational complexity, and on setting some principles for visual pattern design.

After a successful coding, many applications require the coding contents be delivered over a computer networks. The available bandwidth of most computer networks pose a problem when video is delivered. A user may request a video sequence be delivered at a specific data rate with a specific quality. However, the variety of requests and the diversity of the traffic on the network may make it difficulty for a video server to predict, at the time the video is encoded and stored on the server, the video quality and data rate it will be able to provide to a particular user at a given time. Therefore, scalability, the capability of decoding a compressed sequence at different rates, has become a very important issue in video coding.

We proposed a new wavelet based scalable video compression coder which can provide continuous rate scalability. The novel features of this scheme are: 1) a extended zero tree structure is used to organize wavelet coefficients in a more efficient way than the simple zerotree coding; 2) an arithmetic coder employing similarity between and within subbands is used to increase the efficiency of data compression; 3) an proposed embedded motion compensation scheme is employed to prevent error propagation between predictive frames, thus to maintain the quality of the decoded video even at very low bit-rate. We show that that the proposed coder not only provides a wide range of rate scalability, but also achieves comparable performance to the traditional hybrid coders.