

The background features a central white circle containing text. To the left and right of the circle are wavy, layered shapes in shades of green and yellow. Several short, dashed green lines are positioned along the top-left edge of the white circle.

# Digitálne faximile - obraz

DUŠAN KATUŠČÁK

# Snímanie

1. Snímanie je jeden z procesov digitalizácie.
2. Vykonáva sa pomocou vhodného technického zariadenia na digitalizáciu,
3. Konkrétne sú to zariadenia na zachytenie **digitálneho obrazu**
  - digitálne fotoaparáty
  - kamery,
  - knižné skenery,
  - iné skenery.

# Snímanie

- Snímanie vo verejných informačných inštitúciách – v súlade s právnymi predpismi
- V archíve je možné so súhlasom vedenia archívu (bádateľne) snímanie archívnych dokumentov
- klasickou kamerou a digitálnou kamerou (ďalej len "kamera") a fotografickou technikou
- na účely automatickej transkripcie, pokiaľ je to možné, použijeme dokumenty nasnímané profesionálnymi skenermi a obrazmi v najvyššej dosiahnuteľnej kvalite
- minimálna kvalita skenovania by mala byť 300 DPI
- nakoľko pri historických rukopisoch ide *de facto* o grafiku, je vhodné skenovať vo vyššej kvalite.
- pre platformu Transkribus je možné snímať dokumenty do formátu veľkosti A3 zariadením **ScanTent** so softvérom **DocScan**.

# Formáty obrázkov : Formát JPG, JPEG.

- Najrozšírenejší je formát, ktorý sa vyskytuje s príponou .jpg, .jpeg alebo .JPG, .JPEG.
- Medzi nimi nie je žiadny rozdiel.
- V tomto formáte ukladajú súbory všetky fotoaparáty aj mobilné zariadenia, ak používame napríklad DocScan.
- V niektorých aparátoch je možné voliť jeden formát alebo snímame v dvoch formátoch JPG a RAW, ARW.
- Výhodou formátu JPG je, že sa obrázok dá zobrazit' prakticky v každom zariadení - v mobilnom telefóne, televízore alebo vo webovom prehliadači.
- Zaberá málo miesta na disku, je úsporný, pretože ide o kompresiu so stratou.
- Nevýhodou tohto formátu je, že každou úpravou obrázkov stráca kvalitu pri každom uložení.
- V projektoch transkripcie používame na snímame mobilnými zariadeniami formát JPG na archivovanie a v transkripcii spravidla pracujeme s derivovaným formátom PDF.

# Formáty obrázkov : Formát RAW

- znamená, že nasnímaný súbor je „surový“, nespracovaný
- dáta nie sú komprimované.
- dáta v tomto formáte sú veľmi veľké a na ich spracovanie je potrebný špeciálny softvér,
  - napríklad komerčný Zoner Photo Studio alebo open source FastStone Image Viewer.
- výsledné obrázky majú vysokú kvalitu a sú po úprave hodné na kvalitné editovanie.

# Formáty obrázkov : Formát TIFF

- Vyskytuje sa s príponami .tiff, tif
- pri ukladaní do tohto formátu spravidla nedochádza ku kompresii dát
- ak áno, tak ide o bezstratovú kompresiu aj pri opakovanom ukladaní
- súbor zachováva maximum informácií z formátu RAW pri editácii.
- nevýhodou je veľkosť súborov vo formátoch TIFF.
- v profesionálnych projektoch digitalizácie je formát TIFF najvhodnejší na dlhodobé archivovanie

# Formáty obrázkov

- Väčšina snímok v projektoch digitalizácie je vo formáte JPEG,
- Formát JPEG (JPG) je *de facto* široko používaným štandardom na ukladanie digitálnych snímok
- V digitalizácii hovoríme o výsledkoch snímania – obrazy
  - A) digitalizáty
  - B) digitálne faksimile

# Pixel – základ pre ukladanie digitálneho obrazu

- **pixel** – je najmenšia jednotka obrazových informácií
- **pixel** je skratka pre prvok obrázka
- **skratku "pix"** znamená jednu plnofarebnú bodku obrázka.
- samotný **pixel nemá predpísaný tvar** - môže byť štvorcový, kruhový alebo ľubovoľný,
- predstavme si ho ako obdĺžnik, ktorý je vytvorený rozrezaním obrazu na určitý počet vertikálnych a horizontálnych segmentov



# Pixel

(Podľa: [Formáty pro ukládání fotografií - 1.díl: základy | Digimanie](#))



# Rozlíšenie

- Ak rozrežeme skutočný obrázok na určitý počet vertikálnych a horizontálnych prvkov, v skutočnosti vytvoríme mozaiku obrazu.
- Každý segment (obdĺžnik) musíme opísať iba s jedným pixelom - jednou farbou.
- Týmto spôsobom pixel predstavuje priemernú farbu každého segmentu mozaiky a nevyhnutne zjednodušuje obraz.
- Logicky teda platí, že čím viac pixelov, tým jemnejší/presnejší obraz popisujeme.
- Okrem toho, ak poznáme rozmery obrázka, je možné vypočítať veľkosť jedného pixelu so znalosťou jeho rozlíšenia v pixeloch

## *Príklad veľkosti pixelov pre rôzne zariadenia (médiá)*

	<b>Rozlišení</b>	<b>Rozměr</b>	<b>Pixel [mm]</b>
<b>Televize</b>	720x576	42" úhlopříčka	1,18x1,11
<b>LCD obrazovka</b>	1600x1200	20" úhlopříčka	0,255x0,255
<b>Fotografie</b>	3000x2000 (6 MPix)	13 x 9 cm	0,043x0,045

# Farba pixelov

V každom pixeli je v procese snímania zakódovaná **farba, jas a ďalšie parametre**, aby zodpovedali schopnostiam ľudského oka.

Najbežnejšou reprezentáciou používanou v digitálnych fotoaparátoch na snímanie obrázkov a v počítačoch na zobrazenie je takzvaná reprezentácia RGB,

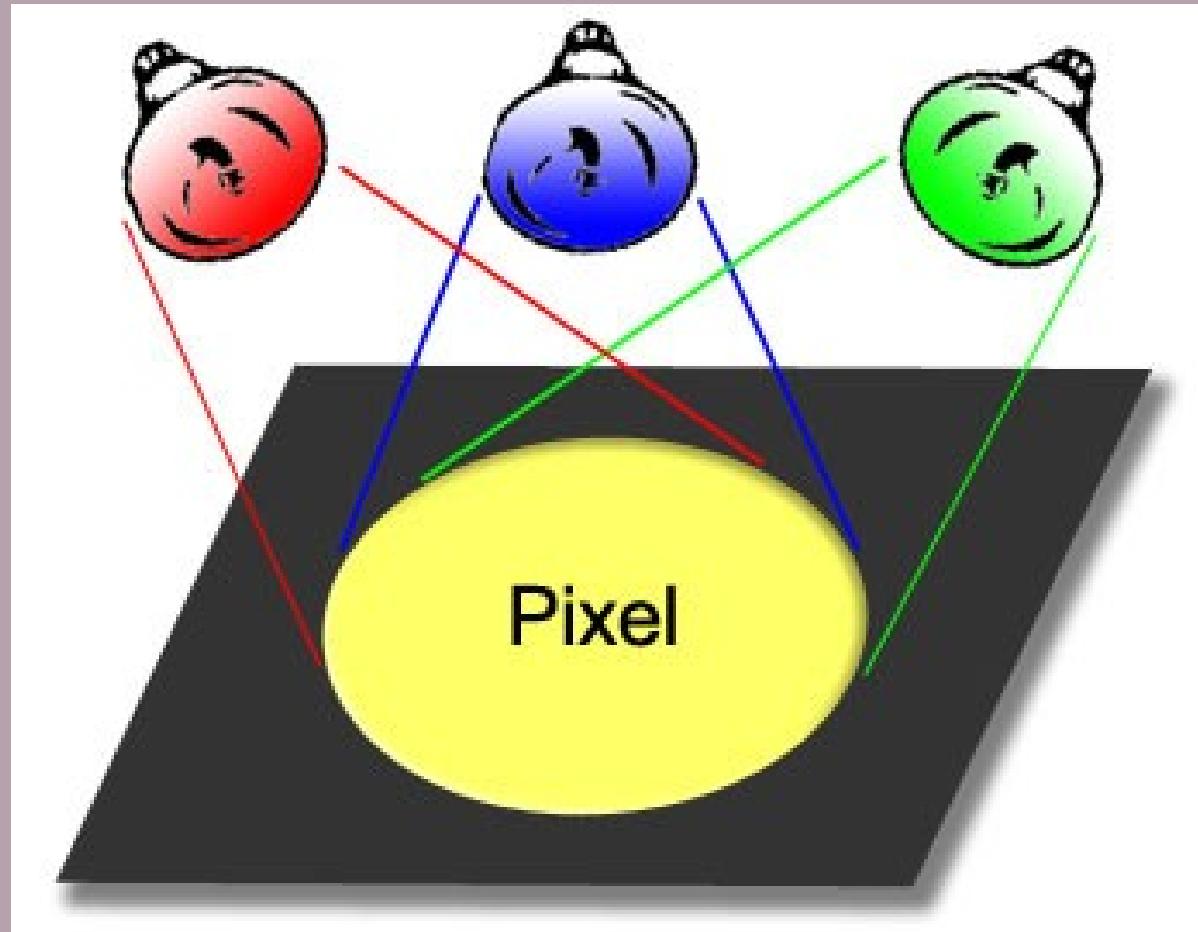
Každý pixel je opísaný tromi číslami RGB - červenou, zelenou, modrou.

Tieto čísla predstavujú červenú, zelenú a modrú zložku každého pixelu

Ich zmiešaním je možné vytvoriť množstvo farieb podobného rozsahu ako ľudské videnie

*Model RGB definuje farbu jedného pixelu popisom farby 3 svetiel, ktoré, ak by svietili na rovnakom mieste, tak by zmiešali farbu*

# RGB



# Farebná hĺbka

**Farba každého pixelu v reprezentácii RGB je kódovaná 3 číslami**, ktoré vyjadrujú jas jeho červenej, zelenej a modrej zložky.

Tu je potrebné urobiť určité zjednodušenia, pretože ukladanie reálneho čísla, napríklad na 10 desatinných miest, je veľmi náročné na dáta.

Preto je **každá farba pixelov jednoducho kódovaná buď 1 bajtom alebo 2 bajtmi**.

**Bajt** je základná jednotka počítačových informácií a skladá sa z 8, 16, 32, 64 bitov,

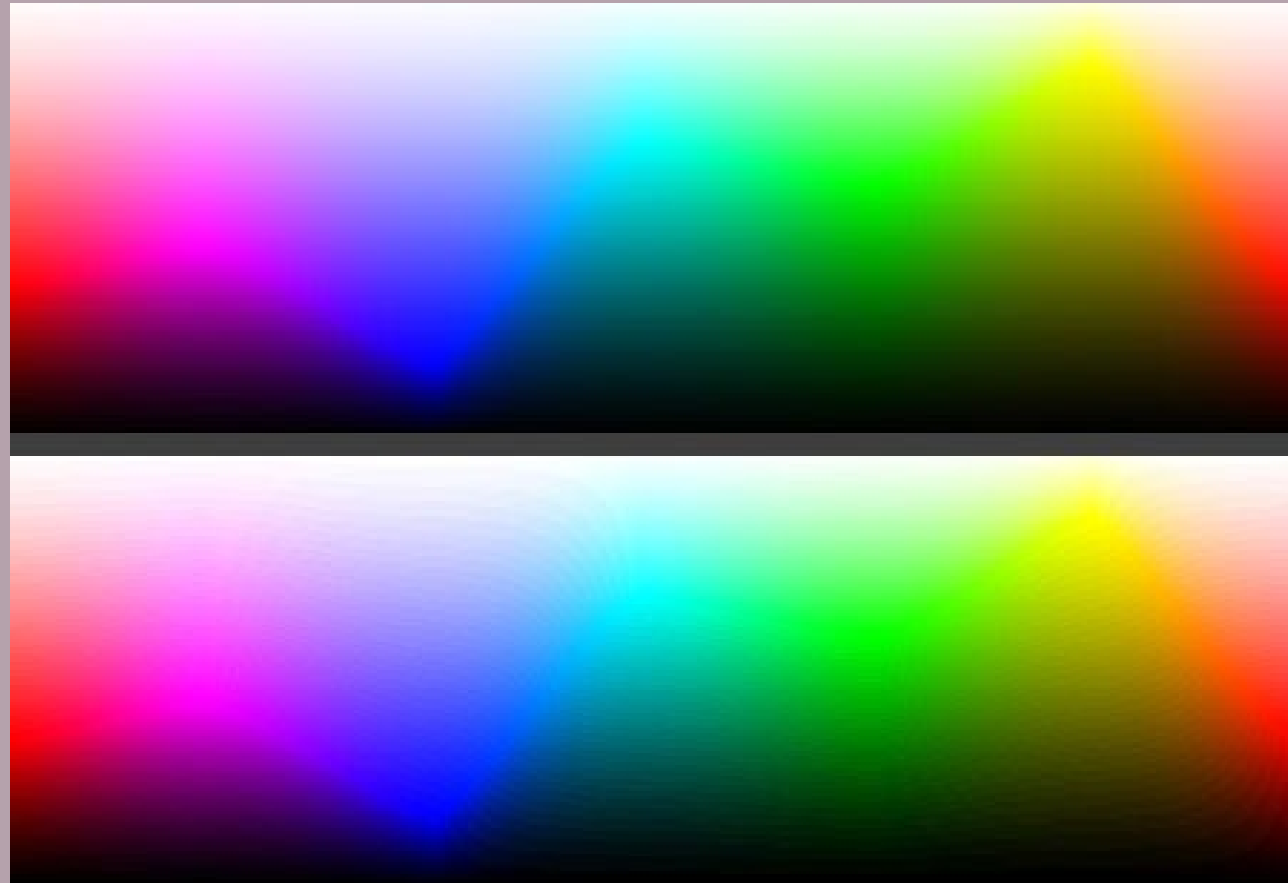
**Bit** je elementárna bunka schopná niesť iba informácie "0" alebo "1".

Napríklad kapacita pamäťovej karty je daná počtom bajtov, ktoré je karta schopná niesť.

*Farebné rozlíšenie určuje celkový počet farieb, ktoré môžu existovať v obrázku.*

*Keď je nízky počet možných farieb (obrázok) vedie to k neplynulým prechodom, a tak sa nedostatok farieb najčastejšie objavuje na oblohe, na jednotnom pozadí atď.*

# Farebná hĺbka

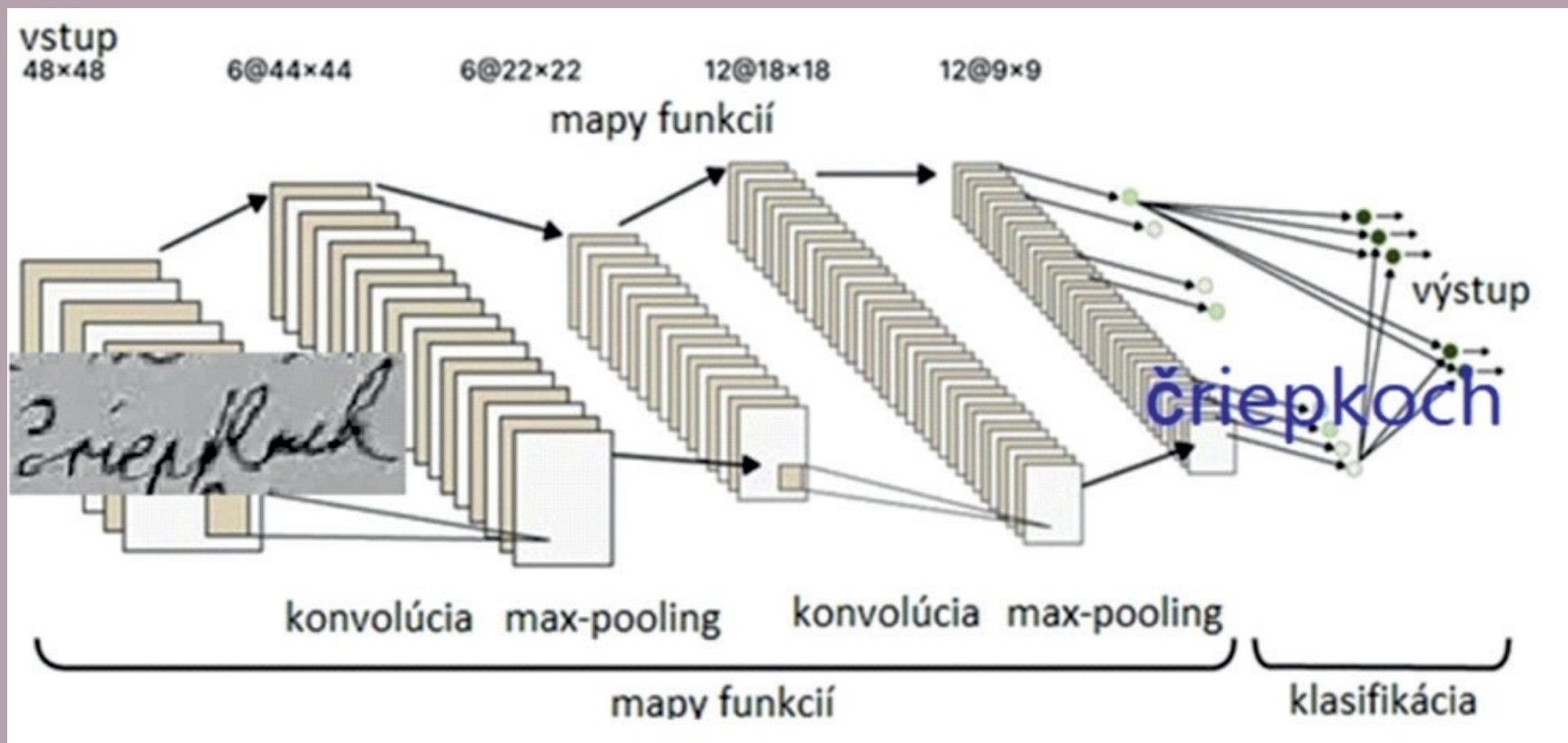


# Farebná hĺbka pixel, bity, bajty

- Pre čiernobiele obrázky - každý pixel sa zvyčajne skladá z 8 bitov (1 bajt)
- Pre farebné obrázky, ak sa používa farebná schéma RGB (červená, zelená, modrá), tak pre každú farbu sa použije jeden bajt , čiže (3x8), teda 24 bitov (3 bajty) na pixel.
- V tomto prípade hovoríme o **farebnej hĺbke 8 bitov na kanál alebo 3x8 = 24 bitov na pixel** (24 bpp = bit na pixel).
- V odbornej praxi to však často nestačí, preto sa používa vyššia farebná hĺbka, t.j. **16 bitov (2 bajty) na kanál, t.j. 3x16=48 bitov na pixel (bpp).**



# Pixely v transkripcii



# Príklad štruktúry konvolučnej siete

- Historické staré a vzácne tlače, strojopisy a hlavne rukopisy spravidla nie je možné uspokojivo transkribovať. Prichádza na pomoc umelá inteligencia
- V snahách sprístupniť historické písomné dedičstvo sa koncentruje pozornosť výskumníkov na *transkripciu* a strojové učenie s použitím *konvolučných neurónových sietí*
- Ide o proces, v ktorom sa nasnímaný *obrázok* mení na *text*.

# Konvolúcia v transkripcii

Obrázok ilustruje vysvetlenie procesu fungovania konvolučnej siete

Aby mohol byť obrázok spracovaný počítačom napr. v transkripcii, musia sa obrazové informácie, teda **pixely** previesť do **číselnej formy**.

Na vstupe (Input) procesu rozpoznávania nejakého predmetu, napríklad písma, tváre, zvieratá, auta sú **pixely** obrázka.

Vstupný obraz má rozlíšenie 48 x'48. Z neho sa vyberajú pixely.

Potom sa postupne použijú množiny filtrov (Mapy funkcií - Feature Maps) na extrahovanie lokálnych obrazových príznakov prostredníctvom operácie konvolúcia (convolution), čo je matematická operácia.

Filtre sú v podstate masky, ktoré sú „prehodené“ cez obrázok, aby sa zistilo, či im niečo vyhovuje.

Konečný súbor funkcií sa potom vloží do husto pripojenej siete, odkiaľ pochádza skutočná univerzálna predikčná sila tohto algoritmu (classification).

Takáto sieť sa môže naučiť aproximovať akúkoľvek primerane dobre vycvičenú funkciu s ľubovoľnou presnosťou, pokiaľ je sieť dostatočne veľká.

V prípade transkripcie rukopisov to prakticky znamená, že na cvičenie modelu je potrebný veľký súbor cvičných dát.

Otázka je, aký veľký by ten súbor mal byť, aby výsledky transkripcie boli čo najpresnejšie.

Na základe rozdielu medzi predpoveďou modelu a „ground truth“ sa parametre vo vnútri siete aktualizujú iteratívne.

Po dokončení cvičenia je možné rozpoznať nové obrázky pri pohľade na výstup, ktorý ukazuje najpresnejšiu aktiváciu.