

# Podatkovni višemedijski prijenos i racunalne mreže

Branko Jeren i Predrag Pale

Fakultet elektrotehnike i racunarstva  
Zavod za elektronicke sustave i obradbu signala

B. Jeren i P. Pale: Podatkovni višemedijski prijenos i racunalne mreže

PVPRM, LS&S (c) 2001

---

---

---

---

---

---

---

---

## Kodiranje

kodiranje i kompresija signala

B. Jeren i P. Pale: Podatkovni višemedijski prijenos i racunalne mreže

PVPRM, LS&S (c) 2001

---

---

---

---

---

---

---

---

## Principi tehnike kompresije

- Kompresija bez gubitaka (lossless compression)
- Kompresija sa gubicima (lossy compression)
  
- Entropijska kompresija
- Kompresija prema izvoru (source compression)

B. Jeren i P. Pale: Podatkovni višemedijski prijenos i racunalne mreže

PVPRM, LS&S (c) 2001

---

---

---

---

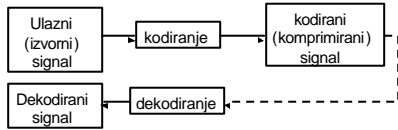
---

---

---

---

## Kompresija sa i bez gubitka



- Kompresija bez gubitka (lossless)
  - dekodirani signal identican je ulaznom
  - racunalni programi, baze podataka, tekstovi
- Kompresija s gubitkom (lossy)
  - dekodirani signal se razlikuje od ulaznog
  - ali se (ljudska) percepcija signala ne razlikuje (puno)
  - audio, video, fotografije

---

---

---

---

---

---

---

---

## Entropijska kompresija

- ne uzima u obzir prirodu signala
  - znacenje i nacin koristenja informacija
- svi signali se tretiraju kao slijed bitova
- uvijek je bez gubitka
  - dekodirani signal identican je ulaznom
- glavne metode
  - potiskivanje ponavljajucih sekvenci
  - statisticko kodiranje

---

---

---

---

---

---

---

---

## Potiskivanje ponavljajucih sekvenci

- uzastopono ponavljanje istog podatka zamjenjuje se šifrom i brojem ponavljanja
- dva oblika:
  - Potiskivanje nula ili praznina
  - Run-length encoding
- primjene
  - i binarne datoteke
  - i tekstovi
  - pa i (neke) slike

---

---

---

---

---

---

---

---

## Potiskivanje nula ili praznina

- kod signala kod kojih se samo jedan simbol (znak) ucestalo ponavlja
  - u binarnim datotekama, najčešće nula
  - u tekstualnim obično praznina ili nula u numerickim podacima
  - kod bi-level slika (samo crno ili bijelo)
- simbol koji se ponavlja
  - ostavi se jednom
  - a iza njega doda posebni simbol (F) i broj ponavljanja

7 5 4 0 0 0 0 0 0 0 0 0 0 0 5 S 7 5 4 F 10 5

---

---

---

---

---

---

---

---

---

---

## Run-Length Encoding

- u slucaju kada se bilo koji simbol može ponavljati
- simbol koji se ponavlja
  - ostavi se jednom
  - a iza njega doda posebni simbol (F) i broj ponavljanja
- ima smisla ako se ponavlja više od tri puta
- jednostavan i brz algoritam

l m e : CR S l m e : F 10 CR

---

---

---

---

---

---

---

---

---

---

## Statisticko kodiranje

- promatra se signal u cjelini
- pronalaze se skupine simbola koje se ponavljaju (uzorci = patterns)
- skupine se zamjenjuju oznakama
  - s manje bitova nego što ih sarzi sama skupina
- pripremi se "rijecnik" (code book)
  - koji sadrži uzorke i simbole kojima se zamjenjuju
  - mora biti na raspolaganju i na strani dekodiranja
  - ponekad se unaprijed pripeme (za neku vrstu signala)
  - ponekad se definiraju za svaki signal
- dva osnovna tipa
  - zamjena uzoraka (pattern substitution)
  - Huffmanovo kodiranje (Huffman encoding)

---

---

---

---

---

---

---

---

---

---

## Zamjena uzoraka

- koristi se za kompresiju teksta
- ceste rijeci se zamjenjuju posebnim znakovima
  - npr. rijec "multimedija" možemo prikazati kao "M"
- primjer
  - knjiga od 1000 stranica ima cca 4,000,000 slova što je oko 500,000 rijeci
  - dobar rječnik sadrži cca 60,000 različitih rijeci
  - dakle svaka se može opisati simbolom od 16 bita
  - znaci da se ova knjiga može predstaviti s 1,000,000 Byta
- ucinkovitost ovisi o jeziku
  - jako dobro za jezike s "nepromjenjivim" jezicima (engl.)
  - neki jezici već jesu takvi: ideogramski (kineski, japanski)
  - problematino za jezike bogate sufiksima (hrv.)

---

---

---

---

---

---

---

---

## Huffman encoding

- najčešće korišten način kodiranja
- radi na pojedinačnim simbolima (znakovima)
  - za dani niz podataka računa se frekvencija ponavljanja svakog okteta
  - iz tako dobivene tablice ponavljanja
  - huffmanovim algoritmom
  - računa se minimalni broj bitova za svaki simbol
  - te mu se pridjeljuje optimalni kod
- kodovi se pohranjuju u rječnik (code-book)
  - kojeg mora imati i idekoder
- koristi se i za nepokretne i pokretne slike
- novi rječnik se može konstruirati
  - za svaku novu sliku
  - za svaki novi okvir (frame), kod pokretnih slika

---

---

---

---

---

---

---

---

## Kompresija prema izvoru

- uzima u obzir prirodu ulaznog signala
  - kod stranice teksta ne zanima nas izgled, već sadržaj
  - kod govora je važna razumljivost i prepoznavanje govornika, ali nisu više frekvencije i npr. stanke
- zanemaruju se "nevažni" dijelovi signala
- postižu se visoki omjeri kompresije
- moguće su kompresije sa ili bez gubitka
- tri vrste kompresije:
  - kodiranje transformacijom (Transform encoding)
  - diferencijalno kodiranje (Differential Encoding)
  - vektorska kvantizacija (Vector quantization)

---

---

---

---

---

---

---

---

## Kodiranje transformacijom

- signal se transformira
  - iz polazne domene u drugu, apstraktnu, domenu koja je prikladna za kompresiju
  - na primjer iz vremenske u frekventijsku
- proces je inverzan
- nakon transformacije grupiraju se najvažniji koeficijenti (koji imaju najviše "energije")
  - njima se daje više bitova za prikaz
  - ili se "nevažne" potpuno zanemari
- najčešće se koristi diskretna kosinusna transformacija

---

---

---

---

---

---

---

---

---

---

## Diferencijalno kodiranje

- ne kodira se apsolutna vrijednost uzorka
- već samo razlika između stvarne vrijednosti i predviđene vrijednosti uzorka
  - to može biti vrijednost prošlog uzorka
  - ili srednja vrijednost zadnjih 10 uzoraka
  - ili bilo koja funkcija
- razlika se naziva
  - "prediction difference" ili
  - "error term"
- vrlo učinkovito kada se uzastopni uzorci malo razlikuju

---

---

---

---

---

---

---

---

---

---

## Diferencijalni PCM

- za predviđenu vrijednost uzima prethodni uzorak
- vrlo jednostavan i brz

---

---

---

---

---

---

---

---

---

---

## Delta modulacija

- varijanta DPCM
- razlika se prikazuje samo jednim bitom
- za signale koji se ne mijenjaju previše brzo (npr. kod niskih frekvencija)

---

---

---

---

---

---

---

---

## Adaptive DPCM

- ne koristi konstantnu funkciju predviđanja
  - već je ADPCM funkcija varijabilna u ovisnosti o
  - kratkotrajnim promjenama uzorkovanog signala
- koristi se ekstrapolacija
- kao i kod DPCM, kodira se razlika predviđene i stvarne vrijednosti uzorka

---

---

---

---

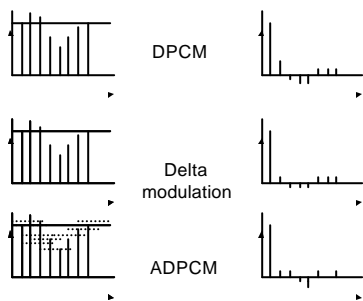
---

---

---

---

## Diferencijalno kodiranje



---

---

---

---

---

---

---

---

## Vektorska kvantizacija

- posebni slučaj zamjene uzoraka
- ulazni signal je podijeljen na male blokove koje nazivamo vektorima
  - npr. kod slike to mogu biti mali pravokutni dijelovi
- postoji tablica (rijecnik) najčešćih uzoraka
  - mora biti poznata i dekoderu
  - može biti poznata unaprijed ili ovisiti o signalu
- svaki se vektor uspoređuje s rijecnikom
  - traži se najslabiji uzorak
  - prepisuje se njegova šifra
- razlika vektora i uzorka
  - se također bilježi
  - kodira se na različite načine
- vrlo korisno za kompresiju govora

B. Jeren i P. Pale: Podatakovni višemedijski prijenos i računalne mreže

PVPRM, LS&S (c) 2001

---

---

---

---

---

---

---

---

## Fraktalna kompresija

- najčešće se koristi za fotografije
- slika je podijeljena na male kvadrate
- koji se uspoređuju sa svim ostalima na slici
- pri tome se primjenjuju transformacije
  - pomaka, povećanja, rotacije i sl.
- kodira se fraktal i funkcija transformacije
- koristi prividni rijecnik
  - ne treba ga slati dekoder, jer nastaje sam u procesu
- očekivanja su velikih omjera kompresije

B. Jeren i P. Pale: Podatakovni višemedijski prijenos i računalne mreže

PVPRM, LS&S (c) 2001

---

---

---

---

---

---

---

---

## Složenost kodiranja i dekodiranja

- bolje omjere kompresije daju složeniji algoritmi
- složeni algoritmi traže jaki CPU i više vremena
- za real-time trebamo brze algoritme
- složene možemo koristiti za off-line
- asimetrični algoritmi traže bitno više vremena za kodiranje nego za dekodiranje
  - pogodni su za:
    - medije za pohranu
    - broadcast
    - off-line

B. Jeren i P. Pale: Podatakovni višemedijski prijenos i računalne mreže

PVPRM, LS&S (c) 2001

---

---

---

---

---

---

---

---

# Podatkovni višemedijski prijenos i racunalne mreže

[pvprm.zesoi.fer.hr](http://pvprm.zesoi.fer.hr)

PVPRM@zesoi.fer.hr

---

---

---

---

---

---

---

---