HIERARCHICAL SCALABLE MOTION-COMPENSATED VIDEO CODING

By

Robert A. Cohen

An Abstract of a Thesis Submitted to the Graduate

Faculty of Rensselaer Polytechnic Institute

in Partial Fulfillment of the

Requirements for the Degree of

DOCTOR OF PHILOSOPHY

Major Subject: Electrical Engineering

The original of the complete thesis is on file
in the Rensselaer Polytechnic Institute Library

Examining Committee:

Prof. John W. Woods, Thesis Adviser
Prof. David Isaacson, Member
Prof. William A. Pearlman, Member
Prof. Gary J. Saulnier, Member

Rensselaer Polytechnic Institute
Troy, New York

January 2007
(For Graduation May 2007)

ABSTRACT

With the rapidly increasing availability of digital video sources and the wide range of bandwidths over which we would like to transmit the video, subband/wavelet-based video coding methods are rapidly becoming attractive alternatives to older well-established coding methods such as MPEG-2, or to the more recent block-based MPEG-4 and H.264 baseline standards. The embedded nature inherent in 3-D subband coding enables video to be scalable spatially, temporally, and with respect to quality. Recently, very good scalable results have been obtained from several motion compensated wavelet-based subband coding techniques. Also, there currently is much interest in using the JPEG 2000 image coding standard to perform high-quality intraframe video compression. Although JPEG 2000 provides spatial and quality (SNR) scalability, it ignores temporal redundancies. In this thesis, we look at new ways to exploit the spatial and resolution-scalable properties of JPEG 2000 when it is combined with 3-D motion compensated subband coding.

After providing an overview of related topics, we present results from our motion-compensated JPEG 2000 coder, which codes the temporal subbands output from a motion-compensated temporal filter (MCTF). On digital cinema sequences, this codec performed within .5 dB of MC-EZBC, which used EZBC to code the temporal subbands. We then describe aspects related to our resolution-hierarchical system, in which motion vectors and MCTF subbands generated in one resolution are used to initialize and predict data of the next resolution. Differential encoding of the prediction error is combined with scalable entropy coding to produce a resolution, spatial, and temporally scalable bitstream. We then take a closer look at how this hierarchy can be exploited for predicting and coding motion vectors. Replacing an adaptive arithmetic coder (AAC) with a context-adaptive binary arithmetic coder (CABAC) yields additional performance improvements.

We next introduce and show results for an open-loop hierarchical motion-compensated codec that eliminates the decoder-feedback loop found in the spatial predictor of our earlier encoder. We analytically and experimentally determine op-

timal rate allocations for these coders, and we discuss the trade-offs encountered between these systems, including how the closed-loop system works best with minimal low-resolution rates, and the open-loop system works better for higher low-resolution rates.