An Optimizing Lossy Generalization of LZW

Steven Pigeon
Université de Montréal
pigeon@iro.umontreal.ca

Abstract

We present a lossy generalization of Welch's LZW algorithm [Welch84]. We first discuss two lossy generalizations of LZW,G-LZW [Pigeon95] and LLZW [Chiang98] that do not optimize any specific objective function. We present an algorithm, also lossy, that optimizes compression according to an objective function supplied by the user. This enables him to balance between compression ratio and image quality.

The algorithm in [Pigeon95] uses a trie as its data structure. Instead of looking for an exact match in the trie with the current input, algorithm G-LZW greedily matches with the pixel in the trie having the less difference with the corresponding pixel in the input. This greedy strategy maximizes image quality but prevents the algorithm from finding longer matches because of a locally bad match. The algorithm in [Chiang98] finds the longuest string in the dictionary that does not exceed a certain error threshold. This only maximize compression.

The proprosed algorithm searches its dictionary to find a string that maximizes (or minimizes) some error function that takes into account both compression ratio and image quality. This algorithm, P-LLZW, optimizes in linear time over the trie a separable objective function (parameterized by ϕ, as in the table, the maximum instantaneous error). Our implementation of the algorithm as a GIF optimizer has the particularity that it is decode compatible with GIF readers, meaning that the program creates files that can be read by web browsers and file viewers without having to recourse to a plug-in, even if the compressor takes different decisions during compression. The table shows results for files from the DJVU corpus.

				P-LLZW			
8 BPP File	Raw Size	Gif Size K	Gif BPP	ϕ	Size in K	BPP	SNR
NY Cargo	1.1M	766.8	5.7	20	275	2.0	27 dB
Jupiter	7.9M	4542	4.6	20	3436	3.4	25 dB
Lagaffe	8.5M	5064	4.8	20	3495	2.4	26 dB
New River	16.7M	8402	4.1	20	6164	3.0	29 dB
P-Boxes	7.4M	3410	3.7	20	2343	2.5	29 dB
Whitman	2.5M	1518	4.8	20	1007	3.2	26 dB

[Chiang98] Loben Chiang, *untranslated thesis title*, City University of Hong Kong, August 1998

[Pigeon95]pigeon95 Steven Pigeon, FLATLAND, *ou comment réduire une image GIF en modifiant l'algorithme LZW*, Journal l'Interactif, March 1996

[Welch84] Welch, Terry A.,*A technique for high-performance data compression*, Computer, June 1984, pp. 8-19

Proceedings of the Data Compression Conference (DCC'01)
1068-0314/01 $10.00 © 2001 IEEE