For further information see: www.cs.brandeis.edu/~dcc

PROGRAM

Data Compression Conference (DCC'98)

(Sponsored by the IEEE Computer Society TCCC)

Snowbird, Utah
March 30 - April 1, 1998

COMMITTEE: J. Storer (*DCC Chair*, Brandeis U.), M. Cohn (*DCC Program Chair*, Brandeis U.), A. Apostolico (Purdue/Padova), R. Arps (IBM), T. Bell (U. Canterbury), B. Carpentieri (U. Salerno), J. Cleary (U. Waikato), M. Effros (CalTech), R. Gray (Stanford U.), S. Haggerty (Ball Aerospace), P. Howard (AT&T), A. Kiely (NASA), J. Kovacevic (Bell Labs), G. Langdon (UC Santa Cruz), A. Lempel (Technion), A. Moffat (U. Melbourne), M. Rabbani (Kodak), K. Ramchandran (U. Illinois), J. Reif (Duke U.), E. Riskin (U. Washington), A. Rodriguez (Scientific Alanta), D. Sheinwald (IBM), G. Sullivan (PictureTel), J. Villasenor (UCLA), J. Vitter (Duke U.), I. Witten (U. Waikato), H. Yokoo (Gunma U.), K. Zeger (UC San Diego), J. Ziv (Technion).

THEME: An international forum for current work on data compression and related areas. Topics of interest include but are not limited to: Source coding, quantization theory, parallel compression algorithms and hardware, lossless and lossy compression algorithms for specific types of data (including text, images, video, speech, music, maps, instrument data, graphics, animation, and bit-maps), data compression standards, bi-level coding, transform methods, wavelet and fractal techniques, string searching and manipulation, closest-match retrieval, minimal length encoding and applications to learning, system issues relating to data compression (including error control, data security, indexing, and browsing), medical imagery, scientific and space data.

SCHEDULE OVERVIEW:

Sunday Evening, March 29: Registration and Reception

Monday, March 30:

Morning: Technical Sessions
Mid-Day: Invited Presentation
Afternoon: Technical Sessions

Tuesday, March 31:

Morning: Technical Sessions
Mid-Day: Technical Sessions
Late Afternoon: Poster Session and Reception

Wednesday, April 1:

Morning: Technical Sessions
Mid-Day: Invited Presentation
Afternoon: Technical Sessions

Thursday, April 2: Industry Workshop

SUNDAY EVENING

Registration / Reception, 7:00-10:00pm (Golden Cliff Room)

MONDAY

Welcome: 7:45am

Session 1: 8:00am - 10:05am

8:00am
"Piecewise Integer Mapping for Arithmetic Coding"
Lang Stuiver and Alistair Moffat
University of Melbourne

8:25am
"The Z-Coder Adaptive Binary Coder"
Leon Bottou, Paul G. Howard and Yoshua Bengio
AT&T Labs, Universite de Montreal

8:50am
"Compression of Sparse Matrices by Arithmetic Coding"
Timothy Bell and Bruce McKenzie
University of Canterbury

9:15am
"Fast Convergence with a Greedy Tag-Phrase Dictionary"
Ross Peters and Tony C. Smith
University of Waikato

9:40am
"Tag Based Models of English Text"
W.J. Teahan and John G. Cleary
University of Waikato

Break: 10:05am - 10:30am

Session 2: 10:30am - 12:35pm

10:30am
"Practical Multi-Resolution Source Coding: TSVQ Revisited"
Michelle Effros
The California Institute of Technology

10:55am
"Quantization, Classification, and Density Estimation for Kohonen's Gaussian Mixture"
Robert M. Gray, Karen O. Perlmutter and Richard A. Olshen
Stanford University

11:20am
"On the Performance of Vector Quantizers Empirically Designed From Dependent Sources"
Assaf J. Zeevi
Stanford University

11:45am
"Successfully Refinable Trellis Coded Quantization"
Hamid Jafarkhani and Vahid Tarokh
AT&T Labs

12:10pm
"Embedded Trellis Coded Quantization"
Hugh Brunk and Nariman Farvadin
University of Maryland

Lunch Break: 12:35pm - 2:00pm

Mid-Day Invited Presentation: 2:00pm - 3:00pm

"Multiple Descriptions as Joint Source Channel Codes"
Dr. Jelena Kovacevic
Bell Laboratories

Break: 3:00pm - 3:30pm

Session 10: 3:30pm - 5:10pm

3:30pm
"Multiple Pattern Matching in LZW Compressed Text"
Takuya Kida, Masayuki Takeda, Ayumi Shinohara, Masamichi Miyazaki, and Setsuo Arikawa
Kyushu Univeristy

3:55pm
"Pattern Matching in Text Compressed with the ID Heuristic"
Piera Baracaccia, Antonella Cresti, and Sergio De Agostino
Universita di Roma

4:20pm
"Some Theory and Practice of Greedy Off-Line Textual Substitution"
Alberto Apostolico and Stefano Lonardi
Purdue Univeristy, Universita di Padova

4:45pm
"A Fast Algorithms for Making Suffix Arrays and for Burrows-Wheeler Transformation"
Kunihiko Sadakane
University of Tokyo

Break: 5:10pm - 5:35pm

Session 11: 5:35pm - 7:15pm

5:35pm
"Joint Source/Channel Coding for Variable Length Codes"
N. Demir, K. Sayood
University of Nebraska

6:00pm
"The Multiple Description Rate Region for High Resolution Source Coding"
Tamas Linder, Ram Zamir, and Kenneth Zeger
University of California, Tel-Aviv University, University of California

6:25pm
"Turbo Decoding of Hidden Markov Sources with Unknown Parameters"
Javier Garcia-Frias and John Villasenor
University of California

6:50pm
"Minimum Message Length Hidden Markov Modeling"
Timothy Edgoose, and Lloyd Allison
Monash University

TUESDAY

Session 5: 8:00 - 10:05

8:00am
"Phrase Hierarchy Inference and Compression in Bounded Space"
Craig G. Nevill-Manning and Ian H. Witten
Stanford University, University of Waikato

8:25am
"The Context Trees of Block Sorting Compression"
Jesper Larsson
Lund University

8:50am
"The Prevention of Error Propagation in Dictionary Compression with Update and Deletion"
James A. Storer
Brandeis University

9:15am
"Musical Image Compression"
David Bainbridge and Stuart Inglis
University of Waikato

9:40am
"On-Line Compression of High Precision Images by Evolvable Hardware"
M. Salami, H. Sakanashi, M. Iwata, and T. Higuchi
Hokkaido University, Electrotechnical Laboratory

Break: 10:05am - 10:30am

Session 6: 10:30am - 12:35Pm

10:30am
"Efficient Algorithms for Optimal Video Transmission"
Dexter Kozen, Yaron Minsky, and Brian Smith
Cornell University

10:55am
"A Locally Optimal Design Algorithm for Block-Based Multi-Hypothesis Motion-Compensated Prediction"
Markus Flierl, Thomas Wiegand and Bernd Girod
University of Erlangen, Stanford University

11:20am
"Intensity Controlled Motion Compensation"
Jarkko Kari, Mihai Gavrilescu
University of Iowa

11:45am
"The H.263+ Video Coding Standard: Complexity and Performance"
Berna Erol, Michael Gallant, Guy Cote and Faouzi Kossentini
University of British Columbia

12:10pm
"Codeword Asigment for Fixed-Length Entropy Coded Video Streams"
Ramon Llados-Bernaus and Robert L. Stevenson
University of Notre Dame

Lunch Break: 12:35pm - 2:00pm

Session 7: 2:00pm - 4:30Pm

2:00
"Non-Uniform PPM and Context Tree Models"
J. Aberg, Yu Shtarkov, B. J. Smeets
Lund University, Russian Academy of Sciences, Lund University

2:25
"Correcting English Text Using PPM models"
W.J. Teahan, S. Inglis, J.G. Cleary, and G. Holmes
University of Waikato

2:50
"Context Tables: A Tool for Describing Text Compression Algorithms"
Hidetoshi Yokoo
Gunma University

3:15
"Context Models for Palette Images"
Paul J. Ausbeck Jr.
University of California

3:40
"Compression by Model Combination"
Tong Zhang
Stanford University

4:05
"Bayesian State Combining for Context Models"
Suzanne Bunton
University of Washington

POSTER SESSION AND RECEPTION

4:45-7:30pm
In the Golden Cliff Room

(Abstracts of each presentation appear in the proceedings.)

WEDNESDAY

Session 8: 8:00am - 10:05am

8:00am
"Image Transmission Using Arithmetic Coding Based Continuous Error Detection"
Igor Kozintsev, Jim Chou and Kannan Ramchandran
University of Illinois

8:25am
"Transmission Error Robust Fractal Coding Using a Model-Residual Approach"
Mirek Novak
Lund University

8:50am
"A Lossless 2-D Image Compression Technique for Synthetic Discrete-Tone Images"
Jeffrey M. Gilbert and Robert W. Brodersen
University of California

9:15am
"Linear Time Construction of Optimal Context Trees"
Harald Helfgott and Martin Cohn
Brandeis Universtiy

9:40am
"Lossless Interframe Image Compression via Context Modeling"
Xiaolin Wu, Wai-Kin Choi and Nasir Memon
University of Western Ontario, Chinese University of Hong Kong, Northern Illinois University

Break: 10:05am - 10:30am

Session 9: 10:30am - 12:35pm

10:30am
"Optimal Multiple-Description Transform Coding of Gaussian Vectors"
Vivek K. Goyal and Jelena Kovacevic
University of California, Bell Laboratories

10:55am
"Line Based, Reduced Memory, Wavelet Image Compression"
Christos Chrysafis and Antonio Ortega
University of Southern California

11:20am
"A Low-Complexity Modeling Approach for Embedded Coding of Wavelet Coefficient"
Erik Ordentlich, Marcelo Weinberger, Gadiel Seroussi
Hewlett-Packard Laboratories

11:45am
"Robust Wavelet Zerotree Image Compression with Fixed-Length Packetization"
Jon Rogers and Pamela Cosman
University of California

12:10pm
"Efficient Lossless Coding of Medical Image Volumes Using Reversible Integer wavelet Transforms"
Ali Bilgin, George Zweig, and Michael Marcellin
University of Arizona

Lunch Break: 12:35pm - 2:00pm

Mid-Day Invited Presentation: 2:00pm - 3:00pm

"Computational Costs of Surprise"
Prof. Alberto Apolostico
Purdue University / University of Padova

Break: 3:00pm - 3:30pm

Session 10: 3:30pm - 5:10pm

3:30pm
"On Suboptimal Multidimensional Companding"
Stephan F. Simon
Aachen University of Technology (RWTH)

3:55pm
"Accurate Subband Coding with Low Resolution Quantization"
Zoran Cvetkovic
AT&T Labs

4:20pm
"Extending TMW for Near Lossless Compression of Greyscale Images"
Bernd Meyer and Peter Tischer
Monash University

4:45pm
"Reversible Variable Length Codes for Efficient and Robust Image and Video Coding"
Jiangtao Wen and John Villasenor
University of California

Break: 5:10pm - 5:35pm

Session 11: 5:35pm - 7:15pm

5:35pm
"Variable-to-Fixed Length Codes for Predictable Sources"
Serap Savani
Lucent Technologies

6:00pm
"Switching Between Two Universal Source Coding Algorithms"
Paul Volf and Frans Willems
Eindhoven University of Technology

6:25pm
"Optimal Lossless Compression of a Class of Dynamic Sources"
John H. Reif and James A. Storer
Duke University, Brandeis University

6:50pm
"Universal Data Compression and Linear Prediction"
Meir Feder and Andrew Singer
Tel-Aviv University, Advanced Systems Directorate

Mid-Day Presentations

Monday:

"Multiple Descriptions as Joint Source Channel Codes"
Dr. Jelena Kovacevic
Bell Laboratories

A large fraction of the information that flows across today's networks is useful even in a degraded condition. Examples include still images, speech, audio, and video. When this information is subject to packet losses or retransmission is impossible due to real-time constraints, superior performance with respect to total transmitted rate, distortion, and delay may sometimes be achieved by adding redundancy to the bit stream rather than repeating lost packets. In multiple description coding, a joint source-channel method, the data is broken into several streams with some redundancy among the streams. When all the streams are received, one can guarantee low distortion at the expense of having a slightly higher bit rate than a system designed purely for compression. On the other hand, when only some of the streams are received, the quality of the reconstruction degrades gracefully, which is very unlikely to happen with a system designed purely for compression. We will motivate the use of multiple descriptions and review some of its information theoretical beginnings. We will then discuss two practical approaches: one quantizer based and the other transform based. Finally, we will outline possible applications.

Wednesday:

"Computational Costs of Surprise"
Prof. Alberto Apolostico
Purdue University / University of Padova

The problem of characterizing and detecting over- or under-represented words in sequences arises ubiquitously in diverse applications and has been studied rather extensively in Computational Molecular Biology. In most approaches to the detection of unusual frequencies of words in sequences, the words (up to a certain length) are enumerated more or less exhaustively and individually checked in terms of observed and expected frequencies, variances, and scores of discrepancy and significance thereof. We take instead the global approach of annotating a suffix trie or automaton of a sequence with some such values and scores, with the objective of using it as a collective detector of all unexpected behaviors, or perhaps just as a preliminary filter for words suspicious enough to warrant further and more accurate scrutiny. We show that such annotations can be carried out in a time- and space efficient fashion for the mean, variance and some of the adopted measures of significance, even without a-priori limits on the length of the words considered. Specifically, we concentrate on the simple probabilistic model in which sequences are produced by a random source emitting symbols from a known alphabet independently and according to a given distribution. We discuss data structures and tools for computing and storing the expected value and variance of all substrings of a given sequence of n symbols in (optimal) O(n^2) overall worst-case, O(nlog n) expected time and space. The O(n^2) time bound constitutes an improvement by a linear factor over the direct method. WE SHOW THAT UNDER SEVERAL ACCEPTED MEASURES OF DEVIATION FROM EXPECTED FREQUENCY, THE CANDIDATE OVER AND UNDER-REPRESENTED WORDS ARE O(n) ONLY. This surprising fact is a consequence of combinatorial properties that constrain the distributions of word occurrences in a string. Based on this, we design global detectors of favored and unfavored words for our probabilistic framework, and display the results of some preliminary experiments.

Work carried out jointly with M.E. Bock, S. Lonardi and X. Xu