Mikołaj Bujok

Szymon Miękus

Grupa BSI-2

Sprawozdanie z zajęć laboratoryjnych „Teoria informacji i kodowanie”

Kodowanie Huffmana i kompresja LZW

1. Temat sprawozdania

Sprawozdanie zbiorcze z zajęć laboratoryjnych, których tematem było zaimplementowanie programu kompresującego używającego metodę kodowania Huffmana oraz z zajęć, na których pracowaliśmy nad programem do kompresji LZW.

1. Przebieg i rezultaty

Zaimplementowaliśmy dodatkowe funkcje w naszej aplikacji okienkowej z poprzednich zajęć laboratoryjnych. Bazując na dostępnych oknach: wynikowym i wpisywania tekstu, funkcji wczytywania i zapisywania do pliku program korzysta z metody Huffmana do bezstratnej kompresji oraz LZW do słownikowej kompresji na osobnych kartach okna.

Rysunek 1: Zrzut ekranu aplikacji

Do porównania metod kompresji wykorzystano 3 pliki tekstowe: mały, średni i duży (o wielkościach ok. 5kB, 10kB, 20 kB).

* 1. Metoda Huffmana

W poniższej tabeli zostały przedstawione wielkości plików (w bajtach) bez kompresji oraz z zastosowaniem metody kodowania Huffmana.

|  |  |  |  |
| --- | --- | --- | --- |
|  | **mały** | **średni** | **duży** |
| Bez kodowania |  |  |  |
| Z kodowaniem |  |  |  |

* 1. Metoda LZW

|  |  |  |  |
| --- | --- | --- | --- |
|  | **mały** | **średni** | **duży** |
| Bez kompresji |  |  |  |
| Z kompresją |  |  |  |

1. Wnioski

Dla plików o bardzo małych rozmiarach kompresja jest niezalecana, ponieważ narzut dodatkowych danych związanych z kompresją przewyższa wielkość pliku przed kompresją. W przypadku plików o dużych rozmiarach sytuacja okazała się być nawet lepsza od idealnej, to mogło być spowodowane faktem użycia słów wygenerowanych przez „Lorem ipsum”, które powtarzały się cyklicznie w kompresowanym pliku tekstowym. W normalnym pliku przy jego większych rozmiarach narzutu danych związanych z kompresją jest znikomy, a więc kompresja jest opłacalna. Nie ma występuje również sytuacja lepsza od idealnego rozwiązania.