



NACIONALINĖ
ŠVIETIMO
AGENTŪRA

Ugdymo *meistrai*



DUOMENŲ GLAUDINIMAS

Vaida Masiulionytė-Dagienė, Vilniaus universitetas,
Vilniaus Valdorfo mokykla

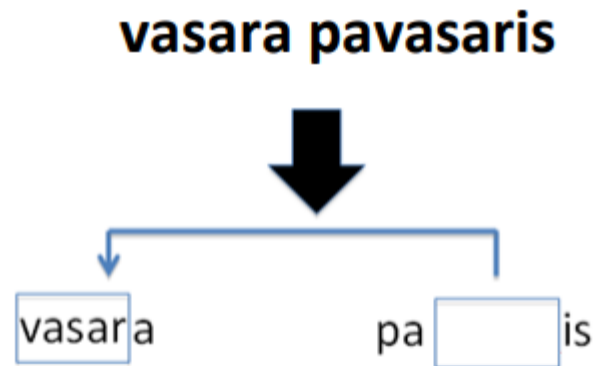
Duomenų glaudinimas (archyvavimas)

- Duomenų pertvarkymas taip, kad jie užimtų mažiau vietos atmintyje ir jiems persiūsti reikėtų siauresnės dažnių juostos arba persiuntimas būtų spartesnis.
- Duomenyse būna vienodų arba panašių fragmentų ir kitokio informacijos pertekliaus. Glaudinant pavyksta rasti dalį informacijos pertekliaus ir jį sumažinti ekonomiškiau koduojant duomenis. Dėl to sutaupoma jiems reikalingos atminties.
- Glaudinami pakuojami, archyvuojami ir elektroniniu paštu persiunčiami duomenys.

Pavyzdys, kaip gali būti glaudinamas tekstas

- Pasikartojantys žodžiai, žodžių dalys arba frazės, pakeičiami kvadratėliais

(<https://bebras.lt/wp-content/uploads/2016/09/3-veikla.pdf>)



Kaip atrodytų nesuglaudintas eilėraštis?

Čia mes atėjom su mama.
 Taip gera _____ pabūti,
 _____ tokios mėlynos žibutės.
 _____ pum _____ r _____ iai k _____ p ak _____
 Ir s _____ n _____ ių šil _____ a...

_____ ,
 _____ .

Pabandykite suglaudinti šias eilutes

Saulė

Vasara pavasaris

Vasara pavasaris

Saulė šypsosi

Vasara pavasaris

Vasara pavasaris

Į laukelį kviečia

Kaip kompiuteryje vaizduojamas glaudinimas

- Kompiuteryje stačiakampiai ir rodyklės vaizduojami skaičiais. Pavyzdžiui, Pasas gali būti užrašytas kaip Pas(2,2).
- Pirmasis skaičius 2 rodo antrą vietą nuo galo – kopijavimo vietos pradžią Pas– –
- Antrasis skaičius 2 reiškia, kad kopijuosime dvi iš eilės einančias raides. Pasa– Pasas
- Kadangi žodžiui užkoduoti naudojami du skaičiai, tai verta suspausti dviejų ar daugiau raidžių grupes, kitaip nebus sutaupoma vietos. Iš tiesų failo dydis gali padidėti, jeigu dviem skaičiais koduosime vieną raidę.

Duomenų glaudinimo metodai

- Sugalvota daug duomenų suspaudimo metodų. Atliekant šią veiklą, buvo naudojamas metodas, kai nuorodomis rodoma į prieš tai buvusius pasikartojančius simbolius tekste.
- Šis metodas vadinamas tiesiog „zip“, jį sugalvojo Phil Katz. Metodas tinka bet kuriai kalbai, juo lengvai suspaudžiami ir perpus sumažinami duomenys. (<https://bebras.lt/wp-content/uploads/2016/09/3-veikla.pdf>)
- Zip yra glaudinimo be praradimų formatas: po išskleidimo duomenys yra tokie patys, kokie buvo prieš glaudinimą. Jis veikia rasdamas perteklinius pradinius duomenis ir efektyviau juos atvaizduodamas. (<https://www.hanshq.net/zip.html>)

Interaktyvi glaudinimo priemonė

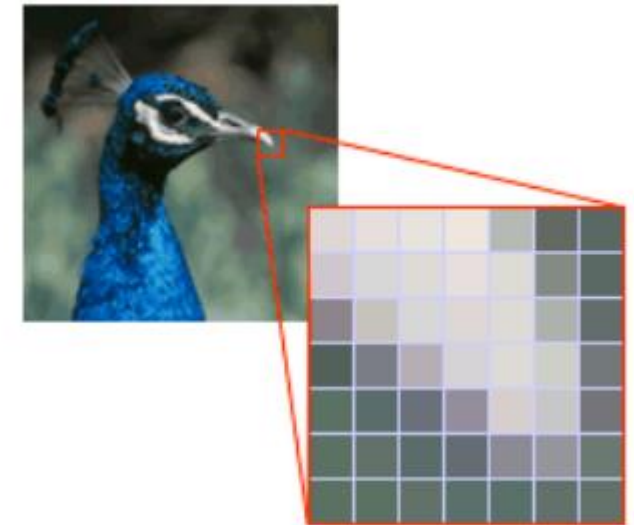
- <https://studio.code.org/s/text-compression/lessons/1/levels/2>
- Teksto pavyzdys:
 - Apvali, saldi gražuolė...
 - Šmurkšt, po suolu jau nupuolė.
 - „Kurgi bėgi? Grižk į puodą...“
 - Bet, pakėlus savo kuodą,
 - Bulvė lekia per laukus:
 - „Nebegrįšiu pas vaikus.“

Praktinė užduotis

- Išbandykite interaktyvią glaudinimo priemonę su savo tekstu ir paieškokite
- kokie yra dar duomenų glaudinimo metodai, kai neprarandami pradiniai duomenys?

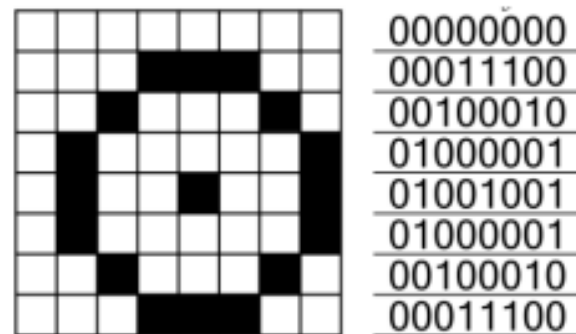
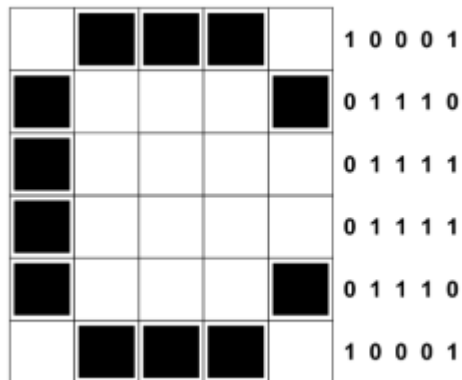
Vaizdų kodavimas ir glaudinimas

- Nuotraukos ir kiti vaizdai ekrane yra pateikiami kaip spalvotų taškų (pikselių) rinkinys.
- Kompiuterių ekranai suskirstyti į tinklelį mažų kvadratėlių (pikselių, taškų). Kiekvienas iš pikselių gali būti skirtingos spalvos, o kadangi pikseliai maži, atskirų taškų ekrane nematome. Jie susilieja ir formuoja rodomą vaizdą.
- (<https://atviri.emokymai.vu.lt/course/view.php?id=220>)



Vaizdų kodavimas

- Paprasčiausią nespalvotą vaizdą galime užkoduoti dvejetainine sistema (<https://atviri.emokymai.vu.lt/course/view.php?id=220>):



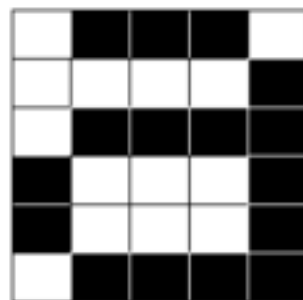
Vaizdų glaudimas

- Ar galima aukščiau užkoduotus vaizdus suglaudinti?

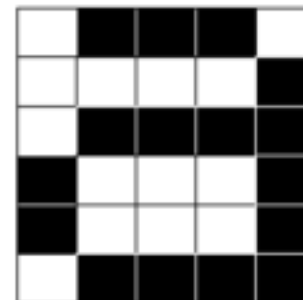
0	1	1	1	0
0	0	0	0	1
0	1	1	1	1
1	0	0	0	1
1	0	0	0	1
0	1	1	1	1

- Atkreipkite dėmesį, kad pirmoje eilutėje šalia yra 3 juodi pikseliai, o antroje – 4 balti pikseliai. Taigi užuot užrašę kiekvieną bitą, galime skaičiais apibūdinti raštą.

<https://atviri.emokymai.vu.lt/course/view.php?id=220>



1, 3, 1



1, 3, 1

4, 1

1, 4

0, 1, 3, 1

0, 1, 3, 1

1, 4

Praktinė užduotis

- Koks gausis paveikslas?

4, 1, 5

3, 3, 4

2, 1, 3, 1, 3

1, 2, 1, 1, 1, 2, 2

0, 1, 7, 1, 1

0, 1, 1, 2, 1, 2, 1, 1, 1

0, 1, 1, 2, 1, 2, 1, 1, 1

0, 1, 4, 2, 1, 1, 1

0, 9, 1

4, 6

Interaktyvi vaizdų kodavimo/glaudavimo priemonė

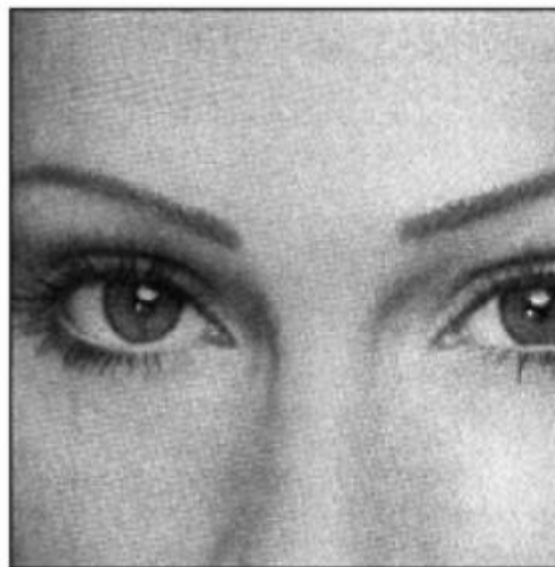
- <https://www.csfieldguide.org.nz/en/interactives/run-length-encoding/>

Glaudzinimas prarandant duomenis

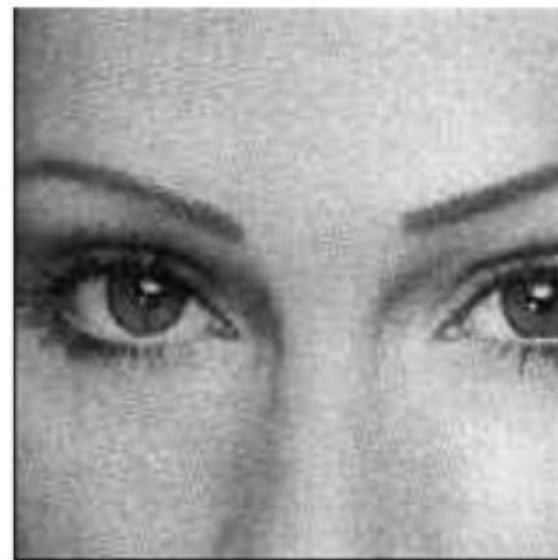
- Glaudzinimas su nuostoliais - tai skaitmeninių duomenų saugojimo ir perdavimo srityje naudojamas duomenų glaudzinimo metodas, kuriuo sumažinamas failo dydis, pašalinant dalį duomenų.
- Skirtingai nuo glaudzinimo be nuostolių, kai išskleidus išsaugomi visi pradiniai duomenys, glaudinant su nuostoliais aukojama dalis duomenų, kad būtų pasiektas didesnis glaudzinimo koeficientas.
- Tai reiškia, kad išskleidžiant failą, kuriam buvo atliktas nuostolingasis glaudzinimas, negaunama tiksli originalių duomenų kopija; vietoj to gaunama apytikslė kopija, kurios kokybė gali šiek tiek nukentėti.

JPEG formatas

- Vaizdo glaudinimas: JPEG (angl. Joint Photographic Experts Group), sumažina vaizdų failų dydį, atsisakydami kai kurių mažiau pastebimų detalių ir spalvų, todėl gaunamas mažesnis failas, tačiau sumažėja vaizdo kokybė.
- <https://sipdtdevelopers.wordpress.com/2014/01/23/digital-image-compression/>



(a) Original Image



(b) Image with 10:1 compression



(c) Image with 45:1 compression