

要 旨

LZW 法にスライド窓を組み合わせた データ圧縮手法の改良に関する研究

高月 康裕

近年，計算機の性能向上に伴い，扱うファイルの大きさが増大する傾向にある．そのため，限られた記憶領域をより効率的に使う必要がある．それを実現するための方法の一つに，データ圧縮技術がある．Ziv と Lempel の提案したインクリメンタル分解を実現した圧縮アルゴリズムの一つである LZW 法は，ユニバーサル性，高速，高圧縮の実用的な無ひずみ圧縮アルゴリズムとして広く利用されている．しかし，出現回数が多くない文字列を処理する場合，それが長ければ長いほど，直ぐには効率の良い圧縮ができないという欠点を持っており，それを解決する手法として，LZW 法にスライド窓を組み合わせたデータ圧縮方式が提案されている．この手法では，LZW 法に付加したスライド窓の中で，一致した文字列の長さを LZW 法の符号語の代わりに符号化することで，圧縮能力の向上を行っている．

本研究で提案するデータ圧縮手法では，LZW 法にスライド窓を組み合わせた圧縮法で処理を行う際，スライド窓を用いて符号化を行うときに発見した文字列を，新しく LZW 法の辞書に登録する．LZW 法には，今後使われる可能性の高いフレーズが多く登録されているほど圧縮能力が高くなるという特性があり，これを利用することを考えたためである．この手法を実装し，実験を行った結果，実用的に利用出来る程には圧縮能力が上がらない事が確認できた．本稿では，その原因についても検討を加える．

キーワード データ圧縮, LZW 法, スライド窓

Abstract

Improvement of Compression Method Using LZW Coding and a Slide Window

Yasuhiro TAKATSUKI

In recent years, with a performance improvement of a computer, the size of files to be managed tends to increase. Therefore it is necessary to use storage capacity efficiently. Data compression technique is one of the methods to realize it.

LZW coding is one of variations of compression method by Ziv and Lempel (LZ78). It is a universal lossless compress algorithm and achieves high speed processing and practical compression ratio. However, it cannot compress efficiently strings that appear only few times. So an improved method that combines LZW coding with a sliding window has been proposed. This method encodes the length of coinciding string in the sliding window instead of the LZW code.

This paper shows a improvement to the existing compression method. In proposed method, when a long coinciding string is found in the sliding window, it is added into the LZW dictionary as a new phrase. This improvement uses the characteristics of LZW coding, that is, compression ability increases as the number of phrases in the dictionary increases.

As the result, proposed method can not achieve good enough compression ratio. This paper also gives discussion on the result and the cause.

key words Data Compression , LZW Coding , Sliding Window