f=U.
00-HP

Proceedings

DCC 2000

DATA COMPRESSION CONFERENCE

March 28-30, 2000 Snowbird, Utah

Edited by

James A. Storer
Martin Cohn

Sponsored by

IEEE Computer Society
Technical Committee on Computer Communications

IEEE COMPUTER SOCIETY

Los Alamitos, California
Washington · Brussels · Tokyo

10

Poster Session

High performance Lempel-Ziv compression using optimized longest string parsing and adaptive Huffman window size

Naphtali Rishe, Artyom Shaposhnikov,
Alexander Vaschillo, Dmitry Vasilevsky, Shu-Ching Chen
High Performance Database Research Center[1], School of Computer Science,
Florida International University, University Park, Miami, FL 33199, (305)348-1706,
Fax (305)348-1705, E-mail: {rishen,shaposhn,avasch01,dvasil01,chens}@fiu.edu

Keywords: *data compression, Lempel-Ziv, Huffman, adaptive window, longest match string searching, asymmetric hashing.*

Abstract

We present optimizations that improve the compression and computational efficiency of Lempel-Ziv (LZ77) and Huffman algorithms [Shaposhnikov-93]. The compression performance of the LZ77 algorithm can be improved by using an optimized longest match parsing strategy. Another factor that can be considered is the size of the reference to the matching string, which can vary due to the secondary Huffman compression. We present an efficient parsing algorithm that considers both factors while minimizing the computational requirements. Our second optimization technique optimizes static Huffman encoding by efficiently dividing the input into blocks of varying size with uneven character frequency distributions.

Our parsing method finds the longest match and results in better compression than greedy parsing. The search is optimized using an asymmetric hash function, which works faster than similar methods using hash searching. The output of LZ77 is further compressed by a static Huffman compression with statistically adaptable block size. The algorithm improves Huffman compression by efficiently and accurately estimating the output size of static Huffman blocks.

An experimental program using the proposed optimizations was created and its performance was analyzed using the Calgary data compression corpus. The proposed optimizations result in one of the best performances achievable with LZ77 and Huffman algorithms. The program works approximately 28% faster than GZip and the compression gain is about 4% better. While this is a modest gain, considering that both programs use the same compression algorithms and dictionary size, the improvements are significant and can be utilized to improve many existing data compression systems without requiring the end users to update the decompression software and hardware.

References

[Shaposhnikov-93] A. Shaposhnikov "Data compression optimizations for the Lempel-Ziv-Huffman algorithm" TR-1993-10, FIU SCS, Miami FL, 33199.
Available at: http://www.cs.fiu.edu/~shaposhn/compress93.pdf

[1] This research was supported in part by NASA (under grants NAGW-4080, NAG5-5095, NAS5-97222, and NAG5-6830) and NSF (CDA-9711582, IRI-9409661, HRD-9707076, and ANI-9876409).

Wireless Image Transmission Using Multiple-Description Based Concatenated Codes*

Daniel Grobe Sachs, Univ. of Ill. at Urbana-Champaign, sachs@ifp.uiuc.edu
Raghavan Anand, Univ. of California-Berkeley, anand@eecs.berkeley.edu
Kannan Ramchandran, Univ. of California-Berkeley, kannanr@eecs.berkeley.edu

This work introduces a multiple-description product code which aims at optimally generating multiple, equally-important wavelet image descriptions from an image encoded by the popular SPIHT image coder. Because the SPIHT image coder is highly sensitive to errors, forward error correction is used to protect the image against bit errors occurring in the channel. The error-correction code is a concatenated channel code including a row (outer) code based on RCPC codes with CRC error detection and a source-channel column (inner) code consisting of the scalable SPIHT image coder and an optimized array of unequal protection Reed-Solomon erasure-correction codes. By matching the unequal protection codes to the embedded source bitstream using our simple, fast optimizer, we maximize expected image quality and provide for graceful degradation of the received image during fades.

To achieve unequal protection, each packet is split into many Reed-Solomon symbols. The i^{th} symbol in each packet forms an (n, k) Reed-Solomon code or "column". A fast, nearly-optimal optimizer, based on Lagrange multipliers and optimal to within convex hull and discretization approximations, chooses k for each Reed-Solomon "column" to minimize the expected mean-squared error at the receiver.

We validated our use of this structure by evaluating its performance in the context of transmitting images over a wireless fading channel. The performance of this scheme was evaluated by simulating the transmission of the Lena image over a Clarke flat-fading channel with an average SNR of 10 dB and a normalized Doppler frequency of 10^{-5} Hz. Our implementation of the Sherwood and Zeger "UEP2" unequal-protection code published at ICIP 1998 achieves a mean PSNR of 27.75 dB (28.36 dB peak) at a bit rate of 0.237 bpp. Using the same channel and packet-loss data, the RCPC/CRC+MDFEC scheme using the efficient Lagrange optimizer achieves an expected PSNR of 27.90 dB (28.84 dB peak) By substituting a rate 2/3 RCPC code for Sherwood and Zeger's rate 1/2 code, we can achieve an expected PSNR at the receiver of 28.38 dB (29.75 dB peak), a 0.63 dB improvement over the "UEP2" code.

In addition to its high performance compared to other techniques for sending images over wireless channels, this packetization scheme and optimizer is also ideally suited to hybrid packet-network and wireless channels. Because the optimization is based only on an end-to-end packet loss distribution, such a hybrid network can be evaluated easily without considering which exact packets arrived intact. And since all packets are equally important, no concept of packet priority is required.

The complete paper can be downloaded from `http://www.ifp.uiuc.edu/~sachs`.

* This work is partially supported by NASA under grant NASA NAG 3-2263.