

Universal Data Compression

July, 2008

Abstract

File compression is a technology that reduces the size of data by 1/5 - 1/2 that of the original, yet it preserves all information, and can completely restore the original data without even a 1-bit defect. This is called universal data compression if the method is applicable to all types of data. The most popular software for this is “UNIX compress”, “LHA” and “GZIP”. However, since several patent issues regarding compression software have recently arisen, for users to choose software with a guaranteed patent has become important. Fujitsu Laboratories has developed an original compression software package called SLCA (Super Lossless Compression & Archiver). SLCA offers high compression and has an airtight patent.

Technology

Since the patent issues all revolve around the dictionary compression method that effectively compresses long repetitions of sentences, SLCA uses a completely different dictionary compression method that has been devised so that it does not infringe on any prior patents. Together with the dictionary compression, our method uses a statistical compression that compresses short repetitions of words to achieve high performance.

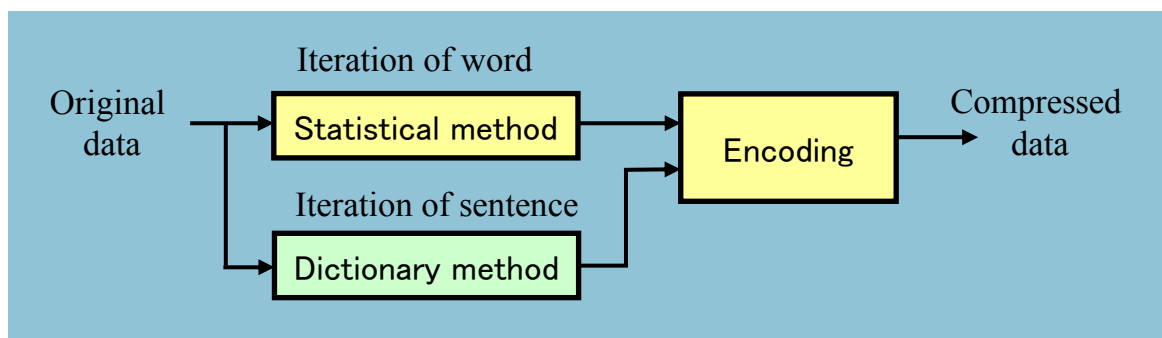
Features

- An original algorithm has been adopted and can be used for business safely.
- The high speed of both compression and decompression makes possible real-time processing.
- It has an archiver function that saves multiple files into a single file.
- For system development, it provides three types of interfaces.
 - API (Application Programming Interface)
 - CUI (Character User Interface)
 - GUI (Graphical User Interface)

Application Examples

SLCA is already broadly used in software for the various hardware products, and is working stably. It is incorporated and used in the following Fujitsu’s software products.

- For increasing efficiency of data transfer → Data transfer component “Linkexpress”
- For increasing efficiency of data storage → Search software “IntelligentSearch”, Groupware “TeamWARE OFFICE”



SLCA algorithm: Statistical and dictionary methods are used together.