

Informacija in kodi

Poročilo tretje laboratorijske vaje

Avtor: Bor Starčič

Mentorja: izr.prof.dr.Simon Dobrišek, as.dr.Klemen Grm

Datum: 5. maj 2024

Kazalo

1	Uvod	3
2	Naloga 1: Implementacija LZW dekodeja	3
2.1	Preizkus delovanja	3
2.2	Preverjanje z MD5 izvlečki	3
3	Naloga 2: Izračun velikosti kompresiranega sporočila	3
3.1	Preizkus uspešnosti kompresije	4
3.2	Ovrednotenje razmerja med dolžino sporočila in uspešnostjo kompresije	5

1 Uvod

To poročilo opisuje implementacijo in testiranje kompresorja in dekompresorja LZW. Glavni cilj je bil raziskati učinkovitost tehnike stiskanja LZW na različnih vrstah datotek.

2 Naloga 1: Implementacija LZW dekodeerja

Prva naloga je zahtevala razvoj dekodeerja, ki deluje na osnovi prej izdelanega enkoderja LZW. Dekoder dinamično obnavlja slovar v procesu dekompresije, ohranjajoč sinhronizacijo z enkoderjem, kar omogoča natančno rekonstrukcijo izvirnega besedila.

2.1 Preizkus delovanja

Dekoder smo preizkusili na več datotekah, ki smo jih predhodno kompresirali z enkoderjem. Preizkus je vključeval primerjavo dekodiranega besedila z originalnim, s čimer smo potrdili pravilnost rekonstrukcije prikazano na sliki 1.

2.2 Preverjanje z MD5 izvlečki

Za potrditev integritete podatkov smo uporabili MD5 izvlečke prvotnih in dekomprimiranih datotek. Ujemanje MD5 kontrolnih vsot je potrdilo, da dekodeer pravilno obnovi prvotno datoteko brez izgub.

3 Naloga 2: Izračun velikosti kompresiranega sporočila

Druga naloga je vključevala implementacijo funkcionalnosti za izračun velikosti kompresiranega sporočila, kar nam omogoča analizo učinkovitosti kompresije v odvisnosti od velikosti izvirne datoteke.

3.1 Preizkus uspešnosti kompresije

Uspešnost kompresije smo testirali na različnih vrstah datotek, pri čemer so se pokazale različne stopnje učinkovitosti:

- **Tekstovne datoteke:** Visoka učinkovitost kompresije z razmerji 2.45 in 2.02 za `besedilo.txt` in `besediloS.txt`, kjer kaže na to da je pri večjih datotekah višja uspešnost kompresije.
- **Slikovne in video datoteke:** Kompresija ni bila učinkovita, ker je `slika.jpg` imela razmerje kompresije 0.84 `video.mp4` pa 0.87. To kaže, da je LZW algoritem povečal velikost datotek, kar je značilno za datoteke, ki so že stisnjene z drugimi metodami.
- **Python skripte:** Za `vaja3_predloga.py` smo zabeležili kompresijsko razmerje 1.76, kar kaže na zmerno učinkovitost LZW pri manjših datotekah.
- **Excel datoteke:** Odlično razmerje kompresije 3.72 za `excel.xls` kaže na visoko učinkovitost pri datotekah, ki niso predhodno stisnjene.

```

Datoteka: besedilo.txt
Prvotna velikost: 1038950 bajtov, Velikost po kompresiji: 423780 bajtov
Razmerje kompresije: 2.45
Prvotni MD5: {'727c3bd61031d70c31549a0ca27eb876'}
Dekompresiran MD5: {'727c3bd61031d70c31549a0ca27eb876'}
Ujemanje MD5: True
-----
Datoteka: besediloS.txt
Prvotna velikost: 11128 bajtov, Velikost po kompresiji: 5516 bajtov
Razmerje kompresije: 2.02
Prvotni MD5: {'d7ae90b54210c4679fa68f6d35e4c30f'}
Dekompresiran MD5: {'d7ae90b54210c4679fa68f6d35e4c30f'}
Ujemanje MD5: True
-----
Datoteka: slika.jpg
Prvotna velikost: 790753 bajtov, Velikost po kompresiji: 937264 bajtov
Razmerje kompresije: 0.84
Prvotni MD5: {'c9a23f5e66d37f29460ca517a2545dc2'}
Dekompresiran MD5: {'c9a23f5e66d37f29460ca517a2545dc2'}
Ujemanje MD5: True
-----
Datoteka: vaja3_predloga.py
Prvotna velikost: 1363 bajtov, Velikost po kompresiji: 774 bajtov
Razmerje kompresije: 1.76
Prvotni MD5: {'5f5bccb812dd15cda120306d3c78bc0a'}
Dekompresiran MD5: {'5f5bccb812dd15cda120306d3c78bc0a'}
Ujemanje MD5: True
-----
Datoteka: excel.xls
Prvotna velikost: 35328 bajtov, Velikost po kompresiji: 9490 bajtov
Razmerje kompresije: 3.72
Prvotni MD5: {'e374d1c03db9dff21807ed2d6ceb2583'}
Dekompresiran MD5: {'e374d1c03db9dff21807ed2d6ceb2583'}
Ujemanje MD5: True
-----
Datoteka: video.mp4
Prvotna velikost: 126886145 bajtov, Velikost po kompresiji: 146044517 bajtov
Razmerje kompresije: 0.87
Prvotni MD5: {'97d0e7388deac6afa9446a161d826cce'}
Dekompresiran MD5: {'97d0e7388deac6afa9446a161d826cce'}
Ujemanje MD5: True

```

Slika 1: Rezultati kompresije različnih vrst datotek

3.2 Ovrednotenje razmerja med dolžino sporočila in uspešnostjo kompresije

Razmerja med originalno in kompresirano velikostjo so nam omogočila oceniti učinkovitost LZW algoritma glede na vrsto datoteke. Rezultati kažejo, da je algoritem najbolj učinkovit pri tekstovnih in nekaterih datotekah z manjšim predhodnim stiskanjem (kot so Excel datoteke) in da uspešnost kompresije raste z večanjem datotek. Pri slikah in videoposnetkih se pa velikost datotek poveča, kar kaže na to da LZW ni primeren za stiskanje že stisnjenih medijskih formatov.