

IZPITNA VPRAŠANJA 2011

1. Bilinearna interpolacija

Najprej linearna interpolacija po vseh vrsticah. Potem linearna interpolacija po vseh stolpcih. Za izračun piksla rezultata upoštevamo vse pikse izvirne slike v okolici +-1

2. Kje pride do izgube pri kodiranju jpeg

Ker smo ljudje bolj občutljivi za spremembe v svetlosti kot barvi, lahko barvni kanal (CbCr) vzorčimo bolj grobo kot svetlost (Y) – (*chroma subsampling*)

Vzorčenje vrednosti označimo s tremi števili, ki povedo koliko vrednosti uporabimo zakodiranje 4 pikslov

-4:4:4 pomeni polno ločljivost, po 4 vrednosti za vsako komponento

3. Kaj je ideja efektov, ki uporabljajo zakasnitve

Vrsta efektov temelji na preprostih zakasnitvah.

K je zakasnitev, ki je

-statična

-ali pa jo lahko dinamično spreminjamo – moduliramo

4. Kaj je gama korekcija, kje se uporablja

Za percepcijo svetlosti (kot za veliko drugih čutil) velja **Weberjev** zakon:

-kar pomeni, da smo pri majhnih intenzitetah bolj občutljivi na spremembe kot pri velikih

-oz. z drugo besedo, da se percepcija svetlosti obnaša cca. logaritmično

Posledično svetlosti oz. barv **ni smiselno kodirati linearno** in porabiti enako bitov za vse intenzitete

Pomanjkljivost odpravlja s tem, da sliko ob zajemu zakodiramo (1/g)

- $s = c \cdot r^{1/\gamma}$

-porabimo več bitov za nižje intenzitete

-tipična gama je 2.2

Digitalni fotoaparati, kamere zajeto sliko že gama zakodirajo (1/g)

-jpg, mpeg, dvd ... vsebujejo gama zakodirane vrednosti

Z gama korekcijo v računalnikih upravlja sistem za upravljanje z barvami

-ob prikazu se slika ustrezno dekodira glede na napravo, ki jo prikazuje

5. Kako lahko spremenimo višino signala z vzorci...

Predvajano **višino** spreminjamo glede na lego na tipkovnici

-najlažje s **pohitritvijo /upočasnitvijo** predvajanja

-to lahko dosežemo z navideznim večanjem ali manjšanjem frekvence vzorčenja

Za koliko spremenimo frekvenco vzorčenja?

-oktava (razpon med 2 in 2*f) v enakomerni uglasitvi je razdeljena na 12 intervalov

-percepcija višine je logaritmična

- vsak polton ima torej frekvenco za $2^{1/12}$ višjo
- torej frekvenco vzorčenja pri dvigu za en polton spremenimo za $2^{1/12}$

6. Kateri barvni prostor zajema vse barve, ki jih vidimo in za kaj se uporablja

- XYZ se uporablja za **pretvorbe** med prostori
- je neodvisen od reprodukcijskih naprav
 - predstavi lahko vse vidne barve

7. Kaj lahko poveš o magnitudi ... vzorca (3 poševne bele pikce v črnem kvadratu)

8. Kako bi uporabil beseda1 AND beseda2 pri obrnjenem indeksiranju.....

9. Kaj je vektor premika/razdalj

- Lahko poskusimo **napovedati / ugotoviti**, kam se posamezen del slike premakne
- računamo razliko do napovedi – **kompensacija gibanja** (*motion compensation*)
 - bolj učinkovito stiskanje

Tipični koraki

- za nek del slike ugotovimo kam se premakne v naslednji sliki (*motion estimation*)
- dobimo **vektor gibanja**
- del slike premaknemo za ugotovljeni vektor gibanja (*motion compensation*) – naredimo napoved naslednje slike
- kodiramo razliko med naslednjo sliko in napovedjo – premaknjenim delom slike

10. Zakaj je pri videu YUV in YcbCr

- Poleg RGB, se v videu uporabljajo tudi drugi zapisi barv, ki ločijo svetlost od barve
- oko je bolj občutljivo na svetlost kot barvo, in ju lahko drugače kodiramo

PAL standard uporablja YUV

- Y je (gama kodirana) svetlost
- UV kodirata barvo, $U=B-Y$, $V=R-Y$

Enostavna pretvorba v/iz RGB

V digitalnem videu se veliko uporablja zapis YCbCr (JPEG, MPEG)

- analogni komponentni video pa YPbPr (YCbCr je le skaliran YPbPr)

Je zelo podoben YUV

- točni koeficienti pretvorbe iz/v RGB so odvisni od barvnega prostora
- npr. HDTV in SDTV in JPEG imajo nekoliko drugačne pretvorbe

11. Kako amplituda vpliva na glasnost

Lahko jo merimo v dB relativno glede na najmočnejši signal

- $A = 20 \cdot \log_{10}(|x| / A_{\max})$
- če je sinusoida med -1 in 1 ($A_{\max} = 1$)
 - je amplituda $A = 20 \cdot \log_{10}(|x|)$

Amplituda v dB je torej negativna vrednost, pove koliko je v nekem časovnem trenutku manjša od najvišje možne.

12. Kaj je identifikacija avdio posnetka

Baza: Velika zbirka posnetkov, ki jih želimo identificirati

Cilj: Imamo kratek del posnetka, s katerim želimo identificirati originalni posnetek (točno in samo tega).

Opombe: Povpraševanje je zelo **specifično** – želimo identificirati natanko isti posnetek pesmi, kot je vsebovan v povpraševanju ne kake druge verzije

Uporaba

Uporabnik sliši glasbo

Posname kratek del (5-15 sekund) skladbe z mobilnim telefonom

Iz posnetka se izloči t.i. prstni odtis dela skladbe, ki se pošlje na strežnik

Na podlagi prstnega odtisa strežnik identificira posnetek v bazi

-Strežnik uporabniku vrne metapodatke (ime skladbe, izvajalec ...)

13. Zakaj se uporablja fitriranje z mediano

Želimo zmanjšati šum brez prevelikega glajenja

-*median filter* zamenja vrednost piksla z mediano vseh pikslov znotraj filtra

Filtriranje z mediano je najbolj učinkovito pri točkastem šumu (sol in poper)

14. Kako izmerimo višino ton

Večina tonov je **kompleksnih** (imajo več frekvenc)

-lahko imajo t.i. **osnovno frekvenco**

-in druge frekvenčne komponente (**harmoniki** oz. alikvotni toni), katerih frekvence so (približni) večkratniki osnovne frekvence

-Višina kompleksnega tona je odvisna od vseh frekvenčnih komponent

-v veliki meri v razmerju med frekvencami komponent

-kot pri enostavnih tonih na višino vpliva tudi glasnost

15. MPEG-1 kodiranje, kako določimo št. bitov za kvantizacijo

16. Kaj je CMYK in kaj YCbCr?

RGB prostori temeljijo na **seštevanju** barv

-npr. žarka svetlobe različnih barv se seštejeta, da dobimo neko novo barvo

Za tiskalnike/črnilo rabimo drug prostor, saj črnilo **odšteva**

-rumeno črnilo na belem listu "odšteje" modro, odbije rdečo in zeleno

-rabimo torej osnovne barve, ki odštevajo – CMY

CMY ne morejo dobro reproducirati temnih delov

-zato dodana še K - black

17. Kaj pri MPEG-1 pomeni B, P in I?

Slike kodiramo na tri načine:

-I (*intra frame*) so slike kodirane v celoti, na podoben način kot JPEG

-tipično ena na vsake 15 sličic

-P (*predictive frame*) je slika kodirana s kompenzacijo gibanja iz prejšnje P ali I slike

-B (*bi-directional frame*) je slika kodirana s kompenzacijo gibanja iz prejšnje in naslednje P ali I slike

18. Kaj določa ton zvoka?

Višina zvoka (*pitch*) je **perceptivna** lastnost, po kateri lahko zvoke razvrstimo po višini (A je višji kot B)

-vsi zvoki nimajo višine (npr. šum)

-tistim, ki imajo višino, lahko rečemo **toni**

Višino zvoka X (v Hz) izmerimo kot frekvenco sinusoide, ki jo poslušalec zazna kot enako visoko kot X

višina **enostavnih tonov** (z le eno frekvenco) je torej v veliki meri enaka frekvenci tona

-na višino vpliva tudi glasnost

19. Kako se amplituda (glasnost) zvoka spreminja skozi čas?

npr. pri orglah se ob pritisku na tipkah takoj sliši zvok, ki ostane enako glasen do konca in takoj zamre, ko tipko spustimo

-pri kitari je na začetku glasen, nato počasi pada

Vzpon - *Attack*: kako hitro zvok doseže največjo jakost

Spust - *Decay*: kako hitro zvok iz največje pade na običajno glasnost

Trajanje – *Sustain*: kako se glasnost spreminja, ko zvok traja (npr. dokler ne spustimo tipke)

Sprostitev – *Decay*: kako hitro zvok upade, ko se konča (spustimo tipko), npr. zvon zelo dolgo, klavir kratko

20. Kaj je Levenshteinova razdalja?

Gradimo matriko vseh možnih pretvorb iz ene besede v drugo

V celicah matrike (i,j) seštevamo cene operacij potrebnih za transformacijo niza (1..j) v drugega (1..i)

če sta znaka (i,j) enaka:

$m[i,j] = \min(m[i-1,j]+1, m[i,j-1]+1, m[i-1,j-1])$

če znaka nista enaka

$m[i,j] = \min(m[i-1,j]+1, m[i,j-1]+1, m[i-1,j-1]+1)$

		F	A	S	T
	0	1	2	3	4
C	1	1	2	3	4
A	2	2	1	2	3
T	3	3	2	2	2

vstavljanje – cena == 1
Razdalja med **brisanje** – cena == 1
fast in cat je 2 **zamenjava** – cena == 1
kopiranje – cena == 0

21. Kaj je tf-idf?

term-frequency – inverse-document frequency

Produkt *tf* in *idf* mer

Najboljša mera za uteževanje

Utež se povečuje s številom pojavitev simbola v dokumentu in z redkostjo simbola

$$W_{t,d} = (1 + \log_{10} tf_{t,d}) \log_{10} N/df_t$$