Huffman Zip Program Development Documentation

1. Requirements

This is a simple Java program to compress files by using Huffman Coding. So you are supposed to have a copy of the JRE (Java Runtime Environment) on your system to run Java applications. And your computer must have enough memory to run the program.

You can view the source of my program on: https://github.com/hangxu/HuffmanZip

2. Algorithm

The deflation algorithm used by hzip is an implementation of Huffman Coding. Huffman coding is an entropy encoding algorithm used for lossless data compression.

And I used the open source package from the Apache organization, org.apache.commons.compress.archivers.tar, to pack files to a single .tar file.

3. Hzip file format

A hzip file is a sequence of compressed members. Each member has the following structure (Version 1.0):

2 bytes	magic header	0x58 0x48 (char 'X' and 'H')
1 byte	compression version	now is 0x01
1 byte	flags	(reserved)
//4 bytes	file modification time	in Unix format (Not yet implanted)
? bytes	compressed data header	
4 bytes	uncompressed input size	
? bytes	compressed data block	
//4 bytes	CRC-32	(Not yet implanted)

4. Codes and APIs

All code belongs to the package net.harryxu.hzip and its subpackage, the package structure is shown as below:

net.harryxu.hzip

Main.java	The main program entry and codes for GUI
Encode.java	Some methods about Huffman encoding
Decode.java	Some methods about Huffman decoding
Driver.java	Maintain constant and coding method entry

Package net.harryxu.hzip.algorithm

HuffmanTreeNode.java	To get instance of nodes in Huffman Tree
MinimumHeap.java	A heap with all the element sorted ascending

Package net.harryxu.hzip.util

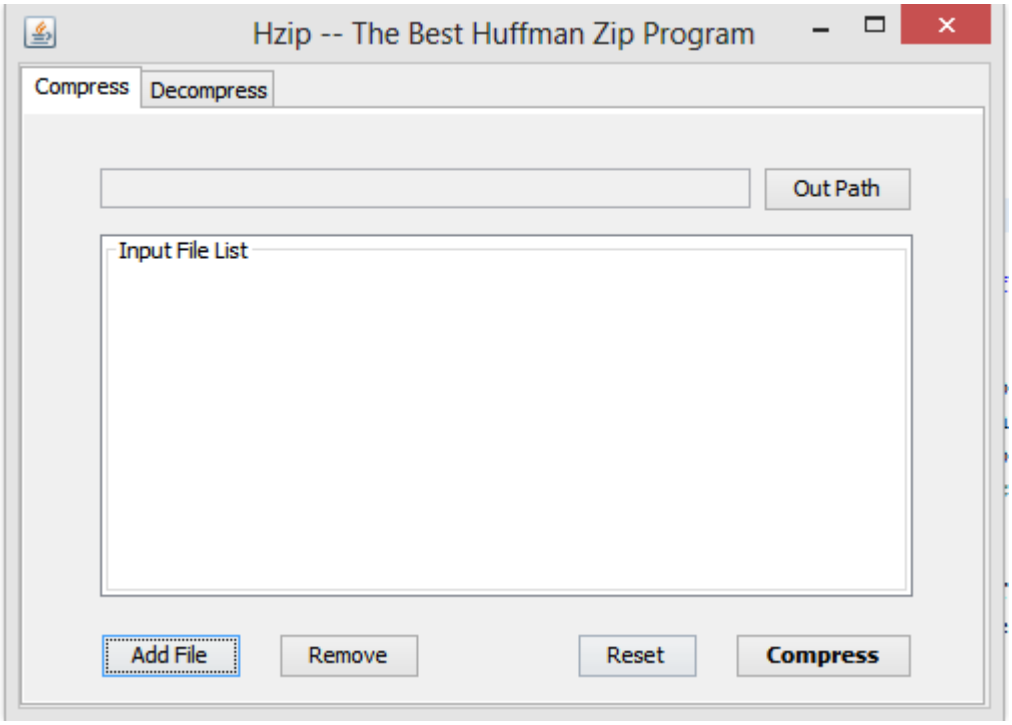
BitReader.java

BitWriter.java

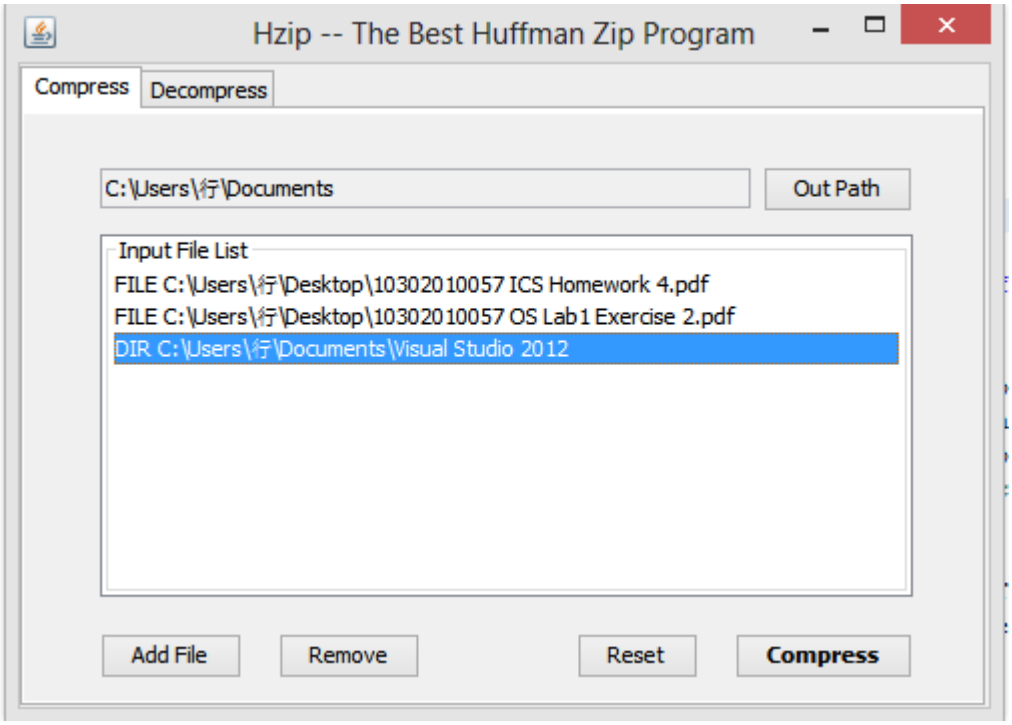
TarUtil.java

5. GUI

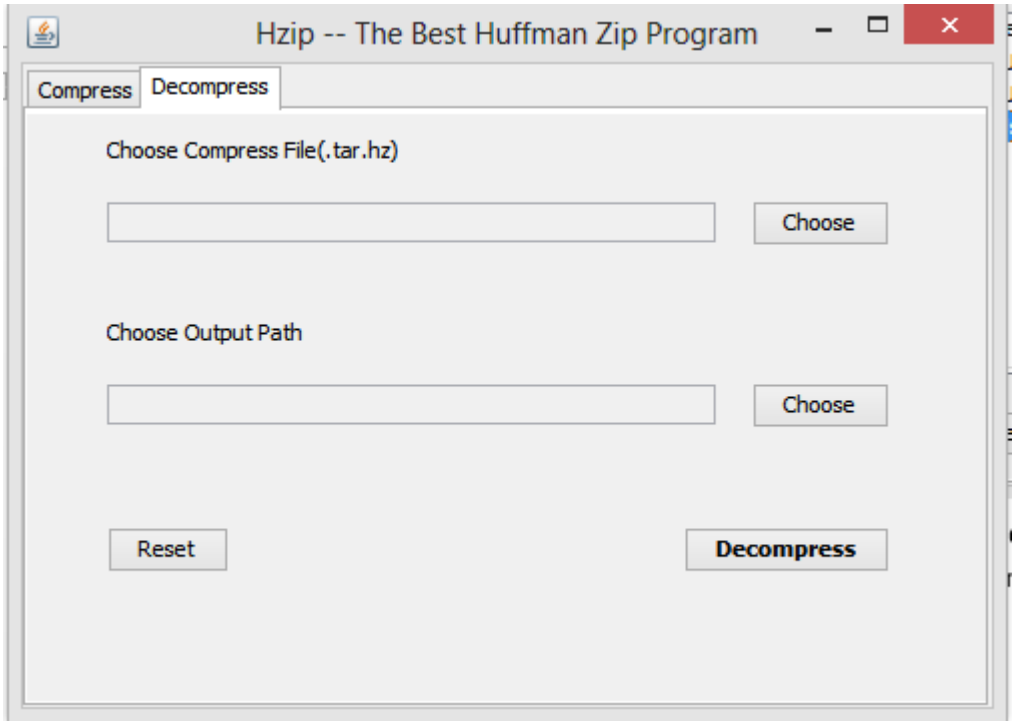
I “dragged” a simple GUI in the swing tools of Eclipse:



The user interface of compress tag. Click ADD FILE to add files into compress list. And click OUT PATH to select the base path of the output file



You can click REMOVE to remove the selected item in the compress list. You can click RESET button to reset all the component. And finally click COMPRESS to compress these selected files.



The user interface of decompress tag