




Digitalizacija

Analogni zapis

- ▶ kontinualna priroda signala : signali koji nas okružuju u prirodi se mogu predstaviti neprekidnim funkcijama
- ▶ analogni signali su konstantne promenljive, i po broju mogućih vrednosti signala u datom trenutku, kao i po broju tačaka u signalu u datom vremenskom periodu.
- ▶ zvučni signal predstavlja promenu pritiska vazduha u zadatoj tački i to kao neprekidnu funkciju vremena
- ▶ slika se može opisati intenzitetom svetlosti određene boje (tj. određene talasne dužine) u datom vremenskom trenutku i to kao neprekidna funkcija prostora
- ▶ analogni zapis uspostavlja analogiju između signala koji je zapisan i određenog svojstva medijuma na kome je signal zapisan: kontinualne promene signala koji se zapisuje opisuju se kontinualnim promenama određenog svojstva medijuma na kojem se signal zapisuje
- ▶ neki analogni medijumi za zapisivanje signala: magnetna traka za zvuk i sliku, papir za sliku,...
- ▶ tehnologija je obično veoma jednostavna
- ▶ relativno niski kvalitet; kvalitet zapisa signala zavisi od kvaliteta medijuma
- ▶ nestalnost medijuma, njegova promenljivost tokom vremena i podložnost spoljašnjim uticajima vodi do neizbežnog pada kvaliteta analogno zapisanog signala; npr. papir vremenom žuti što uzrokuje pad kvaliteta analognih fotografija tokom vremena.
- ▶ gubitak kvaliteta prilikom pravljenja kopija; otežan prenos signala na daljinu
- ▶ npr. fotokopija fotografije je daleko lošijeg kvaliteta od originala

Digitalizacija


- proces pretvaranja analognog signala u digitalni oblik; beleženje analognog signala u digitalnom obliku
- proces prevođenja jednog objekta (slike, zvuka, dokumenta ili signala) u digitalni oblik
- konvertovanje u binarni kod
- moguće je digitalizovati sve vrste materijala, gradiva, od teksta, preko audio i video zapisa, sve do trodimenzionalnih objekata
- zašto?
 - zaštite izvora
 - povećanja dostupnosti
 - stvaranja nove ponude i usluga
- kako?
- vrednost analognog signala se čita u pravilnim vremenskim intervalima (frekvenciji) ili određenim tačkama prostora. Svako ovo čitanje naziva se *uzorak* i može se smatrati neograničeno preciznim u ovom delu procesa; uzorkovanje vrednosti signala za dati momenat intervala.
- izmerene vrednosti se zapisuju. Ovim je svaki digitalno zapisani signal predstavljen nizom brojeva koji se nazivaju odbirci ili semplovi (engl. sample).
- što je frekvencija uzorkovanja veća, dobijeni je signal kvalitetniji, što znači da će uzorkovanje frekvencijom od 10 kHz dati bolji rezultat nego uzorkovanje frekvencijom od 2 kHz

- 
- Koliko često je potrebno vršiti merenje da bi se polazni kontinualni signal mogao verno rekonstruisati? Tvrđenje kaže: signal je dovoljno meriti dva puta češće od najviše frekvencije koja se u njemu javlja

1. diskretizaciju ili uzorkovanje

2. kvantizacija - u kojoj se uzorci zaokružuju na fiksni skup celih brojeva u procesu u računarskoj terminologiji poznatom kao kvantovanje.

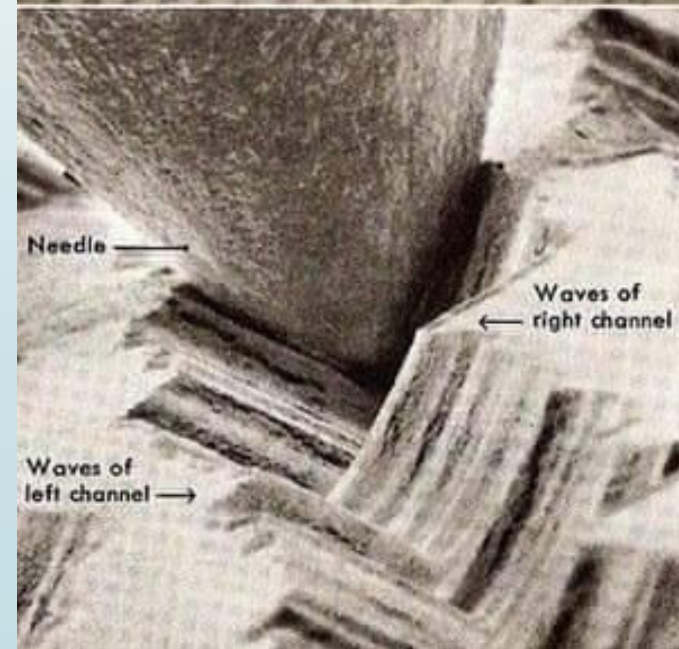
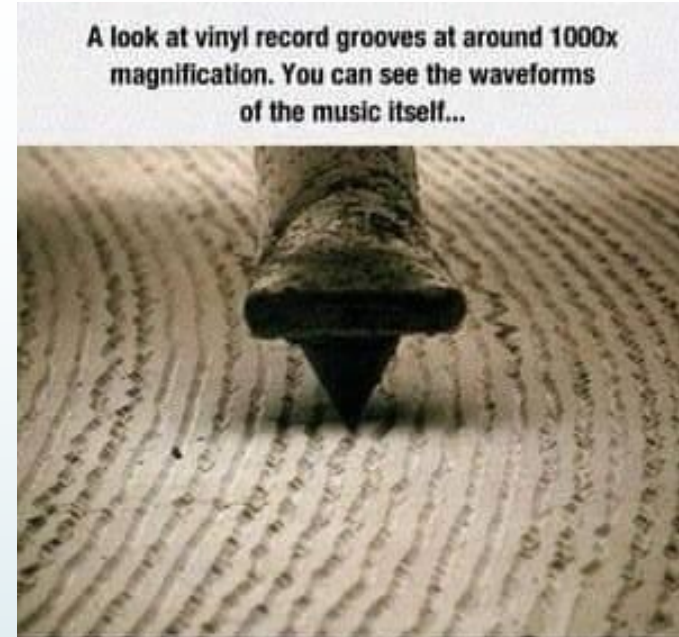
- kvalitet reprodukcije digitalnog zapisa ne zavisi od toga kakav je kvalitet medija na kome su podaci zapisani, sve dok je medijum dovoljnog kvaliteta da se zapisani brojevi mogu razaznati
- npr. ako papir sadrži zapis brojeva koji predstavljaju vrednosti boja u tačkama digitalno zapisane fotografije, činjenica da papir žuti ne bi predstavljala problem dok god se brojevi mogu razaznati
- digitalni zapis omogućava kreiranje apsolutno identičnih kopija što dalje omogućava prenos podataka na daljinu. Nema gubljenja kvaliteta prilikom kopiranja: ukoliko umnožimo CD na kojem su zapisani brojevi koji čine zapis neke fotografije, kvalitet slike ostaje apsolutno isti.
- obrada digitalno zapisanih podataka se svodi na matematičku manipulaciju brojevima i ne zahteva (za razliku od analognih podataka) korišćenje specijalizovanih mašina.
- osnovni problem implementacije digitalnog zapisa predstavlja činjenica da je neophodno imati veoma razvijenu tehnologiju da bi se uopšte stiglo do iole upotrebljivog zapisa. Na primer, izuzetno je komplikovano napraviti uređaj koji je u stanju da 40 hiljada puta izvrši merenje intenziteta zvuka. Jedna sekunda zvuka se predstavlja sa 40 hiljada brojeva, za čiji je zapis neophodna gotovo cela jedna sveska. Ovo je osnovni razlog zbog čega se digitalni zapis istorijski javio kasno. Kada se došlo do tehnološkog nivoa koji omogućava digitalni zapis, on je doneo mnoge prednosti u odnosu na analogni.

- 
- šta onda utiče na kvalitet digitalnog signala?
 - oko 5 procenata tekstualnih dokumenata je digitalizovano do 2006 (en wiki)
 - skoro sva snimljena muzika je digitalizovana. Kao i oko 12 % od 500,000+ filmova koji se nalaze u Internet Movie Database (na DVD snimljeni)

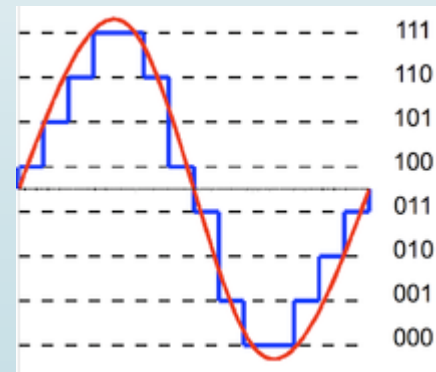
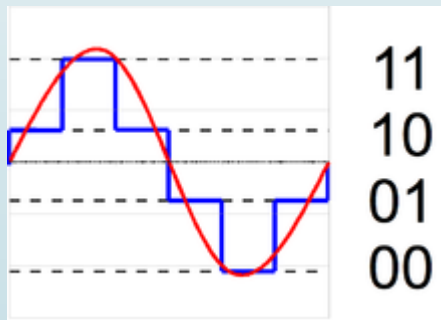
 - postupci digitalizacije razlikuju se u zavisnosti od vrste materijala, gradiva koje se digitalizuje
 - tekstualno gradivo, dokumenti, može se digitalizovati na tri načina; prepisivanjem, skeniranjem i fotografisanjem digitalnim fotoaparatom. U slučaju skeniranja i fotografisanja potrebno je sprovesti i optičko prepoznavanje slova (eng. OCR - *Optical character recognition*).
 - šta još može da se digitalizuje?
 - digitalnih prezentacija geografskih područja, gde se koristi raster ili vektorska grafika u geografskim informacionim sistemima, i tako stvaraju elektronske karte, bilo iz slika različitih satelitskih fotografija ili digitalizacijom sa tradicionalnih papirnih izvora: mapa, grafika i fotografija
 - Digitalizacija trodimenzionalnog gradiva često se naziva i 3D digitalizacija.
 - Kod 3D digitalizacije upotrebljavaju se skeneri namijenjeni skeniranju volumena, a oni se dele na kontaktne i beskontaktne.
 - potrebno je odabrati skener što veće rezolucije, visoko kvalitetni digitalni fotoaparat

Zvuk (audio materijal)

- ▶ mehanički uređaji za analogan signal
- ▶ npr gramofon



- ▶ audio materijal se digitalizuje tako što se zvučni izlaz uređaja za reprodukciju audio sadržaja (gramofon, kasetofon i sl.) poveže s računarom koji je opremljen zvučnom karticom i adekvatnim programom za obradu zvuka
- ▶ uzorkovanje određuje koliko će se često zapisivati informacije, a da to bude dovoljno verno za reprodukciju
- ▶ potrebno je obratiti pažnju i na raspon ljudskog sluha. Poznato je da većina ljudi raspoznaje frekvencije od 20 Hz sve do 20.000 Hz i obično se uzima dvostruka maksimalna frekvencija koju ljudsko uho raspoznaje. U praksi je frekvencija 44,1 kHz (44100 uzoraka/sec) postala standard
- ▶ nakon uzorkovanja zvučnog signala potrebno je kvantizovati njegovu amplitudu. Ako je uzorkovano standardom od 40,1 kHz, to znači da je svaka sekunda zvučnog signala podeljena na 41.000 delova, a svaki deo sadrži informaciju o amplitudi. Svaki taj deo može se podeliti na beskonačno mnogo delova, a kvantizacijom se određuje dužina binarnog niza, odnosno delovi amplitude se svode na određenu dužinu.

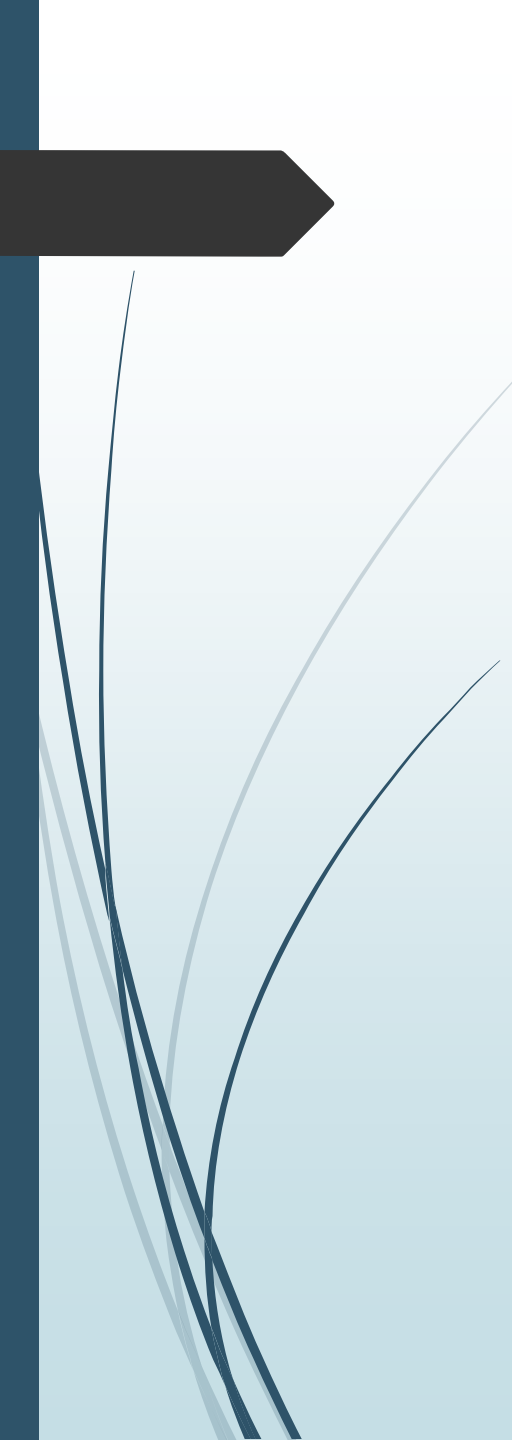


- ▶ 2-bitna rezolucija sa četiri nivoa kvantizacije upoređena sa analognim signalom
- ▶ 3-bitna rezolucija sa osam nivoa

- digitalni audio zapis se može skladištiti, čuvati na bilo kom uređaju za skladištenje digitalnih podataka: CD, Blue Ray, digitalni audio player, hard disk, USB flash disk, ...
- CD Optical disc
- Dijametar 120 milimetara (4.7 in)
- Kapacitet do 700 MiB podataka, do 80 minuta nekompresovanog zvuka
- Read mechanism 780 nm talasne dužine (infrared) laser, 1200 Kibit/s (1×)
- Write mechanism 1200 Kibit/s (1×)
- Brzina 1× znači 153.6 kB/s (150 KiB/s)

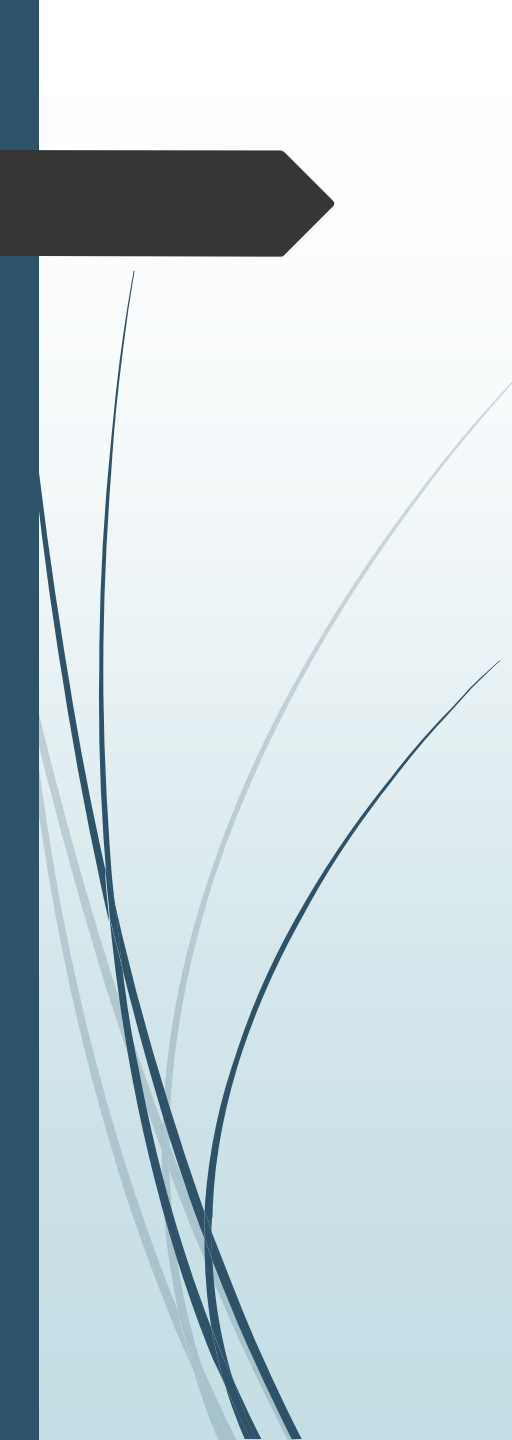
- DVD (digital video disc) Optical disc
- Kapacitet 4.7 GB (single-sided, single-layer – common)
 8.5 GB (single-sided, double-layer)
 9.4 GB (double-sided, single-layer)
 17.08 GB (double-sided, double-layer – rare)
- Read mechanism 650 nm laser, 10.5 Mbit/s (1×)
- Write mechanism 10.5 Mbit/s
- Brzina 1× znači 1,385 kB/s (1,353 KiB/s);

- Blue-ray optički disk
- Kapacitet 25 GB (single-layer)
 50 GB (dual-layer)
- Brzina 1× znači 36 Mb/s (4,5MB/s);

- 
- tehnike kompresije audio podataka, npr MP3, Advanced Audio Coding, Ogg Vorbis, FLAC, su razvijene da bi se smanjila veličina datoteka, fajlova
 - uređaj mora da ima deo, hardware, koji može da dovoljno brzo dekompresuje signal i pretvori u analogni da bi mi mogli da ga čujemo, vidimo...
 - algoritmi za kompresiju se implementiraju kao software - audio kodeci
 - između ostalog, signal se kodira tako što se koristi manje bitova nego u originalnoj datoteci

 - u poređenju sa CD kvalitetom digitalnog zvuka, MP3 (MPEG-1, MPEG-2 Audio Layer III) kompresija obično postiže 75 do 95% smanjenja originalnog fajla. Stoga su $1/4$ do $1/20$ veličine originalne datoteke. Što je važno za skladištenje i prenošenje podataka
 - npr uobičajeno MP3 kodiranje je CBR (Constant bitrate) 128 kbit/s što rezultira veličinom fajla $1/11$ (=9% tj 91% kompresija) originalnog CD kvaliteta
 - oko 100 - 150 mp3 fajlova može da stane na CD (prosek trajanja 3-5 minuta = cc 560 min audio)
 - ekstenzija datoteke .mp3, a internet media tip može da bude i mpeg, mpa

 - AAC, naslednik MP3 kompresije, uglavnom postiže bolji kvalitet zvuka u sličnoj stopi bita
 - ekstenzija datoteke .aac, .mp4, .3gp, a internet media tip može da bude i mpeg4, 3gpp
 - standardni audio format za YouTube, iPhone, iTunes, Nintendo, PlayStation 3, Android, BlackBerry.

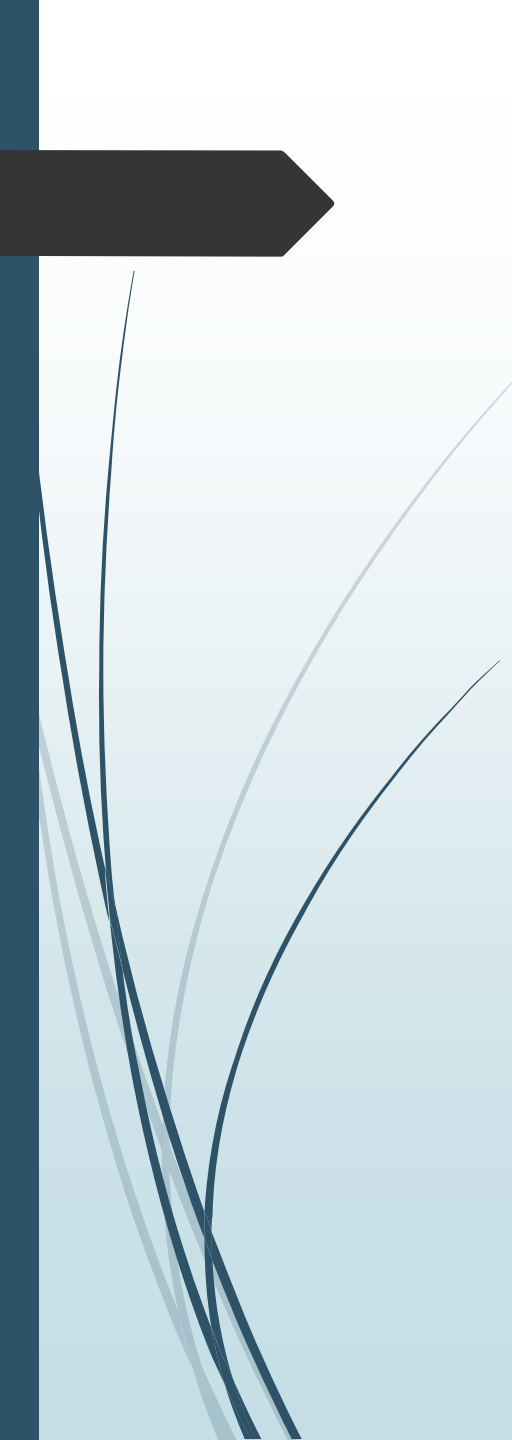
- 
- ▶ Vorbis je besplatna i otvorenog koda kompresija
 - ▶ ekstenzija .ogg
 - ▶ FLAC je i audio format i kodek
 - ▶ FLAC algoritam smanjuje za 50-60% veličinu i dekompresuje u identičnu datoteku originalnoj

 - ▶ CD audio, npr, može da podrži brzinu uzorkovanja 44.1 kHz (44,100 uzoraka po sekundi) i 16-bitnu rezoluciju za svaki stereo kanal
 - ▶ npr. pesmu s gramofonske ploče pretvoriti u digitalni audio format. Ako se koristi 44.100 uzoraka u sekundi, a uzorci su 16 bitni. Na taj način ljudsko uvo ne može prepoznati razliku između izvora (gramofonske ploče) i digitalnog oblika, jer kao što je oko tromo, i 24 slike u sekundi prepoznaje kao pokret (tj. ne vidi trzanje), tako i uvo 44.100 uzoraka u sekundi podeljenih na 65536 nivoa (2 na 16) ne može razlikovati od originala

▶

Slike

- digitalizuje se skeniranjem ili fotografisanjem
- potrebno je odabrati skener što veće rezolucije, visoko kvalitetni digitalni fotoaparati
- Rezolucija - broj piksela ili tačaka po inču, kod digitalizacije rezolucija označava broj uzorkovanja tokom skeniranja
- Bitna dubina tačke - svakoj tački u digitalnoj slici pridružen je binarni niz. Broj cifara tog niza naziva se bitna dubina
- Boja - pre digitalizacije bitno je odrediti kako će se upotrebljavati gradivo; samo za čuvanje, na monitoru ili će postojati i mogućnost ispisa. Tome treba prilagoditi sistem prikaza boje. Najčešći su RGB, CMYK i CIELab
- može biti vektorska ili rasterska, u zavisnosti od toga da li je rezolucija fiksirana ili ne
- rasterska slika ili bitmapirana ima fiksni konačan broj digitalnih vrednosti, koji se zovu pixeli. Pixeli se čuvaju u memoriji kao rasterska slika ili rasterska mapa, dvo-dimenzioni niz
- vektorske slike su nastale iz geometrije (vektora). U matematičkim terminima, vektor sadrži (je određen) pravac i dužinu
- kod npr. bilborda, slika sadrži elemente i rasterske i vektorske grafike: text – vektorska grafika, dok je fotografija rasterska
- Digitalizacija je primaran način čuvanja slika u obliku pogodnom za prenos i obradu na računaru, bez obzira da li se radi o objektima skeniranim iz dvodimenzionalnih analognih originala ili uz pomoć uređaja poput digitalnih foto-aparata ili kompjuterizovanih tomografskih instrumenta kao što je CT skener, ili drugim načinima beleženja preciznih dimenzija stvarnog objekta, kao što je npr. zgrada uz pomoć uređaja za 3D skeniranje.

- 
- GIF (Graphics Interchange Format)
 - slike se kompresuju pomoću algoritma za „lossless data compression“ tehniku da se smanji veličina datoteke bez opadanja kvaliteta slike
 - podržava do 8 bita po pixelu, što omogućava 256 različite boje iz 24-bit RGB Sistema
 - podržava i animacije sa 256 boja za svaki frame
 - zbog ograničenja boja GIF format nije pogodan za fotografije, ali je veoma dobar za brzi prenos, internet, logo, ... jednostavnije slike
 - JPEG (Joint Photographic Experts Group) uobičajen metod za „lossy compression“
 - stepen kompresije se može podešavati, što omogućava balaniranje između veličine i kvaliteta slike. Obično je 10:1 kompresija koja ne gubi vidno u kvalitetu slike.
 - JPEG algoritam za kompresiju je najbolji za fotografije i slike sa malim varijacijama tonova i boja. Za upotrebu na web-u, gde je količina podataka korišćenih za sliku važna, JPEG je veoma popularna. JPEG/Exif je uobičajen format koji koriste digitalne kamere.
 - podržava maksimalnu veličinu slike 65,535×65,535 pixela
 - PNG (Portable Network Graphics) je osmišljen za transfer slika putem interneta, ne za profesionalni kvalitet, ne podržava ne-RGB kolor sistem kao npr CMYK
 - podržava palete boja 24-bit RGB, 32-bit RGBA, grayscale

Video gradivo

- ▶ Digitalizacija video gradiva svodi se na digitalizaciju pokretnih slika i digitalizaciju audio sadržaja.
- ▶ Faktori bitni za kvalitetu pokretnih slika isti su kao i kod klasičnog slikovnog gradiva. Ono na što je potrebno obratiti pažnju je broj slika u sekundi.
- ▶ digitalizacija je najzastupljenija u bibliotekarstvu, arhivistici i muzeologiji.
- ▶ <http://www.ncd.org.rs/>