

# Analýza obrazu v materiálových vědách

Dalibor Martišek

FSI VUT v Brně, 2017

# Obsah

1 Obraz .....	5
1.1 Zobrazení v euklidovském a projektivním prostoru .....	5
1.2 Obraz a jeho digitalizace.....	9
1.3 Atributy pixelu .....	12
1.4 Barva .....	14
1.5 Barevné systémy (modely) .....	17
1.6 Barevné prostory .....	20
1.7 Formáty rastrových grafických souborů .....	23
2 Obraz jako statistický soubor .....	25
2.1 Rovnoměrné a normální rozdělení.....	25
2.2 Základní výběrové charakteristiky.....	26
2.3 Histogram obrazu.....	27
2.4 Úprava jasu a kontrastu.....	29
2.5 Ekvalizace histogramu .....	34
3 Aritmetické a logické operace s obrazy .....	35
3.1 Aritmetické operace .....	35
3.2 Základy vícehodnotové logiky .....	40
3.3 Logické operace s obrazy .....	41
4 Geometrické transformace obrazu .....	43
4.1 Zobrazení v projektivní rovině .....	46
4.2 Transformace diskrétního obrazu .....	68
5 Fourierova transformace .....	48
5.1 Fourierova řada funkcí reálné proměnné .....	48
5.2 Fourierova transformace .....	51
5.3 Konvoluce .....	57
5.4 Okrajový efekt.....	59
5.5 Registrace obrazů .....	62
6 Obrazové filtry .....	69
6.1 Šum .....	69
6.2 Lineární filtry .....	71
6.3 Nelineární filtry.....	75
6.4 Adaptivní filtry.....	78
7. Měření objektů.....	78
7.1 Prahování.....	78
7.2 Geometrické momenty objektu .....	79
7.3 Fraktální dimenze objektů .....	82

# 1 Obraz

Analýza obrazu je relativně nová matematická disciplína, která, jak již její název napovídá, se zabývá metodami zpracování obrazu na výstupním zařízení počítače.

Ze středoškolské matematiky známe pojem zobrazení z množiny  $A$  do množiny  $B$  jako relaci  $R$ , kde pro každé  $a \in A$  existuje nejvýše jedno  $b \in B$  takové, že  $[a; b] \in R$ . Prvek  $a$  v této dvojici nazýváme vzorem prvku  $b$ , prvek  $b$  naopak obrazem prvku  $a$ . Množinu  $O(A)$  všech obrazů prvků množiny  $A$  nazýváme obrazem množiny  $A$ .

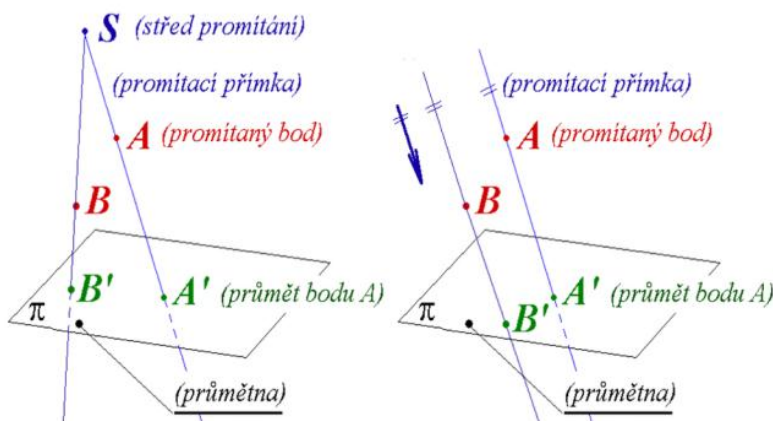
Tento popis vypadá matematicky velmi formálně, v praktickém životě se však s touto situací setkáváme neustále. Například každý náš zrakový vjem je zobrazením, a to dokonce zobrazením složeným. První „fází“ vzniku zrakového vjemu je zobrazení vnějšího světa na sítnici očí, druhou „fází“ pak přenos těchto dvou obrazů do mozku a jejich zpracování zrakovým centrem.

Každý bod prostoru, který nás obklopuje, je určen svojí polohou a barvou. Polohu bodu v prostoru lze matematicky popsat třemi prostorovými souřadnicemi. Pokud jde o barvu, pak zdravé lidské oko je schopno rozeznat asi pět miliónů různých odstínů. Sítnice lidského oka obsahuje dva druhy fotoreceptorických buněk - tyčinky umožňující vnímat intenzitu osvětlení a čípky umožňující vnímání barev. Čípky jsou trojího druhu – liší se látkou citlivou na různé části spektra viditelného světla – červenou, zelenou a modrou. Matematicky tak můžeme barvu (při jistém zjednodušení) modelovat rovněž třemi čísly – barevnými souřadnicemi.

Podívejme se podrobněji nejdříve na zpracování prostorové informace.

## 1.1 Zobrazení v euklidovském a projektivním prostoru

Svět, který nás obklopuje, je trojrozměrný, obraz na výstupním zařízení dvojrozměrný. Zobrazení, které převádí trojrozměrný obraz na dvojrozměrný nazýváme promítání. Promítání je v reálném světě zprostředkováno (alespoň ve většině případů) přímočaře se šířícím světlem, matematicky ho proto modelujeme pomocí promítacích přímk dle obr. 1.1.1.



Obr. 1.1.1: Promítání

a) středové

b) rovnoběžné

Sem zadejte rovnici. Geometrie, kterou jste se zabývali na střední škole, studovala útvary v tzv. euklidovském prostoru. Tato geometrie vychází z několika základních tvrzení, která se nedokazují (tzv. axiomů). Jedním z nejznámějších je axiom o rovnoběžkách:

**Ar: Bodem  $A$ , který neleží na přímce  $p$ , lze v rovině  $\overleftrightarrow{pA}$  vést právě jednu přímku, která nemá s  $p$  žádný společný bod.**

Tento axiom však úvahy o promítání značně komplikuje, neboť při jeho platnosti je třeba nahlížet na rovnoběžné a středové promítání jako na dva různé případy, což se odráží i v nejednotném matematickém popisu.

Pro matematický popis promítání se ukazuje výhodnější přijmout bez důkazu jiné tvrzení (projektivní axiom):

**Ap: Každé dvě přímky ležící v dané rovině, mají společný bod.**

Současně však budeme chtít (alespoň v jistém smyslu) zachovat v platnosti axiom o rovnoběžkách. To je zdánlivě neřešitelný úkol, řešení je však poměrně jednoduché: Kromě „tradičních“ bodů, které známe ze středoškolské geometrie a které budeme nazývat **vlastní body**, budeme uvažovat ještě body „netradiční“, které prohlásíme za **body nevlastní**. Axiom o rovnoběžkách pak stačí jen mírně upřesnit – o rovnoběžkách prohlásíme, že nemají společný žádný **vlastní bod**. Z Ap pak plyne, že společný bod rovnoběžek je bodem nevlastním.

Vlastní bod si tradičně představujeme jako bezrozměrnou tečku, nevlastní bod si lze představit jako množinu všech vektorů téhož směru. Analyticky lze body v projektivním prostoru definovat jako množiny uspořádaných čtveřic tvaru

$$P = k \cdot (p_1; p_2; p_3; \omega_p); k \neq 0,$$

přičemž  $\omega_p \neq 0$  právě tehdy, když bod  $P$  je vlastní, a  $\omega_p = 0$  právě tehdy, když bod  $P$  je nevlastní. Nevlastní bod  $P$  značíme podrobněji  $_{\infty}P$ .

Obě promítání z obr. 1.1.1 můžeme považovat za promítání středové- střed promítání na obr. 1.1.1a) je vlastní, na obr. 1.1.1b) pak nevlastní.

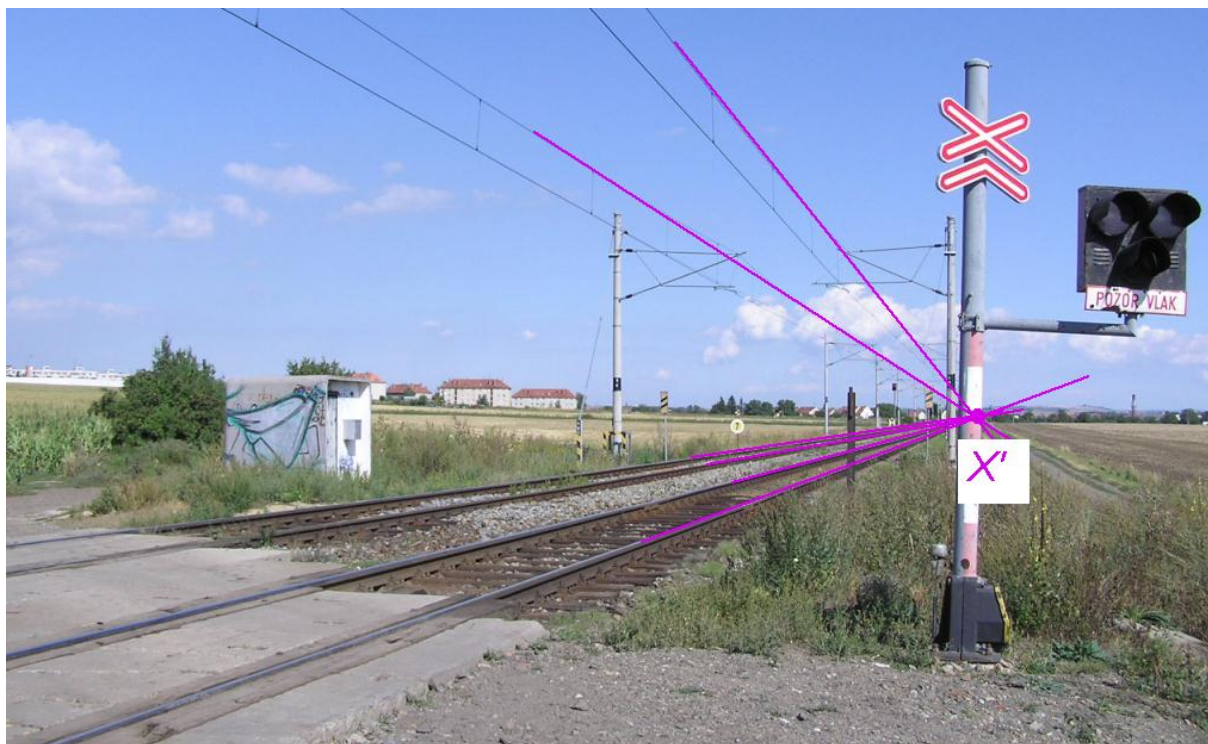
Každé zobrazení, v němž je obrazem přímky opět přímka, anebo bod (tedy i promítání) lze zapsat ve tvaru

$$\begin{pmatrix} x' \\ y' \\ z' \\ \omega' \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ \omega \end{pmatrix} \Rightarrow \mathbf{x}' = \mathbf{Ax} \quad (1.1.1)$$

Některá z těchto zobrazení zachovávají vlastní a nevlastní body, tj. obrazem vlastního (nevlastního) bodu je opět vlastní (nevlastní) bod. Pro tato zobrazení musí tedy platit  $(\omega' = 0) \Leftrightarrow (\omega = 0)$  a rovnice (1.1.1) je tedy tvaru

$$\begin{pmatrix} x' \\ y' \\ z' \\ \omega' \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ 0 & 0 & 0 & a_{44} \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ \omega \end{pmatrix}; \quad a_{44} \neq 0 \quad (1.1.2)$$

Rovnice (1.1.2) popisuje souměrnosti, posunutí, otáčení, změnu měřítka, zkosení a promítání s nevlastním středem (rovnoběžné promítání). Rovnici (1.1.2) naopak nesplňuje např. středové promítání s vlastním středem, kde obrazem vlastního bodu může být bod nevlastní a naopak. Fotografie na obr. 1.1.2 vznikla právě středovým promítáním s vlastním středem, protože vlastní bod  $X'$  na fotografii je obrazem nevlastního bodu  $_{\infty}X$  (totiž společného bodu rovnoběžných kolejí a trolejí).



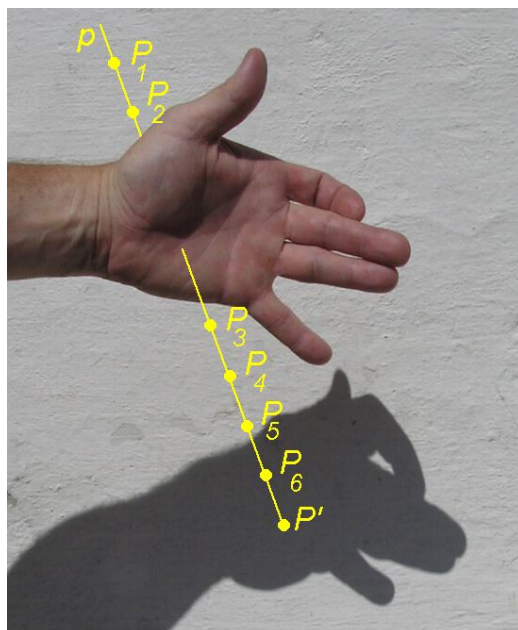
**Obr. 1.1.2:** Středový průmět prostoru do roviny

Středem promítání je v tomto případě ohnisko objektivu fotoaparátu. Toto zobrazení lze popsat rovnicí (1.1.1), kde alespoň jedno z čísel  $a_{41}; a_{42}; a_{43}$  je různé od nuly.

Některá ze zobrazení popsaná rovnicemi (1.1.1); (1.1.2) jsou prostá (tj. obrazy každých dvou různých bodů jsou navzájem různé) a některá prostá nejsou. Matice  $\mathbf{A}$  je v případě prostého zobrazení regulární (to platí např. pro souměrnosti, otočení, změnu měřítka, zkosení), v případě, že zobrazení prosté není, je matice  $\mathbf{A}$  singulární. Příkladem zobrazení, které není prosté, je promítání na rovinu (ať už středové, anebo rovnoběžné). Všechny body každé promítací přímky se totiž promítnou do jediného bodu. Například

$$\begin{pmatrix} x' \\ y' \\ z' \\ \omega' \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ 0 & 0 & 0 & 0 \\ a_{41} & a_{42} & a_{43} & a_{44} \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ \omega \end{pmatrix} \quad (1.1.3)$$

je rovnice promítání na rovinu  $z' = 0$  (viz obr. 1.1.3).



**Obr. 1.1.3:** Promítání prostoru na rovinu. Obrazem všech bodů  $P_1; P_2; \dots$ , které leží na téže promítací přímce  $p$  (slunečním paprsku), je jediný bod  $P'$ .

Jak již bylo řečeno, soustavami lineárních rovnic (1.1.1), (1.1.2), (1.1.3) jsou popsána zobrazení, v nichž obrazem přímky je opět přímka, anebo bod. V praxi se často setkáváme s nejrůznějšími geometrickými vadami optických soustav, i se zobrazeními, jejichž projevem je zobrazení v němž obrazem přímky již není přímka. Nejčastěji se jedná o tzv. soudkovité (obr. 1.1.4a) a polštářovité (obr. 1.1.4b) zkreslení.



**Obr. 1.1.4:** Geometrické vady čoček

a) soudkovité zkreslení

b) polštářovité zkreslení

Tato zobrazení jsou popsána soustavou nelineárních rovnic tvaru

$$x' = f(x; y; z)$$

$$y' = g(x; y; z)$$

Metody nápravy těchto deformací přesahují rozsah tohoto textu a nebudeme se jimi zabývat.

## 1. 2 Obraz a jeho digitalizace

V následujícím textu se budeme zabývat analýzou (dvojměrných) průmětů okolního světa do roviny. Bez újmy na obecnosti můžeme předpokládat, že touto rovinou je rovina  $z=0$ , zobrazení je tedy dáno rovnicí (1.1.3). Toto zobrazení přináší do každého bodu  $[x'; y']$  průmětny informaci o bodech  $[x; y; z]$  v prostoru, které leží na příslušné promítací přímce. Pro začátek předpokládejme, že tuto informaci lze vyjádřit jediným číslem.

**Obrazem či obrazovou funkcí budeme v tomto případě rozumět skalární funkci**

$$v = f(x'; y') \quad (1.2.1)$$

Podle předchozího textu může mít obrazová funkce spojitý definiční obor, na kterém může být spojitá. Aby tomu však tak skutečně bylo, musíme každý bod roviny modelovat jako bezrozměrný objekt, který je schopen nést každou reálnou hodnotu  $v \in \langle a; b \rangle$ . Elektronické zařízení je však fyzický objekt a nedokáže zobrazit ani konečně malý bod, ani nekonečně mnoho informací (které funkce spojitá na intervalu  $\langle a; b \rangle$  teoreticky představuje).

Pro uchování a reprodukci obrazové informace v elektronické podobě je tedy třeba převést oba spojitě obory obrazové funkce na obory diskrétní – hovoříme o **digitalizaci obrazu**. Digitalizace obrazu se skládá ze dvou částí - diskretizace definičního oboru, kterou nazýváme **vzorkováním**, a diskretizaci oboru hodnot, kterou nazýváme **kvantování**.

Výsledkem vzorkování je obrazová matice **I** (Image) typu  $W \times H$  (Width; Height) skalárních hodnot  $(v_{ij})$ . Každý prvek této obrazové matice nazýváme **pixel** (z anglického „picture element“). Pixel modelujeme nejčastěji jako čtverec, méně často jako obdélník, lze použít i hexagonální (šestiúhelníkové) uspořádání. V dalším textu budeme pixely modelovat jako čtverce. Kvantování pak každému pixelu přiřadí celočíselnou hodnotu nějaké veličiny (osvětlení, teplota, barevná složka...). Předpokládejme prozatím, že touto veličinou je intenzita bílého světla odražená od plochy pixelu, kterou vnímáme jako světlejší či tmavší odstín šedé barvy.

**Vzorkování:** Vzorkováním rozumíme reprezentaci spojitě obrazové funkce  $f(x'; y')$  maticí diskrétních hodnot. Jedním z rozhodujících faktorů kvality digitalizovaného obrazu je fyzická velikost jeho pixelů, anebo tzv. rozlišení obrazu. Tento pojem je chápán většinou ve dvou významech: jednak je to typ obrazové matice (tj. její „rozměry“ v pixelech) a jednak je to počet pixelů na palec, který se udává v jednotkách DPI (dots per inch, tedy počet bodů na palec). Na obr. 1.2.1 vidíme fotografii s rozlišením (v prvním slova smyslu) 1500 x 2100; 500 x 700, 100 x 140; 30 x 42.

Obrazy vnímáme zrakem většinou ze vzdálenosti cca 20 cm a větší. Lidské oko je za ideálních světelných podmínek a při dostatečně velkém kontrastu schopno rozlišit dva body, jestliže je vidí pod zorným úhlem minimálně 1'. Ze vzdálenosti 20 cm jsme tak schopni rozlišit dva body vzdálené od sebe asi pět setin milimetru. Tato hodnota je tedy zároveň maximální velikost pixelu, kterou by měl kvalitní obraz mít. Jednoduchým přepočtem zjistíme, že kvalitní obraz by měl mít rozlišení minimálně 500 DPI (1 palec je 2,54 cm).

Maximálního kontrastu, který jsme předpokládali výše, však dosahuje jen ostrý černý text na čistě bílém pozadí, tenké křivky na přesných technických výkresech apod. Pro jiné obrazy (např. barevné fotografie) většinou postačuje rozlišení 300 DPI.





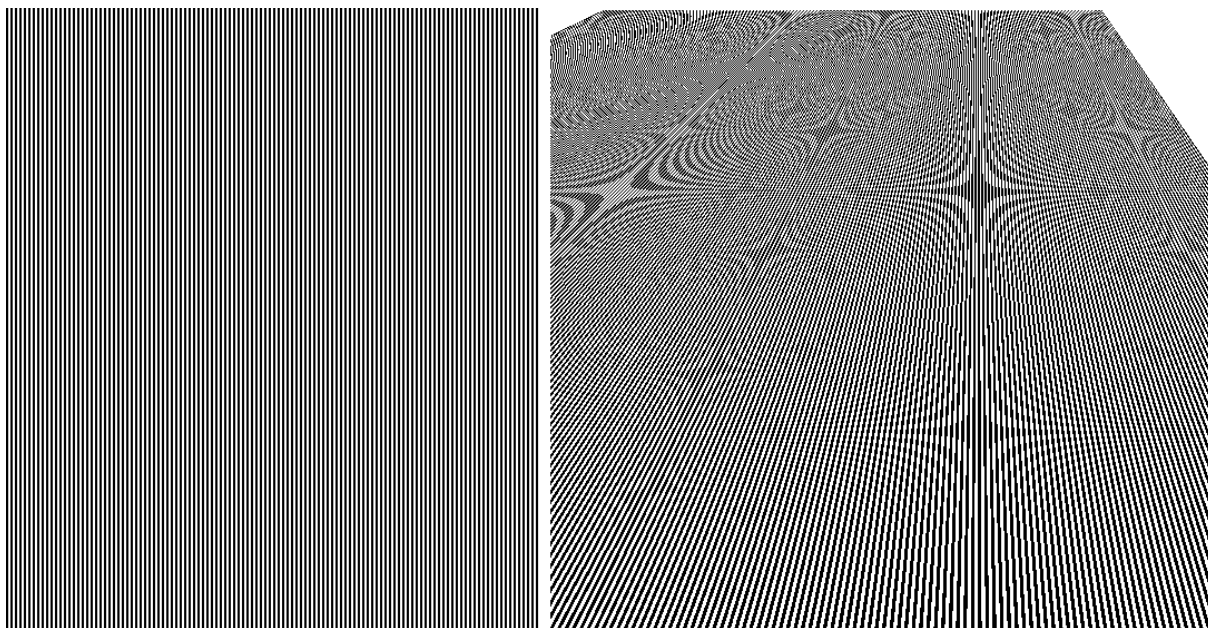
**Obr. 1.2.1:** Snímek s postupně klesajícím rozlišením – 1600x1200; 400x300, 200x150; 100x75. Nedostatečné vzorkování degraduje nejdříve drobné detaily (zde fasády budov) a postupuje k velkým plochám (zde trávník a obloha).

Reálný obraz představuje spojitý signál, z něhož odebíráme vzorky ve vzdálenostech odpovídajících velikosti pixelů. Na kvalitu digitalizovaného obrazu má tedy vliv nejen jednotka DPI, ale také velikost nejmenších detailů snímaného reálného objektu. Tuto otázku řeší Nyquistův – Shannonův teorém:

**Nyquistův – Shannonův teorém:** Přesná rekonstrukce spojitého a frekvenčně omezeného signálu z jeho vzorků je možná tehdy, když vzorkovací frekvence je vyšší než dvojnásobek nejvyšší harmonické složky vzorkovaného signálu.

Prakticky to znamená následující: **Digitální obraz může být kvalitní jen tehdy, je-li jeho nejmenší detail nejméně dvakrát větší než pixel.** Toto pravidlo lze ilustrovat následovně: na obr. 1.2.2 vlevo máme pravidelně pruhovanou předlohu zobrazenou v kolmém promítání, kde jednotlivé pruhy mají šířku několika pixelů. Tato předloha je zobrazena kvalitně. Na obr. 1.2.2 vpravo je tatáž předloha ve středovém promítání, kde se původně rovnoběžné linie sbíhají a pruhy se zužují. Jakmile se zobrazovaná šířka zmenší pod hranici dvou pixelů, začínají se na obraze objevovat vzory, které na reálném objektu nejsou. S tímto jevem se setkáváme poměrně často a nazýváme ho **alias**. Alias můžeme považovat za chybu vzorkování.





**Obr. 1.2.2:** Předloha skládající se z vodorovných pruhů je zobrazena v rovnoběžném promítání (vlevo) a ve středovém promítání (vpravo). Jakmile se zdánlivá šířka pruhů zmenší pod dva pixely, začíná se projevovat pseudoinformace, která na skutečné předloze není. Tento jev se nazývá alias.

**Kvantování:** Vzorkováním obdržíme konečný počet funkčních hodnot obrazové funkce (počet těchto hodnot je dán počtem prvků obrazové matice, tedy  $W \times H$ ). Obor těchto hodnot však může mít stále  $W \times H$  zcela libovolných prvků. Výstupní zařízení však může reprodukovat jen ekvidistanční hodnoty – kvantizační úrovně o stejných vzdálenostech, tak jak to ilustruje obr. 1.2.3 na příkladu jasu.



**Obr. 1.2.3:** Spojitý obor hodnot jasu (nahore) je kvantován na osm úrovní (dole). Striktně vzato ani horní obor není spojitý - je kvantován na 256 úrovní. Při takto jemném kvantování však lidské oko není schopno rozpoznat rozdíly jasů sousedních úrovní.

Stanovení počtu kvantizačních úrovní se řídí citlivostí lidského zraku. Pro většinu praktických situací stačí, je-li poměr jasu bílé k černé roven stu. Při tomto poměru je zdravé lidské oko schopno rozlišovat asi 230 úrovní jasu. Pro většinu aplikací tedy postačuje osmibitové kódování úrovně jasu (tj. 256 úrovní). Při takto jemném kvantování již oko považuje celou škálu za spojitou (viz obr. 1.2.3 nahore).

Podobně jako nedostatečné vzorkování (viz obr. 1.2.1) může obraz značně degradovat rovněž nedostatečné kvantování – viz obr. 1.2.4.



**Obr. 1.2.4:** Snímek s postupně klesajícím počtem kvantizačních úrovní (zde úrovně jasu) – 256, 16, 8, 4. Nedostatečné kvantování degraduje nejdříve velké plochy (zde trávník a obloha) a postupuje k drobným detailům (zde fasády budov).

### 1. 3 Atributy pixelu

**Vlastní atributy pixelu:** Na začátku této kapitoly jsme předpokládali, že pixel nese informaci, kterou lze vyjádřit jen jedním číslem (skalární informací). Tato informace je však velmi často složitější. V obecném slova smyslu můžeme pixel považovat za plošný element, na kterém je současně uložena informace o jasu, intenzitě elektrického pole, vlnové délce světla atd. Každou informaci, kterou můžeme takto uložit, nazýváme atribut pixelu, přesněji řečeno **vlastní atribut pixelu**. Vlastní atribut pixelu je atribut, který závisí jen na uvažovaném pixelu. Situaci, kdy je těchto vlastních atributů obecně  $n$ , lze popsat vektorovou funkcí

$$\mathbf{v}_{ij} = (v_1; v_2; \dots; v_n)_{ij} = \mathbf{f}(i; j) \quad (1.3.1)$$

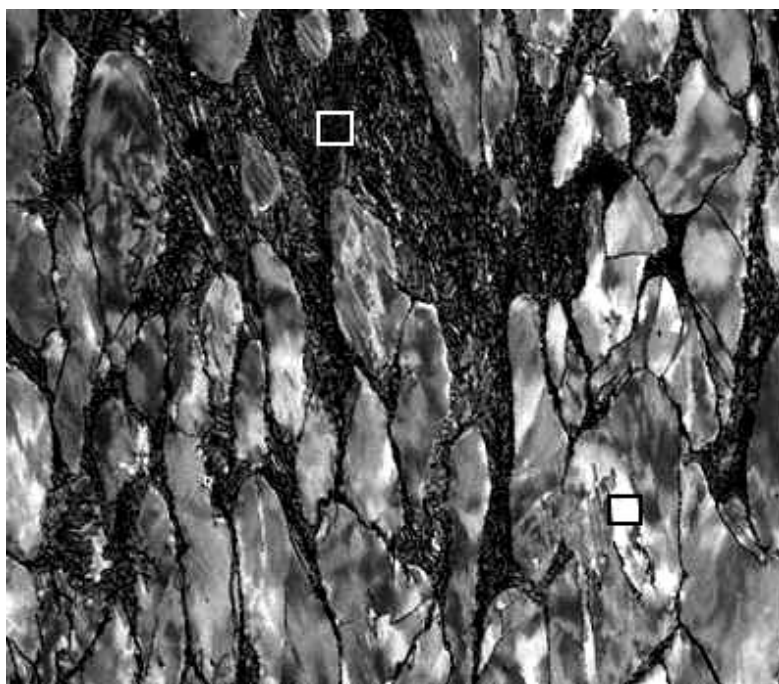
Prakticky se však většinou postupuje tak, že se tato vektorová funkce nahradí  $n$ -ticí skalárních funkcí

$$\begin{aligned} v_{1;ij} &= f_1(i; j) \\ v_{2;ij} &= f_2(i; j) \\ &\dots\dots\dots \\ v_{n;ij} &= f_n(i; j) \end{aligned} \quad (1.3.2)$$

Říkáme, že vektorovou funkci popisujeme po složkách (složkami jsou zde jednotlivé vlastní atributy pixelu). Podobně jako jas může každý jiný atribut pocházet ze spojitě informace a v tom případě musí být tato informace rovněž výše uvedeným způsobem digitalizována. Místo jedné obrazové matice **I** tak obecně dostáváme uspořádanou n-tici (**I**<sub>1</sub>;**I**<sub>2</sub>;...;**I**<sub>n</sub>) matic, z nichž každá popisuje jeden atribut pixelů daného obrazu.

Nese-li pixel skalární informaci, je touto informací nejčastěji jas. Je to skalární hodnota omezená dynamickým rozsahem obrazu a počtem kvantizačních úrovní tak, jak jsme o tom mluvili v předchozí kapitole. Je-li tato informace vektorová, je touto informací nejčastěji informace o barvě - **intenzita barevných složek, jas a světlost barvy, saturace a tón barvy**. Tyto atributy nejsou nezávislé, hodnotu některých z nich lze odvodit z ostatních tak, jak o tom pojednáme v následujících dvou kapitolách.

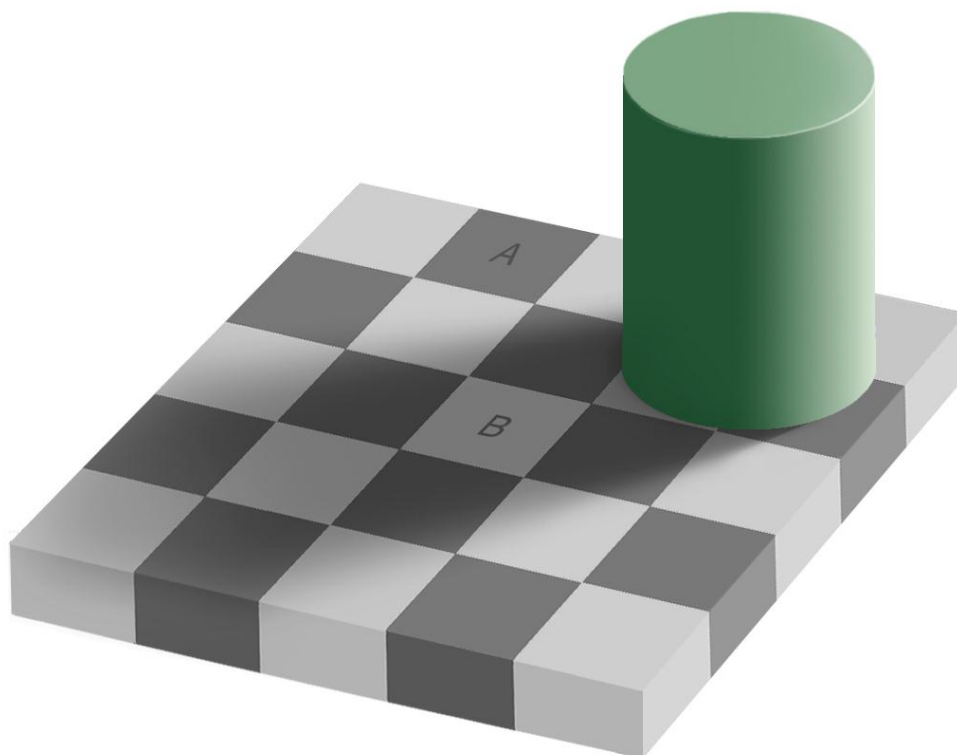
**Nevlastní atributy pixelu:** Tyto atributy jsou závislé na vlastních attributech sousedních pixelů. Jedná se zejména o **střední hodnotu, rozptyl a velikost gradientu (kontrast)**. O těchto attributech pojednáme v dalším textu. Zde si podrobněji všimněme velmi důležitého nevlastního atributu pixelu a tím je jeho kontrast. Pojmu kontrast většinou intuitivně velmi dobře rozumí. Je to jednoduše řečeno rozdíl mezi světlem a tmou. Kontrast je možné vzhledem k vlastnostem lidského oka definovat různě. Např. barevný kontrast je postaven na kontrastu barev a maximálního kontrastu lze dosáhnout použitím tzv. doplňkových barev (budeme o nich mluvit v následující kapitole). Asi častější chápání slova kontrast je ale postaveno na jasovém kontrastu - čili rozdílu jasů různých ploch v jednom obrazu (globální kontrast), anebo některé jeho části (lokální kontrast). Na obr 1.2.5. je mikrosnímek oceli X210Cr12 s vysokým globálním kontrastem, ale lokální kontrast v obou vyznačených oblastech je velmi malý.



**Obr. 1.3.1:** Mikrosnímek oceli X210Cr12 s vysokým globálním kontrastem a malým lokálním kontrastem ve vyznačených oblastech.

Na závěr této kapitoly poznamenejme, že kontrast vnímaný lidským okem závisí i na celkových světelných podmínkách – tentýž rozdíl jasů vnímá lidské oko citlivěji za šera než za plného osvětlení. Konečně – lidské je citlivé na lokální kontrast ale velmi těžko posuzuje

absolutní jas. Ilustruje to obr. 1.2.6. kde máme posoudit jas polí šachovnice označených písmeny A, B. Dojem, že pole A je tmavší než pole B, je mylný. Oko se jen nechalo zmást lokálním kontrastem na hranicích světlejších a tmavých polí a absolutní jas vyhodnotilo špatně. Obě pole jsou stejně jasná.



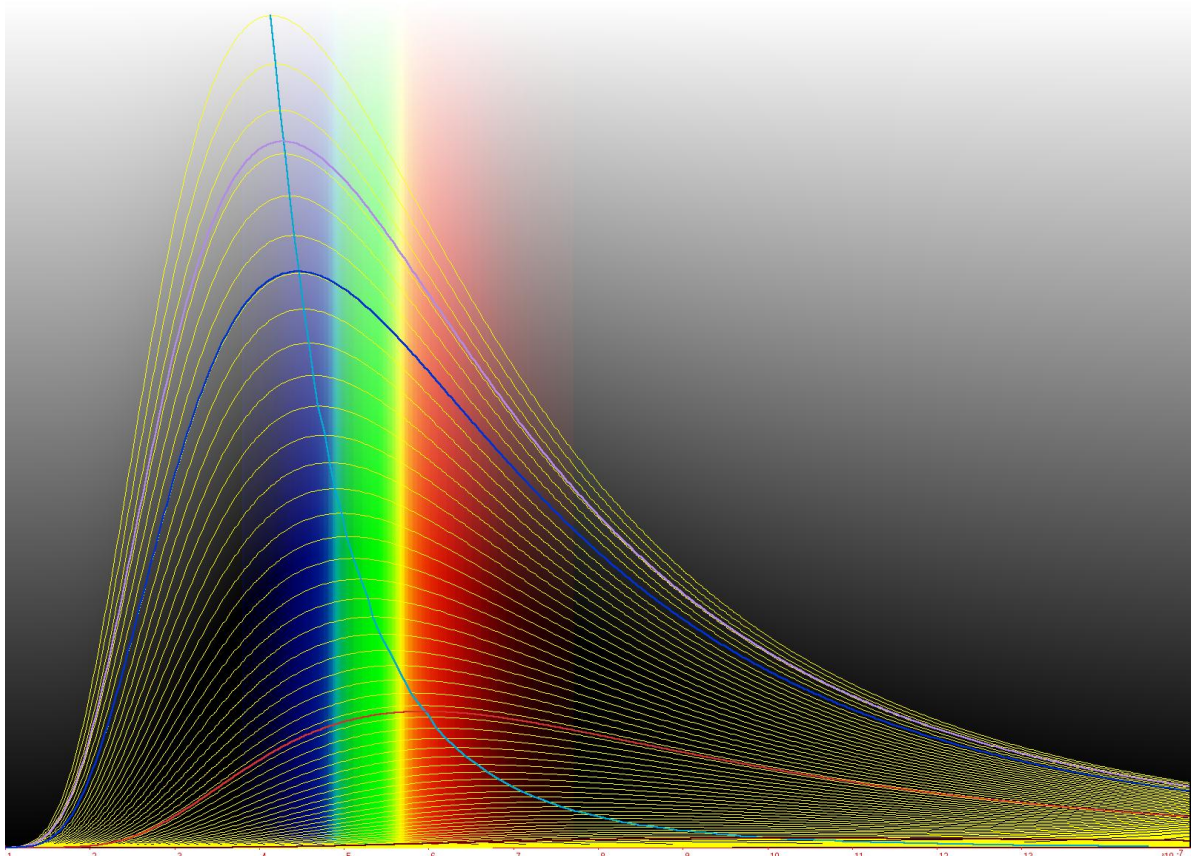
**Obr. 1.3.2:** Ukázka chybné vyhodnocení absolutního jasu lidským okem: pole A i B jsou stejně jasná.

## 1. 4 Barva

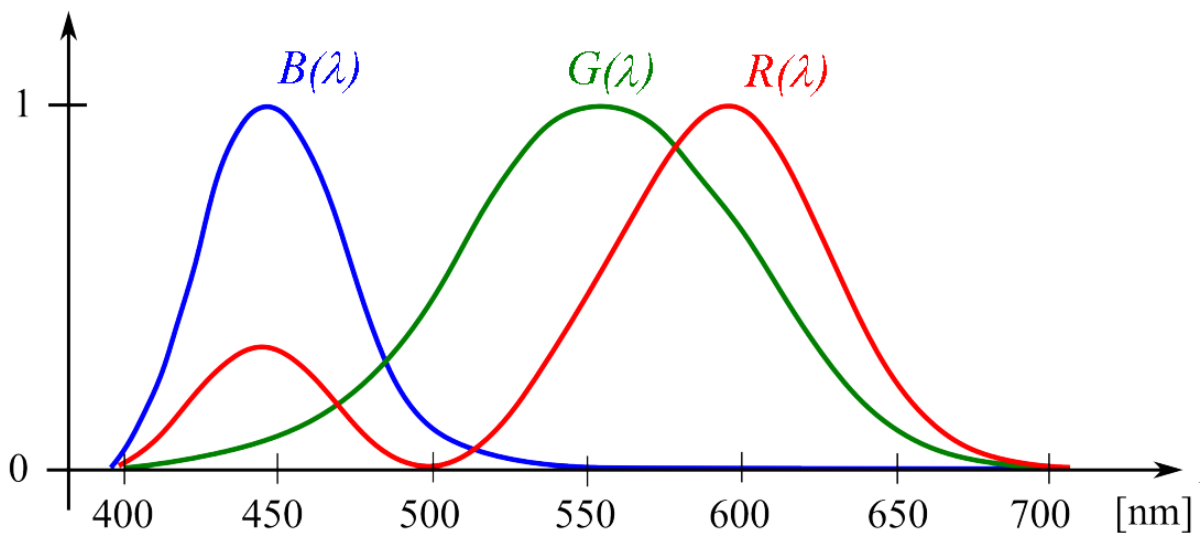
Lidské oko vnímá elektromagnetické vlnění o vlnových délkách 400 – 720 nm. Zdrojem běžného denního světla je Slunce, které můžeme považovat za absolutně černé těleso. Intenzita záření takového tělesa závisí na jeho teplotě (vyšší teplota znamená vyšší intenzitu a naopak), závislost intenzity záření konkrétního tělesa na jednotlivých vlnových délkách -  $I_s(\lambda)$ , charakteristiku záření - určuje Planckův zákon. Křivky, které tento zákon určuje pro tělesa různých teplot, si můžeme prohlédnout na obr. 1.3.1.

Nejintenzivnější záření Slunce má podle tohoto zákona vlnovou délku asi 550 nm, což je (ne náhodou) uprostřed rozsahu viditelného světla. Kratší vlnové délky vnímáme jako modrou až fialovou, delší pak jako žlutou a poté červenou. Jednotlivé barvy vznikají tak, že ze slunečního světla je nějakým způsobem část záření odfiltrována. Barva, kterou vnímáme, je pak dána jednak charakteristikou  $I_F(\lambda)$  tohoto přefiltrovaného záření a jednak způsobem, jakým lidské oko takové záření vnímá, přesněji řečeno až na tom, jak tento zrakový vjem zpracuje mozek.

**Vnímání světla lidským okem:** Na sítnici oka jsou dva druhy receptorů – tyčinky, které reagují na intenzitu světla, a čípky, které nám umožňují rozlišovat vlnové délky vnímaného záření, tedy barvy. Čípky se vzájemně neliší stavbou, ale obsahem tzv. fotopsinů – chemických látek různě citlivých na různé vlnové délky viditelného světla. Čípek je tak specializován buď na červenou, anebo zelenou resp. modrou barvu. Citlivost čípků na jednotlivé vlnové délky ilustruje obr. 1.3.2.



**Obr. 1.4.1:** Závislost intenzity záření na vlnové délce absolutně černých těles různých teplot.



**Obr. 1.4.2:** Normovaná citlivost čípků na vlnové délky

Charakteristika záření zaznamenaná lidským okem je pak rovna

$$I(\lambda) = [R(\lambda) + G(\lambda) + B(\lambda)] I_F(\lambda) \quad (1.3.1)$$

**Charakteristiky barvy:** Záření (1.3.1) je přeneseno zrakovým nervem do mozku, kde je zpracováno na zrakový vjem. Pro toto zpracování jsou nejdůležitější tyto parametry.

1) **Jas barvy** je dán hodnotou integrálu  $B = \int_{\lambda_1}^{\lambda_2} I(\lambda) d\lambda$

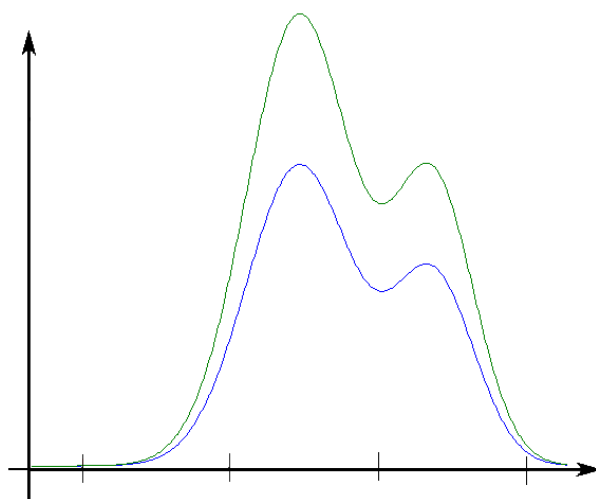
kde  $\langle \lambda_1; \lambda_2 \rangle$  je oblast viditelného světla. Speciálně tedy – jestliže pro charakteristiky dvou barev platí  $I_2(\lambda) = c \cdot I_1(\lambda)$ , pak druhá z nich má  $c$  - krát větší jas. Situace je znázorněna na obr. 1.3.3.

**2) Saturace barvy** je dána rozptylem charakteristiky kolem lokálních maxim

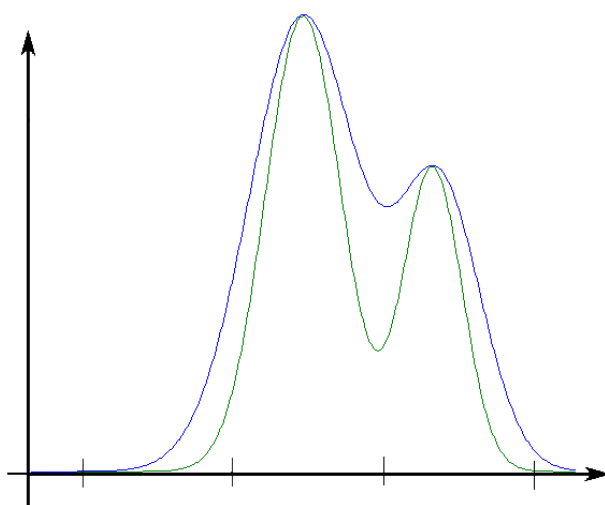
Speciálně tedy – jestliže pro charakteristiky dvou barev platí  $I_2(\lambda - m) = I_1(c \cdot (\lambda - m))$ , pak druhá má  $c$  - krát větší saturaci. Situace je znázorněna na obr. 1.3.4..

**3) Tón barvy** je dán počtem lokálních maxim charakteristiky a vzájemným poměrem jejich velikostí. Situace je znázorněna na obr 1.3.5.

Má-li charakteristika barvy jediné maximum a extrémně vysokou saturaci (tj. rozptyl je téměř nulový) dostáváme speciální barevné tóny – tyto barevné tóny vždy najdeme ve spektru, jedná se o tzv. spektrální barvy.

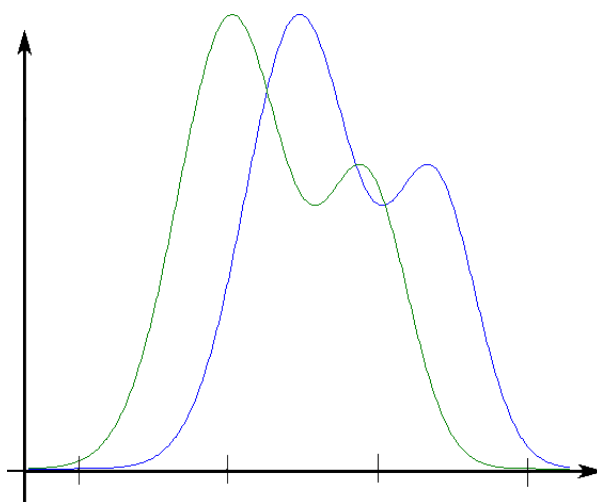


**Obr. 1.4.3:** Charakteristiky barev, které mají různou intenzitu při stejné saturaci a stejném barevném tónu



**Obr. 1.4.4:** Charakteristiky barev, které mají stejný barevný tón a různou saturaci a intenzitu (barva určená zelenou charakteristikou má menší intenzitu a vyšší saturaci).





**Obr. 1.4.5:** Charakteristiky barev, které mají různý barevný tón při stejné saturaci a při stejném jasu.

## 1. 5 Barevné systémy (modely)

Pokud bychom chtěli zpracovávat a uchovávat barevné obrazy tak, aby byly schopny zcela věrně reprodukovat barvy tak, jak je vnímá zdravé lidské oko, museli bychom mít pro každý pixel nekonečně mnoho zdrojů spektrálních barev, jejichž intenzitu by bylo možno spojitě měnit. To samozřejmě není možné. Proto byly vyvinuty dva základní barevné modely a řada tzv. barevných prostorů, které se více či méně úspěšně snaží tomuto ideálu přiblížit. Podle způsobu, jakým je vyvolán barevný vjem, rozlišujeme dva základní druhy barevných modelů.

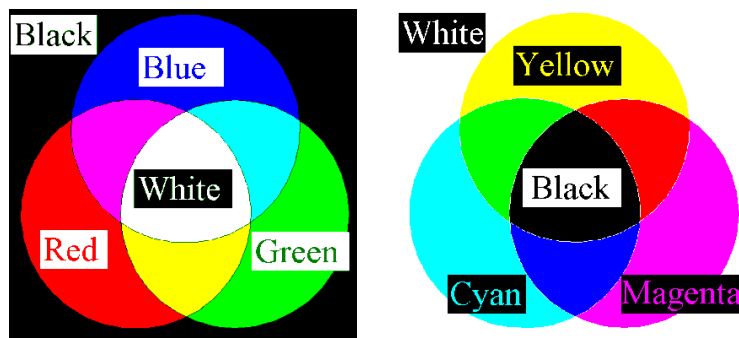
**Aditivní model:** Barva každého pixelu je složena z vlnových délek světla vyzařovaného ze tří zdrojů světla. Maxima záření jednotlivých zdrojů jsou dána maximy křivek na obr. 1.3.2 a oko je tedy vnímá jako červenou zelenou resp. modrou barvu. Pro tento model se tedy vžila zkratka **RGB (Red, Green, Blue)**. Jsou-li tyto zdroje dostatečně malé, oko je není schopno rozlišit a splývají mu v jeden bod a jednu barvu, která je dána intenzitou jednotlivých zdrojů. Tento model využívají všechny aktivní zobrazovací systémy (např. monitory, dataprojektory, barevné displeje mobilních telefonů apod.). Svítí-li na displeji např. všechny červené a zelené čipy v maximální intenzitě a nesvítí žádný modrý, vnímáme celý displej jako žlutý. Základní míchání barev v systému RGB vidíme na obr. 1.4.1 vlevo. Na obr. 1.4.2 je mikrofotografie části monitoru, vpravo pak pixely odpovídajících vnímaných barev.

Obraz v modelu RGB tak můžeme reprezentovat trojicí obrazových matic (**R;G;B**), vektorová funkce (5) pro je tvaru

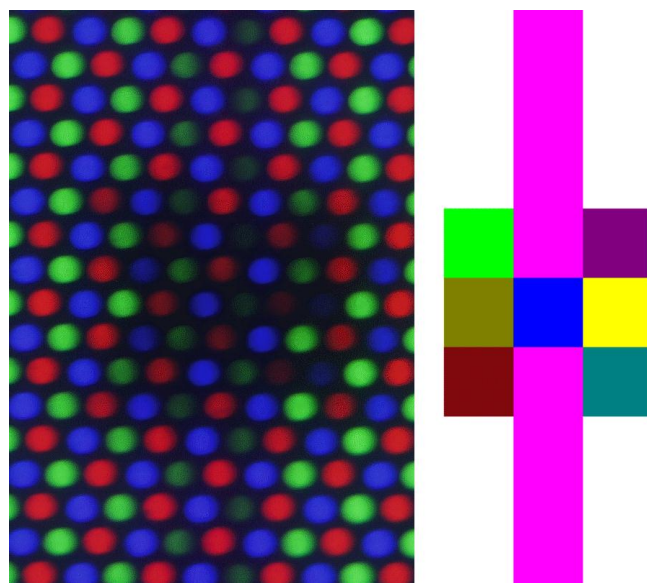
$$\mathbf{v}_{ij} = (R; G; B)_{ij} = (R_{ij}; G_{ij}; B_{ij})$$

**Subtraktivní model:** Barva každého pixelu je složena z vlnových délek zdrojů bílého světla, ve kterém byly nějakým způsobem potlačeny (odfiltrovány) vlnové délky odpovídající červené, zelené resp. modré barvě. Takto upravené světlo je pak vnímáno v tzv. doplňkových barvách - jako modrozelené (odfiltrována červená), purpurové (odfiltrována zelená) resp. žluté (odfiltrována modrá). Pro tento model se vžila zkratka **CMY (Cyan, Magenta, Yellow)**. Využívají ho barevné tiskárny, které tisknou na bílý papír. Ten je bílý proto, že odráží všechny vlnové délky viditelného spektra. Nanesení barevných inkoustů či tonerů působí jako filtr, který v definované intenzitě potlačí červenou, zelenou resp. modrou a máme tedy opět možnost barvy míchat. Základní míchání barev v systému CMY vidíme na obr. 1.4.1 vpravo.





**Obr. 1.4.1:** Základní míchání barev v systémech RGB (vlevo) a CMY (vpravo)



**Obr. 1.4.2:** Realizace barev v systému RGB. Vlevo mikrofotografie výřezu monitoru, vpravo tentýž výřez – několik pixelů v barvách, které vidí člověk po zpracování zrkového vjemu mozkiem.

I když je míchání barev ze složek CMY teoreticky možné, v praxi k těmto základním barvám přidáváme ještě černou, a to především ze dvou důvodů. Složením všech tří barev CMY by sice měla teoreticky vzniknout černá, reálně však vzniká barva hnědo šedivá. Druhým důvodem je skutečnost, že míchání černé z barevných složek je značně neekonomické. Proto se ke třem základním barvám přidává ještě černá. Tento systém je znám jako **CMYK**.

V modelu RGB lze definovat atributy pixelu, o kterých jsme se zmínili v předchozí kapitole, takto:

**Intenzita barevných složek:** Hodnoty barevných složek  $R;G;B$  jsou hodnoty z intervalu  $\langle 0;D \rangle$ , číslo  $D$  nazýváme **dynamickým rozsahem obrazu**. V modelu RGB tak máme k dispozici celkem  $(D+1)^3$  různých hodnot. Zdravé lidské oko je schopno rozlišit asi šest až osm milionů různých hodnot. V praxi proto většinou postačuje dynamický rozsah 255 – v tom případě obsazuje každá barevná složka 8 bitů – hovoříme o osmibitovém obrazu. K dispozici je  $(D+1)^3 = 256^3 = 16\,777\,216$  barev. Pro vědecké účely je však často potřeba vyšší hodnoty – obrazy dvanáctibitové ( $D = 4\,096$ ) a šestnáctibitové ( $D = 65\,535$ ).

**Jas:** Je dán celkovou intenzitou všech tří barevných složek daného pixelu. Rozeznáváme dva typy jasu – jas **objektivní** a jas **subjektivní**. Objektivní jas je určen intenzitou záření pixelu, kterou by naměřil fotometr, v modelu RGB tedy zřejmě

$$J_o = \frac{R + G + B}{3}$$

Subjektivní jas je určen intenzitou záření pixelu tak, jak ho vnímá lidské oko, které je na barevné složky různě citlivé. Pro subjektivní jas platí:

$$J = c_1 R + c_2 G + c_3 B; \quad c_1 + c_2 + c_3 = 1$$

Hodnota konstant  $c_1 + c_2 + c_3$  se určuje experimentálně, většinou se uvádí

$$c_1 = 0.3; \quad c_2 = 0.59; \quad c_3 = 0.11$$

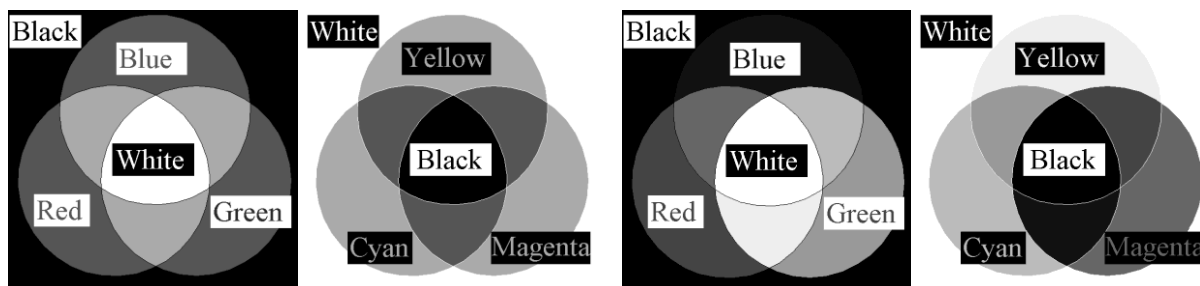
Různé prameny udávají pro tyto konstanty mírně rozdílné hodnoty, pro běžnou potřebu jsou však tyto rozdíly zanedbatelné.

Zatímco objektivní jas červené, zelené i modré je stejný ( $1/3$ ), je subjektivní jas červené (v maximální intenzitě) roven  $J_R = 0.3D$ ; zelené  $J_G = 0.59 \cdot D$  a modré  $J_B = 0.11 \cdot D$ . Je to dáno tím, že právě zelená je zhruba uprostřed spektra viditelného záření Slunce. Je to vlastně obráceně – Slunce zasobuje Zemi po miliardy let zářením, které má podle Planckova zákona maximum ve žlutozelené oblasti. Proto se (nejen) lidský zrak vyvinul tak, že je nejcitlivější na zelenou barvu. Subjektivní jas doplňkových barev je pak

modrozelená:	$J_C = J_G + J_B = 0.59D + 0.11D = 0.70D$
purpurová	$J_M = J_R + J_B = 0.30D + 0.11D = 0.41D$
žlutá	$J_Y = J_R + J_G = 0.30D + 0.59D = 0.89D$

Největší subjektivní jas má tedy žlutá, lidské oko ji tedy vidí jako nejjasnější, což se projevuje zvláště při nedostatečném osvětlení. Objektivní jas všech doplňkových barev je přitom stejný –  $J_0 = \frac{2}{3}$ .

**Obraz v odstínech šedi:** Někdy je potřeba převést barevný obraz na obraz v odstínech šedi, matematicky řečeno – trojici obrazových matic  $(\mathbf{R}; \mathbf{G}; \mathbf{B})$  reprezentovat jedinou maticí, jejíž prvky chápeme jako odstíny šedi, tj. maticí  $\mathbf{J}$ , jejíž prvky budou hodnoty jasu. Tento obraz se často nazývá Grey Scale. Některé systémy na zpracování obrazu skutečně používají jas objektivní (např. ImageJ) a některé jas subjektivní (např. IrfanView). Jak však plyne z předchozího odstavce, při použití objektivního jasu mohou některé barvy splynout, protože jejich objektivní jas je stejný (např. základní i doplňkové barvy při stejných hodnotách RGB). Korektnější je použití jasu subjektivního, který je pro tyto barvy různý, proto barvy nesplynou – žlutá bude světlejší, modrozelená naopak tmavší než purpurová. Na obr. 1.4.3 máme základní barvy systémů RGB a CMY při převodu do odstínů šedé. Vlevo při použití objektivního jasu splynuly všechny základní barvy do jedné, doplňkové rovněž. Osm barev tak bylo přededeno jen do čtyř odstínů. Při použití subjektivního jasu (vpravo) je po převodu téže předlohy na obraze osm odstínů.



**Obr. 1.4.3:** Základní a doplňkové barvy systému RGB po převodu do GreyScale Vlevo je počítán objektivní jas (software ImageJ), vpravo jas subjektivní (software Irfan View).

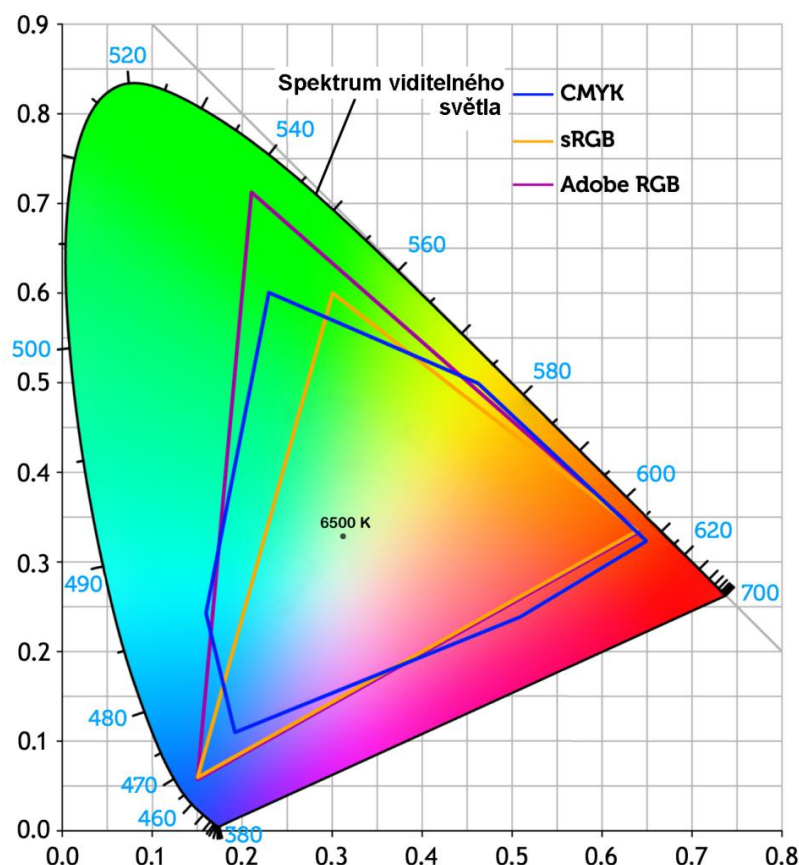
## 1. 6 Barevné prostory

Barevným prostorem rozumíme přesně definovanou množinu barev, kterou je dané zařízení schopno snímat, zobrazit nebo reprodukovat. Vycházejí z barevného modelu, na rozdíl od něj však mají přesným způsobem definované barvy. Rozsah barev, které jsou v daném prostoru definovány, se nazývá **gamut**.

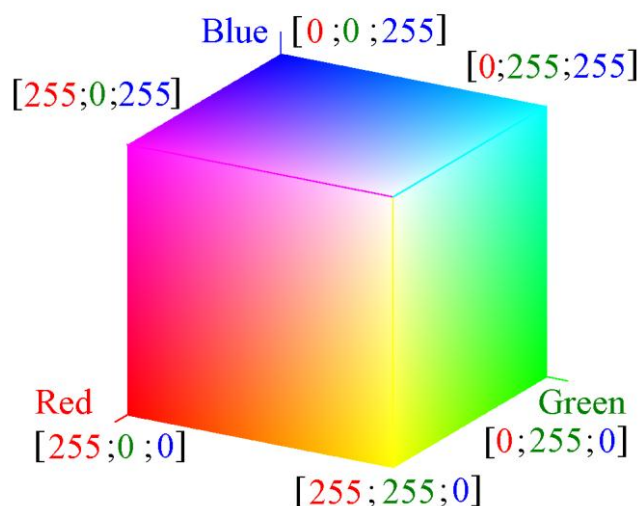
**Barevné prostory založené na modelu RGB:** Jsou prostory, ve kterých se barvy míchají ze složek RGB. Liší se definicí těchto složek a definicí bílé. Některé z nich jsou závislé na zařízení, některé jsou nezávislé. Jedním z prvních matematicky definovaných barevných prostorů je **CIE XYZ**. Byl definován mezinárodním úřadem **Commision Internationale de l'Eclairage** v r. 1931. Koncept vnímání barev rozděluje na dvě části tak, jak bylo popsáno v úvodu kpt. 1. 3. Parametr Y vyjadřuje jas, vnímaný čípky, a vlastní barva, rozpoznávaná tyčinkami, je specifikována dvěma odvozenými parametry x,y. Tento barevný prostor je znázorněn chromatickým diagramem na obr. 1.5.1. Nejrozšířenějším barevným prostorem v současnosti je zřejmě **standard RGB (sRGB)**. Jeho základem je osmibitový RGB. Každá barevná složka obsazuje jeden byte a má tedy k dispozici 256 hodnot. Prostor se znázorňuje jako krychle – viz obr. 1.5.2. Gamut tohoto prostoru je na obr. 1.5.3 vyznačen žlutým trojúhelníkem. Barvy mimo tento trojúhelník není schopen prostor sRGB reprodukovat. Dalším prostorem je **Adobe RGB**. Má širší gamut než sRGB – dokáže zachytit některé barvy, které jsou mimo sRGB, ale některé monitory a tiskárny je dokáží zobrazit. Gamut AdobeRGB je na obr 1.5.3 vyznačen purpurovou barvou. Existuje několik dalších standardů na bázi RGB, např. **Apple RGB, Pro Photo RGB** apod.

**Systémy HSV; HSL** – jsou ekvivalentní s RGB. Transformují RGB krychli na pravidelný šestiboký jehlan, dvojkužel či válec. Jsou to barevné systémy, jehož složky se snaží respektovat běžné vnímání barvy. Těmito složkami jsou:

- a) **Hue (barevný tón)**
- b) **Saturation (saturace, sytost barvy)**
- c) **Value (jas, světelnost)**



**Obr. 1.5.1:** Barevný prostor CIE – obvod tvoří spektrální barvy, uvnitř vyznačeny gamuty některých dalších používaných prostorů.



**Obr. 1.5.2:** Barevný prostor sRGB – RGB krychle

Tyto systémy respektují vymezení barvy tak, jak bylo uvedeno v předchozí kapitole. Jas klesá se vzdáleností od podstavy kuželu či jehlanu, ve vrcholu nulový. Saturace a barevný tón tvoří jakousi obdobu polární soustavy souřadné, kde ovšem barvy se stejnou saturací neleží na kružnici, ale na obvodu pravidelného šestiúhelníka. Tuto nepříjemnost odstraňuje model **HSL** (Hue-Saturation-Lightness). Barevný prostor má tvar dvojkuželu, kde saturace a barevný tón tvoří (tentokrát klasickou) polární soustavu souřadnou a barvy v rovině podstavy mají

poloviční světelnost. Ta se vzdáleností od podstavy roste směrem k „bílému“ vrcholu a klesá směrem k „černému“ vrcholu. V grafických systémech využívajících systémy HSV resp. HSL nenajdeme většinou jehlan či dvojkužel, ale některou z jejich možných transformací.

**Barevný tón** je v těchto modelech geometricky znázorňován jako úhel. Přepočty RGB na HSL je následující. Odstíny šedé, tj. případy kdy  $R = G = B$  se přepočítávají na  $H = S = 0$  a  $L = R = G = B$ . V ostatní případech se postupuje následovně : V HSL se položí

$$L = \max \{R; G; B\}.$$

Konečně výpočet úhlu pro barvu ( $H$ ):

$$L = R \quad \Rightarrow \quad H = 60 \cdot \frac{G - B}{L - \min \{R; G; B\}}; \quad H < 0 \Rightarrow H := 360 + H$$

$$L = G \quad \Rightarrow \quad H = 120 \cdot \frac{R - B}{L - \min \{R; G; B\}}$$

$$L = B \quad \Rightarrow \quad H = 270 \cdot \frac{R - G}{L - \min \{R; G; B\}}$$

**Saturace barvy:** V modelu RGB je saturace dána vztahem

$$S = \max \{R; G; B\} - \min \{R; G; B\}$$

popř. v normalizované podobě

$$S_n = \frac{\max \{R; G; B\} - \min \{R; G; B\}}{\max \{R; G; B\}}; \max \{R; G; B\} = 0 \Rightarrow S_n = 0$$

**Systém YUV:** Systém sRGB poskytuje velmi širokou paletu barev, je však poměrně nevýhodný pro bezztrátovou kompresi (sousední pixely mají velmi často velmi blízké hodnoty). To je velká nevýhoda např. při přenosu televizního signálu. Model YUV tuto nevýhodu do značné míry potlačuje. Odděluje jasovou složku  $Y$  - jas a dvě barevné složky  $U$  a  $V$  (písmeno  $Y$  tedy nezaměňovat s  $Y$  v modelu CMY, kde značí žlutou). Vektor  $\mathbf{y} = (y; u; v)$  obdržíme z vektoru  $\mathbf{r} = (r; g; b)$  pomocí rovnice

$$\mathbf{y}^T = \mathbf{M} \cdot \mathbf{r}^T : \begin{pmatrix} y \\ u \\ v \end{pmatrix} = \begin{pmatrix} 0,299 & 0,587 & 0,144 \\ -0,147 & -0,289 & 0,436 \\ 0,615 & -0,515 & -0,100 \end{pmatrix} \begin{pmatrix} r \\ g \\ b \end{pmatrix}$$

Matice převodu je regulární, existuje k ní tedy matice inverzní, tudíž rovněž jednoznačný převod zpět z YUV do RGB. Nalezneme ho řešením maticové rovnice (3):

$$\left. \begin{array}{l} \mathbf{M} \cdot \mathbf{r}^T = \mathbf{y}^T / \cdot \mathbf{M}^{-1} \text{ zleva} \\ \mathbf{M}^{-1} \cdot \mathbf{M} \cdot \mathbf{r}^T = \mathbf{M}^{-1} \cdot \mathbf{y}^T \\ \mathbf{r}^T = \mathbf{M}^{-1} \cdot \mathbf{y}^T \end{array} \right\} \Rightarrow \begin{pmatrix} r \\ g \\ b \end{pmatrix} = \begin{pmatrix} 0,299 & 0,587 & 0,144 \\ -0,147 & -0,289 & 0,436 \\ 0,615 & -0,515 & -0,100 \end{pmatrix}^{-1} \begin{pmatrix} y \\ u \\ v \end{pmatrix} = \begin{pmatrix} 1,000 & 0,000 & 1,137 \\ 1,000 & -0,397 & -0,580 \\ 1,000 & 2,034 & 0,000 \end{pmatrix} \begin{pmatrix} y \\ u \\ v \end{pmatrix}$$

(vypočtete inverzní matici!)

**Barevné prostory založené na modelu CMYK:** Jsou prostory, ve kterých se barvy míchají ze složek Cyan, Magenta, Yellow a Black. Na rozdíl od některých prostorů založených na RGB jsou tyto prostory vždy závislé na zařízení. Proto je těchto prostorů velmi mnoho. Převod barev z prostorů RGB do CMYK může být velký problém, především pokud se týká barev, které CMYK nedokáže zobrazit. Gamut prostorů založených na CMYK je přibližně vyznačen na obr. 13 modrou barvou.

## 1. 7 Formáty rastrových grafických souborů

Formátem grafického souboru rozumíme strukturu používanou k uložení grafických dat na paměťovém médiu. Rastrový soubor se obecně skládá z

**hlavičky**, která obsahuje základní informace o souboru (rozměry obrázku, bitovou hloubku, typ případné komprese, odkaz na začátek obrázku apod.)

**obrazové matice**

**paty obsahující** doplňující informace

V souboru je uložena buď celá obrazová matice pixel po pixelu (v tom případě mluvíme o nekomprimovaném formátu), anebo je informace o obrazové matici zakódována tak, že tento kód má menší nároky na paměť (v tom případě se jedná o komprimovaný formát). Uvedme nejdříve dva nejdůležitější formáty nekomprimované.

**BMP (Windows Bitmap)** – Pro dynamický rozsah 8 bitů na barevnou složku (true color zobrazení) nepoužívá komprese, přibližná velikost je pak výška x šířka x dyn. rozsah + velikost hlavičky. Při menší barevné hloubce může být použita komprese RLE (viz dále).

**RAW** – formát pro uložení originálních, nijak neupravených obrazů přímo z CCD snímače. Formát není standardizován, každý výrobce hardware má poněkud jiný RAW se svojí vlastní příponou, popis formátu bývá součástí dokumentace k přístroji. Pro vědecké účely je nejvhodnější, protože není „předzpracován“ firemním softwarem.

**Komprimace (komprese) souborů:** rozlišujeme komprimaci **bezeztrátovou** a **ztrátovou**. Při ztrátové kompresi jsou některé informace nenávratně ztraceny a nelze je zpět zrekonstruovat. Používá se tam, kde je možné ztrátu některých informací tolerovat a kde nevýhoda určitého zkreslení je bohatě vyvážena velmi významným zmenšením souboru. Bezeztrátová komprese obvykle není tak účinná jako ztrátová komprese dat. Velkou výhodou je, že komprimovaný soubor lze opačným postupem rekonstruovat do původní podoby. To je nutná podmínka při přenášení velké většiny vědeckých dat., kde by ztráta i jediného byte mohla znamenat nenávratné poškození souboru.

**RLE (Run Length Encoding)** Nejčastější implementace je komprese bytového proudu. Implementuje se pomocí dvojic bytů, jeden byte určuje počet opakování, druhý byte udává hodnotu kódovaného bytu. Nevýhoda je, že pro jednu samostatnou hodnotu v proudu bytů se kóduje tato hodnota pomocí dvojice bytů, proto pro nevhodná vstupní data může být komprese neúčinná a velikost proudu z výstupu kóderu RLE naopak až dvojnásobná oproti vstupu.

V počítačové grafice a analýze obrazu se uplatní nejlépe v případě, že obraz obsahuje větší plochy stejné barvy. V grafických souborových formátech existují různé varianty RLE komprese, buď jako hlavní komprese dat (např. PCX) nebo jako pomocná metoda komprese dat (JPEG). Jako příklad uvedme zakódování posloupnosti

06060606060606060606060620202056ABABABABAB

při osmitovém kódování. Kódujeme tedy pomocí dvojice cifer v šestnáctkové soustavě ( $2^8=16^2=256$ ), ve které se používají cifry 0;1;...;9;A;B;C;D;E;F. V naší posloupnosti se nejprve dvanáctkrát opakuje dvojice znaků '06', kterou tedy zakódujeme jako 0C06. Celá zakódovaná posloupnost pak vypadá takto:

0C060320015605AB

**Huffmanovo kódování:** znaky, které se ve vstupním souboru vyskytují nejčastěji, jsou konvertovány do bitových řetězců s nejkratší délkou (nejfrekventovanější znak tak může být konvertován do jediného bitu), znaky, které se vyskytují velmi zřídka, jsou konvertovány do delších řetězců (mohou být i delší než 8 bitů) ukládáno jako binární strom

**LZW (Lempel-Ziv-Welch)** je založeno na "slovníku" (vyhledává opakující se posloupnosti, které chápe jako slova a každému slovu je přiřazen jeden znak). Algoritmus se hodí především pro síťové účely, kde ještě není znám konec přijímaných dat, ale jejich začátek už se může dekomprimovat.

Bezeztrátovou kompresi používají především tyto formáty:

**TIFF (Target Image File Format):** Velmi univerzální formát. Neoficiální standard pro tisk s bohatými možnostmi (komprese, barevná hloubka, možnost uložení více obrazů v jednom souboru - multitiff, informace určené přímo pro tisk, některé vektorové informace...)

**PNG (Portable Network Graphics):** má většinou účinnější kompresi než tif, kombinuje LZW, LZ77 a Huffmanovo kódování. Umožňuje uložit obrazy s dynamickým rozsahem až 48 bitů.

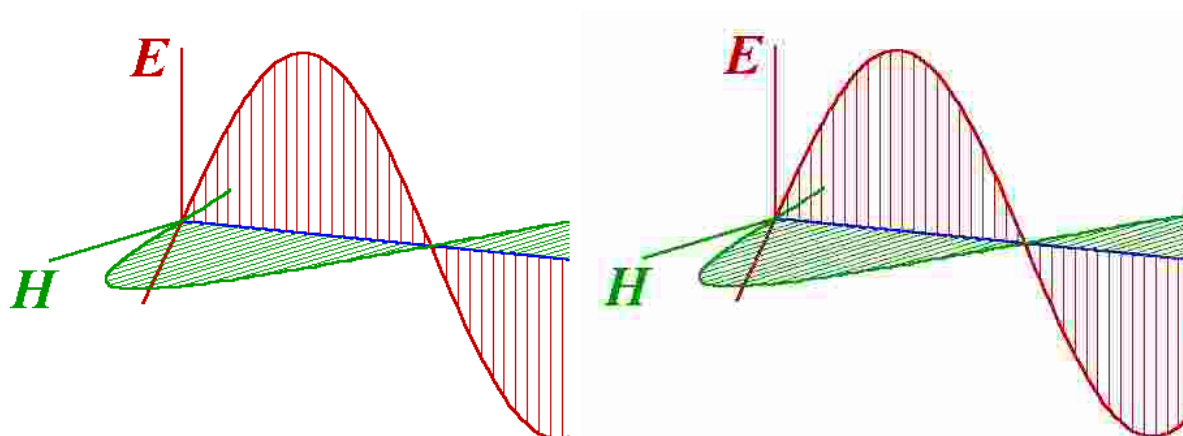
**GIF** používá kompresi LZW, vhodný jen pro obrázky s malým počtem barev (max 256) a většími jednobarevnými plochami. Umožňuje uložit více obrazů v jednu souboru jako animaci.

**PCX (PC Paintbrush File Format)** umožňuje ukládat obrázky s barevnou hloubkou 1bit, 4bity, 8bitů a 24bitů. Komprese bitmapy buď RLE nebo bez komprese.

**Ztrátová komprimace obrazů:** používá se tam, kde není nezbytně nutné přesně zachovat původní informaci. Obvykle je pak určitá ztráta informace vyvážena výraznou úsporou místa. Např. obrazy, které nejsou určeny k vědeckým účelům, mohou pro lidský zrak zůstat vysoce kvalitní i při poměrně značné ztrátě původní informace. Typickou informací postradatelnou pro lidský zrak je střídání pixelů s velmi blízkým jasem či odstínem, které lidský zrak není schopen rozlišit. Toto střídání navíc nemusí představovat užitečnou („reálnou“) informaci, ale nežádoucí šum. Kompresní metoda tak může snížit reálný počet barev či odstínů šedi, což může v jistých případech dokonce zvýšit kvalitu obrazu.. Při vyšším požadavku na úsporu místa se ale projeví nedostatečnou kvantizací tak, jak můžeme vidět na obr. 1.2.4. Ztrátovou kompresi nelze rovněž doporučit u plánek, technických výkresů apod. Tenké čáry a ostré přechody barvy či jasu, které se na těchto obrazech vyskytují, totiž představují vysokofrekvenční informace, které ztrátová komprese rovněž omezí. Do obrazu je tak zanesen nežádoucí barevný šum tak, jak je vidět na obr. 1.6.1

Ztrátovou kompresi dnes používá především formát **JPEG (Joint Photographic Experts Group)**. Zajímavou alternativou do budoucna může být tzv. fraktálová komprese, která při ztrátě srovnatelné s formátem JPEG komprimuje podstatně účinněji. Jejím rozšíření dnes brání neúnosná doba komprimace, která se při běžných velikostech obrázků pohybuje řádově v minutách až desítkách minut.





**Obr. 1.6.1:** Tenké čáry a ostré barevné přechody znehodnocené barevným šumem o nižších frekvencích při ztrátové kompresi.

## 2 Obraz jako statistický soubor

Ze střední školy víme, že statistika se zabývá zpracováváním číselných údajů o tzv. **hromadných jevech**, tj. údajů o různých skutečnostech, nikoli však jednotlivě, ale ve velkém počtu případů. Množinu všech takto pozorovaných objektů nazýváme **statistickým souborem**, její prvky **statistické jednotky** a jejich počet  $n$  je **rozsah statistického souboru**. Společnou vlastnost  $x$  všech prvků statistického souboru, jejíž proměnlivost je předmětem statistického šetření, nazýváme **statistickým znakem**, konkrétní hodnota  $x_1; x_2; \dots; x_n$  zjištěná pro každý znak je **hodnota statistického znaku**. „Mechanismus“, který prvkům statistického souboru přiřazuje hodnotu statistického znaku (například každému člověku jeho výšku či váhu) je z matematického hlediska funkcí. Funkční hodnota této funkce však závisí na náhodě, tj. na množství parametrů, které neumíme matematicky popsat. Proto tuto funkci nazýváme **náhodná veličina**.

I obraz můžeme chápat jako statistický soubor. Statistickými jednotkami jsou v tom případě jednotlivé pixely, rozsahem souboru je číslo  $n = w \cdot h$ , kde  $w$  je šířka a  $h$  je výška obrazu v pixelech. Statistickým znakem může být jas, červená složka barvy, saturace apod., hodnotou znaku v každém pixelu je hodnota veličiny, která nás momentálně zajímá. Tu ovlivňuje mnoho okolností při pořizování obrazu. Například na fotografii se projeví volba expozice a zaostření, které nemůžeme zcela přesně kvantifikovat, dále světelné podmínky, tepelné chvění atmosféry i součástí fotoaparátu, prach na objektivu apod. Fotografie je tedy z hlediska analýzy obrazu náhodnou veličinou.

### 2. 1 Rovnoměrné a normální rozdělení

Jak uvidíme dále, v analýze obrazu hrají v určitých situacích zásadní roli tzv. rozdělení náhodné veličiny. Uvedme alespoň ty z nich, které budeme dále bezprostředně potřebovat.

**Rovnoměrné rozdělení** má náhodná veličina určená funkcí

$$p(x) = \begin{cases} \frac{1}{2h}; & x \in \langle \mu - h; \mu + h \rangle \\ 0; & x \notin \langle \mu - h; \mu + h \rangle \end{cases} \quad (2.1.1)$$

Funkce  $p(x)$  se nazývá hustota pravděpodobnosti rovnoměrného rozdělení. Číslo  $2h$  se nazývá rozpětí, číslo  $\mu$  se nazývá střední hodnota, číslo  $\frac{h^2}{3}$  pak jeho rozptyl.

Rovnoměrným rozdělením se řídí náhodné veličiny, které mají stejnou možnost nabytí kterékoliv hodnoty z daného intervalu, např. odečítání údajů z lineárních stupnic měřicích přístrojů nebo chyby při zaokrouhlování čísel. Ty se mohou projevit při kvantizaci obrazu jako kvantizační šum.

**Normální rozdělení** má náhodná veličina určená rovnicí

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (2.1.2)$$

Funkce  $p(x)$  se nazývá hustota pravděpodobnosti normálního rozdělení. Číslo  $\mu$  se nazývá střední hodnota,  $\sigma^2$  je rozptyl normálního rozložení a  $\sigma$  je směrodatná odchylka. Toto rozdělení je použitelné všude tam, kde kolísání náhodné veličiny způsobuje součet velkého počtu vzájemně nezávislých jevů. Speciálně v analýze obrazu je toto rozdělení typické pro tzv. aditivní šum.

## 2. 2 Základní výběrové charakteristiky

Statistické zkoumání můžeme provádět na souboru, který zahrnuje všechny objekty, které v dané situaci přicházejí v úvahu. Tento soubor se nazývá **základní soubor**. Takto rozsáhlé šetření však většinou není možné, proto ze základního souboru vybíráme soubor menší. Výběru, čímž dostáváme tzv. **soubor výběrový**. Jestliže je náš výběr náhodný a výběrový soubor má velký počet prvků, lze na základě vlastností tohoto výběru usuzovat na vlastnosti celého základního souboru. Nejdůležitějšími vlastnosti výběrového souboru jsou tzv. výběrové charakteristiky.

Ze střední školy známe pojmy aritmetický průměr, modus, medián (tzv. charakteristiky polohy) a rozptyl, resp. směrodatná odchylka (tzv. charakteristiky variability). Zopakujme, že **aritmetický průměr**  $\bar{x}$  zjištěný z hodnot  $x_1; x_2; \dots; x_n$  je dán vztahem

$$\bar{x} = \frac{1}{n} \cdot \sum_{i=1}^n x_i, \quad (1.2.1)$$

**modus** je nejčastěji se opakující hodnota znaku, tj. hodnota znaku, který má největší četnost a **medián** souboru uspořádaného podle velikosti statistického znaku je prostřední hodnota znaku, přesněji řečeno u souboru s lichým rozsahem je to prvek s indexem  $(n+1)/2$  a u souborů se sudým rozsahem je to aritmetický průměr prvků s indexy  $n/2; (n/2)+1$ .

Na histogramu obrazu je tedy modus nejvyšší sloupec a medián rozdělí histogram na dvě části se stejnou plochou. Aritmetický průměr má praktický význam pouze u souborů, které mají výrazný modus, tj. histogram má zřetelné maximum a dále je tento histogram přibližně symetrický, tj. modus se příliš neliší od mediánu.

**Výběrový rozptyl** je aritmetickým průměrem čtverců odchylek hodnot znaku od aritmetického průměru, je tedy dán jako

$$\sigma^2 = \frac{1}{n} \cdot \sum_{i=1}^n (x_i - \bar{x})^2 \quad (1.2.2)$$

a **směrodatná odchylka** je odmocnina z výběrového rozptylu, tj.

$$\sigma = \sqrt{\sigma^2} = \sqrt{\frac{1}{n} \cdot \sum_{i=1}^n (x_i - \bar{x})^2} \quad (1.2.3)$$

## 2.3 Histogram obrazu

Obraz je ze statistického hlediska výběrovým souborem. Předpokládejme, že nás zajímá hodnota jasu pixelu v obraze s rozlišením  $1024 \times 768$  a dynamickým rozsahem  $D=255$ . Je zřejmé, že některé hodnoty jasu se v obraze musejí mnohokrát opakovat (pixelů je totiž 786 432 a jejich možných hodnot jen 256). Označme  $k$ ;  $0 < k \leq 255$  počet jasových úrovní, které se v obraze skutečně vyskytují. Počet pixelů, které mají stejný jas  $v_i$ ;  $i=0;1;\dots;k$  se nazývá **absolutní četnost znaku** (v tomto případě hodnoty  $v_i$ ) a značí se  $n_i$ . Podíl četnosti znaku  $n_i$  a rozsahu souboru  $n$  se nazývá **relativní četnost znaku**  $v_i$  a budeme ji značit  $c_i$ . Platí tedy:

$$c_i = \frac{n_i}{n} = \frac{n_i}{w \cdot h}; \quad i=0;1;\dots;k$$

$$\sum_{i=1}^k n_i = n_1 + n_2 + \dots + n_k = n = w \cdot h$$

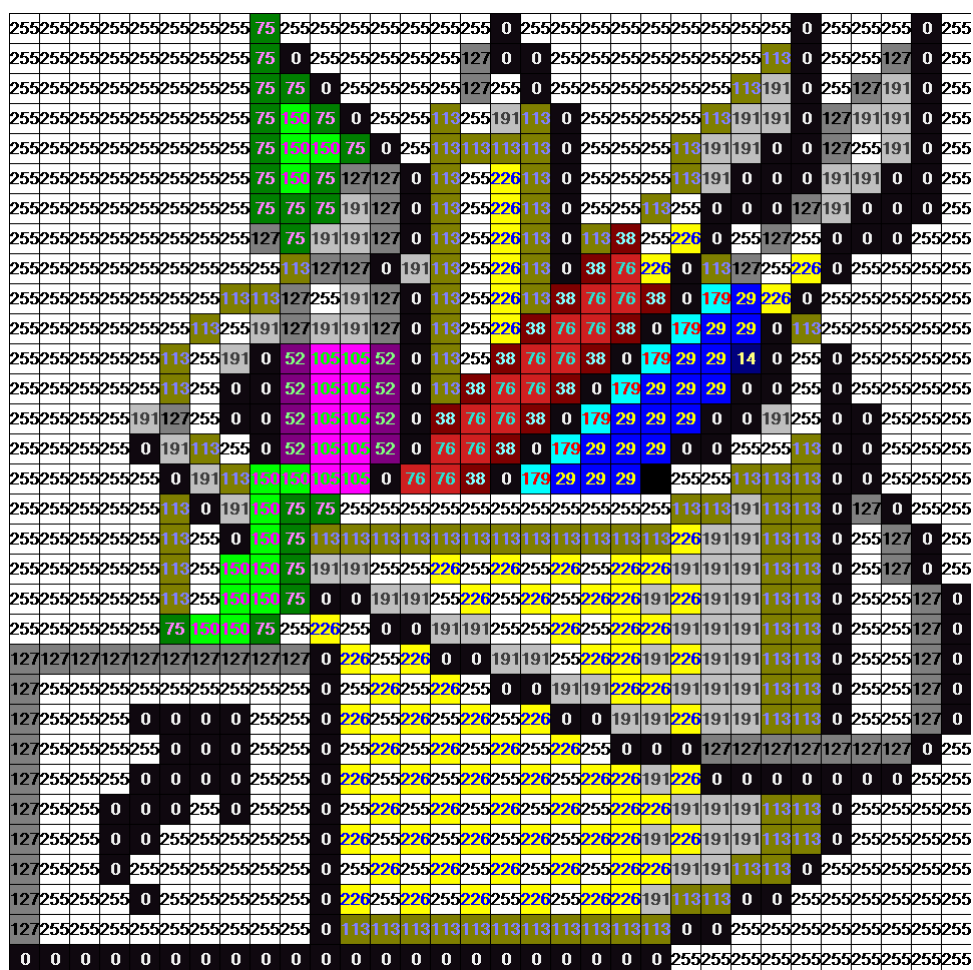
$$\sum_{i=1}^k c_i = \sum_{i=1}^k \frac{n_i}{n} = \frac{n_1}{n} + \frac{n_2}{n} + \dots + \frac{n_k}{n} = \frac{1}{n} \cdot (n_1 + n_2 + \dots + n_k) = 1$$

Relativní četnosti se často vyjadřují v procentech, v tom případě je

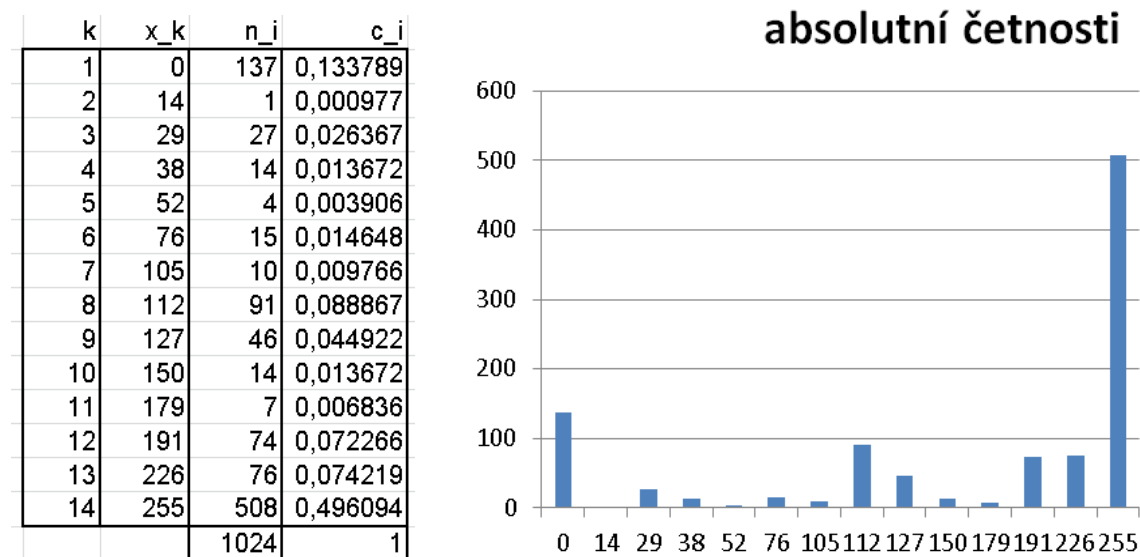
$$\sum_{i=1}^k 100\% \cdot c_i = 100\% \cdot \sum_{i=1}^k \frac{n_i}{n} = 100\% \cdot \left[ \frac{n_1}{n} + \frac{n_2}{n} + \dots + \frac{n_k}{n} \right] = \frac{100\%}{n} \cdot (n_1 + n_2 + \dots + n_k) = 100\%$$

Zapíšeme-li všechny četnosti resp. relativní četnosti do tabulky, nazýváme tuto tabulku **tabulkou četností** resp. **tabulkou relativních četností**.

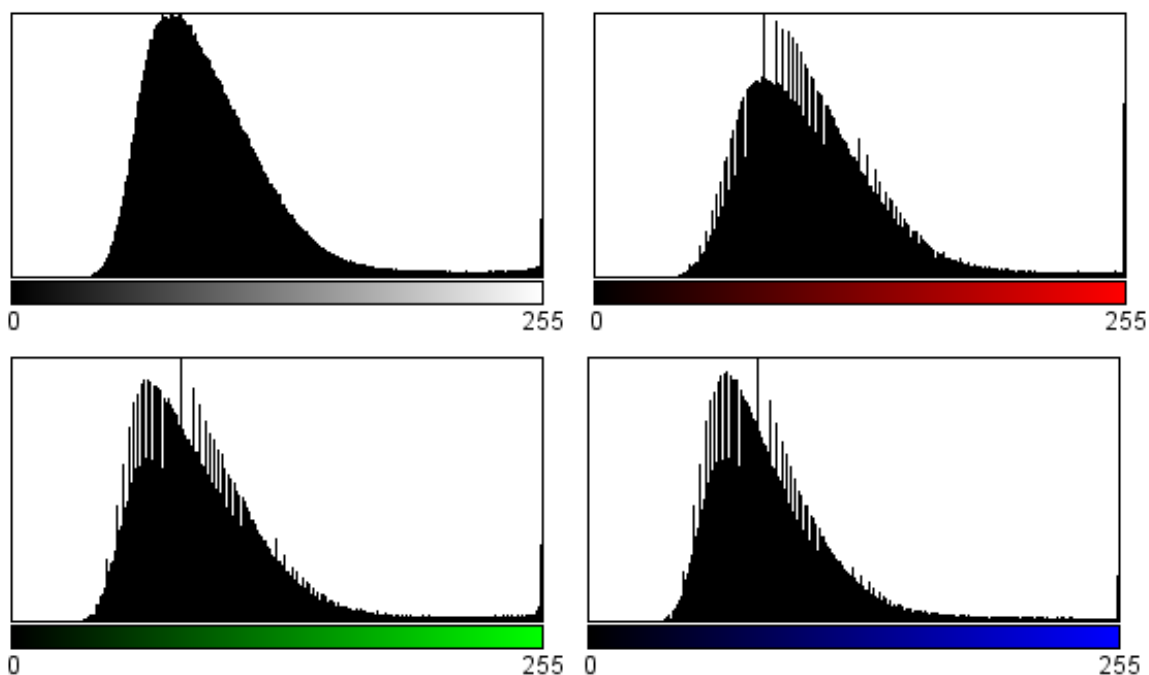
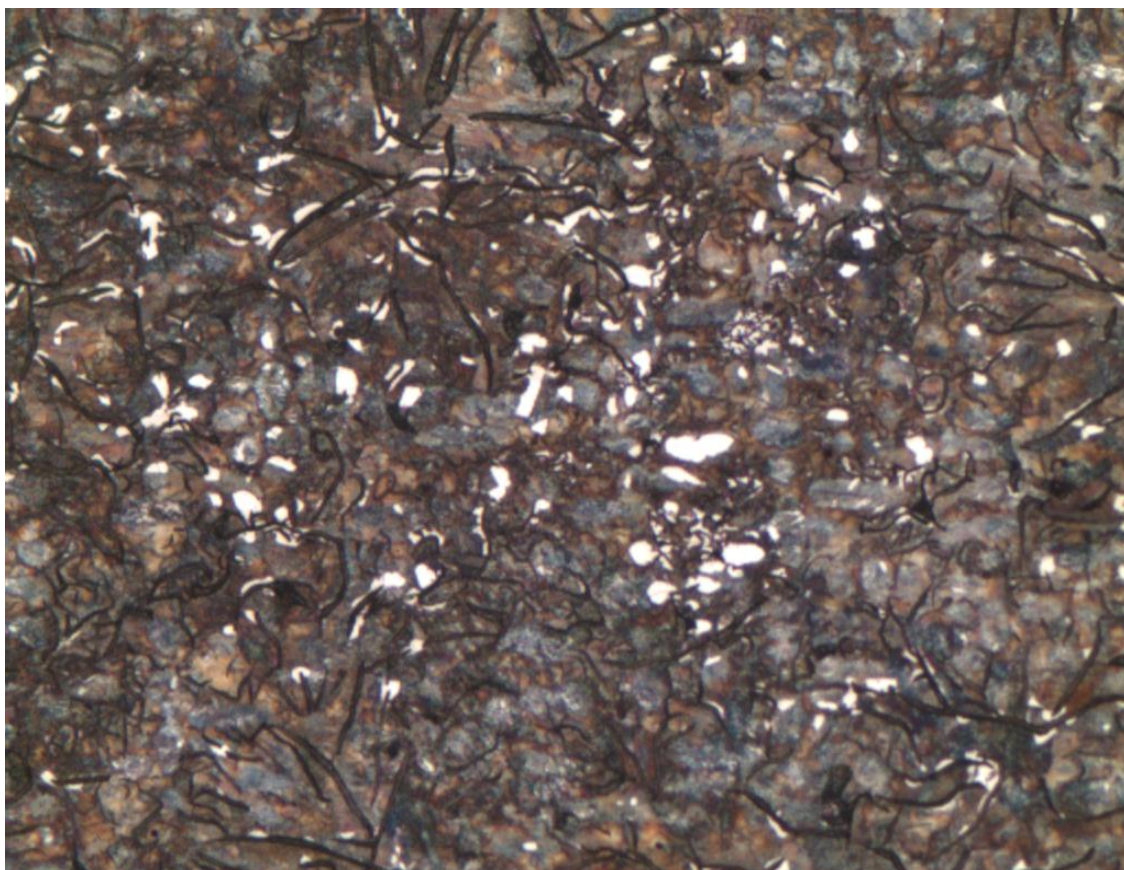
Na obr. 2.3.1 je ikona programu Malování v rozlišení  $32 \times 32$  pixelů a formátu 256 barev. Hodnoty pixelů zde znamenají pořadí barvy v paletě. Na obr. 2.3.2 vlevo je tabulka absolutních a relativních četností hodnot pixelu a vpravo je histogram tohoto obrazu. Na obr. 2.3.3. je mikrosnímek feritu a jeho histogramy objektivního jasu a jednotlivých barevných složek.



**Obr. 2.3.1:** Ikona programu Malování ve formátu 256 barev. Hodnoty pixelů zde znamenají pořadí barvy v paletě.



**Obr. 2.3.2:** Tabulka absolutních a relativních četností (vlevo) a histogram absolutních četností (vpravo) obrazu 2.3.1.



**Obr. 2.3.3:** Mikrosnímek feritu a jeho histogramy objektivního jasu a jednotlivých barevných složek.

## 2. 4 Úprava jasu a kontrastu

Histogram jasu obrazu umožňuje velmi snadno odhadnout výběrové charakteristiky obrazu. Ideální obraz má rovnoměrné rozložení jasu s rozpětím  $D$  a střední hodnotou  $\mu/2$ . Sloupce histogramu jasu by tedy měly být přibližně stejně vysoké, a to na celém intervalu  $\langle 0; D \rangle$ .

Většinou tomu tak ale není. Pak je na místě otázka, jakými úpravami je možné se k tomuto ideálnímu stavu alespoň přiblížit.

Abychom byli v této snaze úspěšní, je třeba z histogramu správně odhadnout výběrové charakteristiky obrazu uvedené v předchozí kapitole.

Má-li histogram zřetelně patrný jeden nejvyšší sloupec, je tímto sloupcem jednoznačně určen modus obrazu – hovoříme o tzv unimodálním histogramu resp. obraze. Je-li sloupců, které tvoří výrazná a zřetelně oddělená maxima, jedná se o histogram bimodální popř. multimodální. Medián se nachází ve sloupci, který dělí histogram na dva přibližně rovnoploché útvary. Aritmetický průměr jako výběrová charakteristika má význam pouze u unimodálních obrazů, jejichž medián a modus jsou blízké hodnoty a histogram je alespoň přibližně symetrický. V tom případě se ani průměr příliš neliší od modu a mediánu a nachází se přibližně uprostřed dynamického rozsahu obrazu.

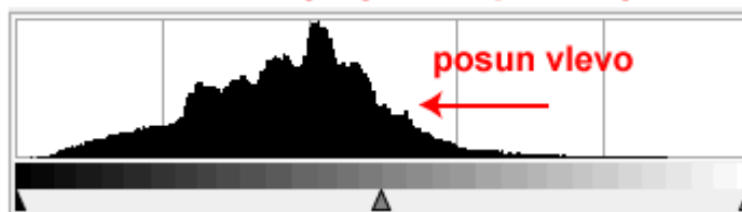
Výběrový rozptyl se na histogramu projevuje jako jeho šířka – širší histogram signalizuje větší rozptyl, užší histogram naopak znamená menší  $\sigma^2$ .

**Medián posunutý** ze středu dynamického rozsahu **směrem dolů signalizuje nízký jas**. Jeho **posun k vyšším hodnotám** naopak **znamená**, že obraz má **jas příliš vysoký**.

**histogram obrázku s vysokým jasnem - převažují světlé pixely**



**histogram obrázku s nízkým jasnem - převažují tmavé pixely**

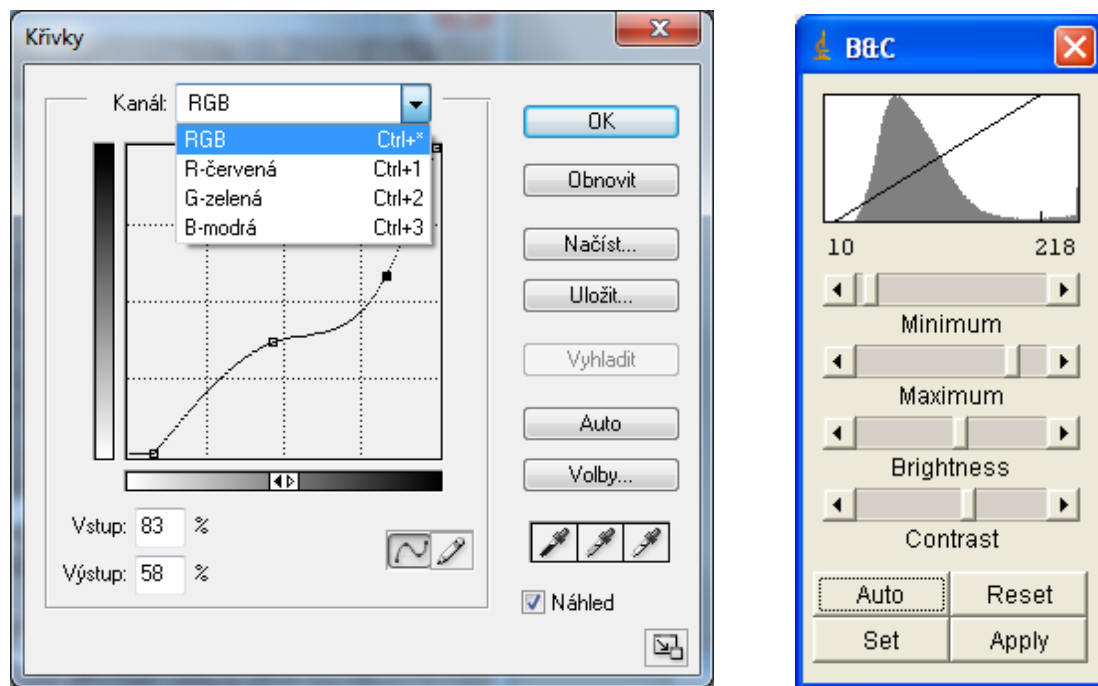


**Obr. 2.4.1:** Typický histogram obrazu s příliš nízkým (nahore) resp. s příliš vysokým jasnem.

Typické histogramy obrazů s nízkým resp. vysokým jasnem si můžeme prohlédnout na obr. 2.4.1. Malý rozptyl histogramu (úzký histogram) signalizuje nízký kontrast obrazu. Z histogramů na obr 2.3.3. je např. zřejmé, že jas snímku feritu (tamtéž) je nízký. Snímek má rovněž nízký globální kontrast.

Úpravu jasu a kontrastu provádíme transformací histogramu pomocí tzv. transformační funkce. Na obr. 2.4.2 vidíme zcela obecnou transformační funkci programu Adobe Photoshop (vlevo) a lineární transformační funkci programy ImageJ (vpravo)





**Obr. 2.4.2:** Editace transformační funkce v prostředí Adobe Photoshop (vlevo) a ImageJ (vpravo)

**Jádro, nosič a paty transformační funkce:** Definičním oborem i oborem hodnot transformační funkce  $f(x)$  je interval  $\langle 0; D \rangle$ . Množinu hodnot  $x$ , pro které je  $f(x) > 0$ , nazýváme **nosičem**, množinu hodnot, pro které je  $f(x) = D$  **jádrem**. Interval  $\langle 0; a \rangle$ , ve kterém je  $f(x) = 0$ , nazýváme **dolní** nebo též **černou patou**, interval  $\langle b; D \rangle$ , ve kterém je  $f(x) = 0$ , nazýváme **horní** nebo též **bílou patou**.

Neprázdné paty jsou známkou toho, že hodnoty pixelů obrazu nevyužívají celý dynamický rozsah, který poskytuje daný grafický formát, a obraz tedy má nižší kontrast.

**Úprava jasu:** Připomeňme, že jas obrazu je definován jako integrál  $B = \int_{\lambda_1}^{\lambda_2} I(\lambda) d\lambda$ . Jestliže tedy chceme zvýšit (snížit) celkový jas obrazu, musíme úpravou transformační funkce zvýšit (snížit) hodnotu tohoto integrálu. Chceme-li celkový jas při úpravě transformační funkce  $f(x)$  na funkci  $g(x)$  zachovat, musí platit

$$B = \int_0^D f(x) dx = \int_0^D g(x) dx \quad (2.4.1)$$

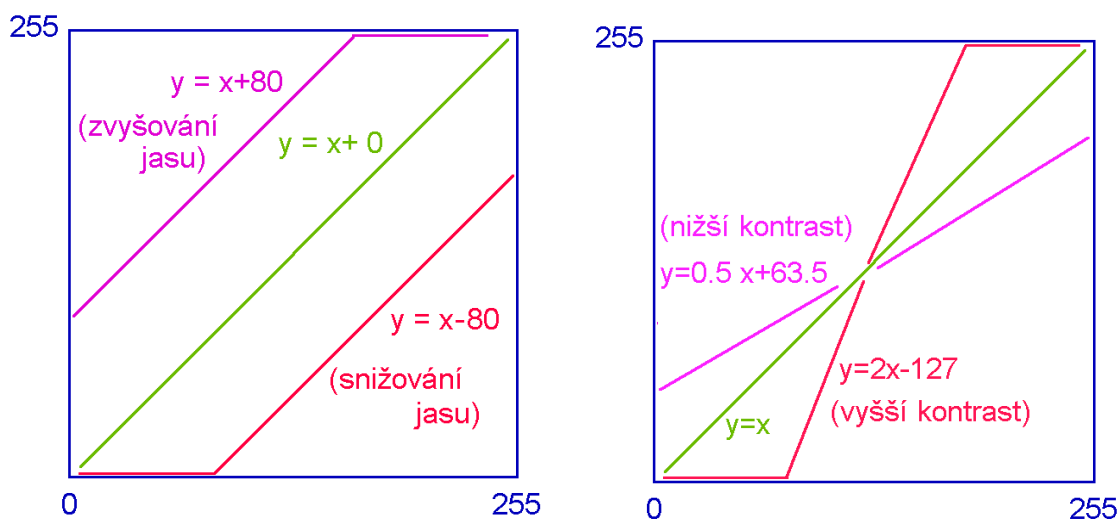
To je obecně zvláště při ruční editaci velmi obtížné, v řadě speciálních případů toho ale lze dosáhnout poměrně snadno (viz dále).

**Úprava kontrastu:** změna kontrastu znamená změnu rozdílů jasu mezi jednotlivými pixely obrazu. Tato změna se projeví změnou šířky histogramu. Změnu kontrastu tedy dosáhneme obecně tak, že budeme zvyšovat nebo zmenšovat nosič transformační funkce  $f(x)$ , tj. upravíme ji na tvar  $f(k \cdot x + q)$ . Při této úpravě zůstane zachována hodnota  $q$ , pro všechny



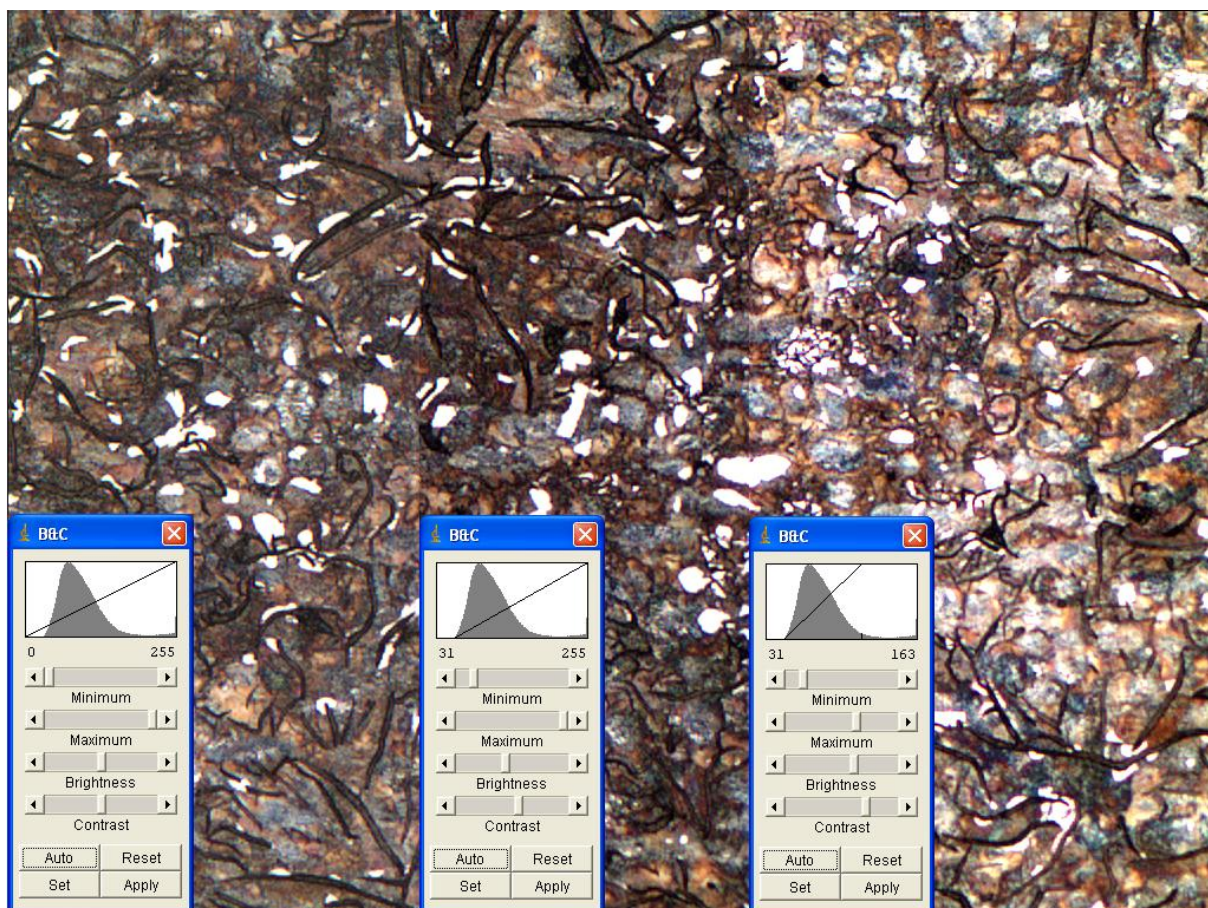
hodnoty  $x \neq q$  se  $k$ -krát změni absolutní hodnota  $|q - x|$ . Pro  $k > 1$  se zvětší (kontrast se zvýší), pro  $0 < k < 1$  se zmenší (kontrast se snižuje).

Úpravu jasu a kontrastu pomocí lineární funkce je ilustrováno na obr. 2.4.4. Úpravu jasu (viz obr. 2.4.4. vlevo) provádíme pomocí funkce  $f(x) = x + q$  změnou aditivní konstanty  $q$ . Při této úpravě musí být zachován obor hodnot, transformační funkce tak musí být ořezána buď shora (zvyšování jasu), anebo zdola (snižování jasu). Nemá-li při této úpravě dojít ke ztrátě informace, nesmí toto ořezání přesáhnout paty histogramu. V tom případě zůstane zachován i celkový kontrast obrazu. Úpravu kontrastu provádíme pomocí funkce  $f(x) = kx + q$  změnou směrnice  $k$ . I při této úpravě musí být v zájmu zachování oboru hodnot transformační funkce ořezána. Pokud nemá dojít ke ztrátě informace, pak ani zde nesmějí tyto ořezy přesáhnout paty histogramu. Nemá-li dojít při změně kontrastu ke změně jasu, musí se měnit i aditivní konstanta  $q$  tak aby zůstal zachován integrál (2.4.1), tj. obsah plochy „pod přímkou“. Je tedy třeba, aby graf transformační funkce stále procházel bodem  $[\frac{D}{2}; \frac{D}{2}]$  tak, jak ilustruje obr. 2.4.3 vpravo.

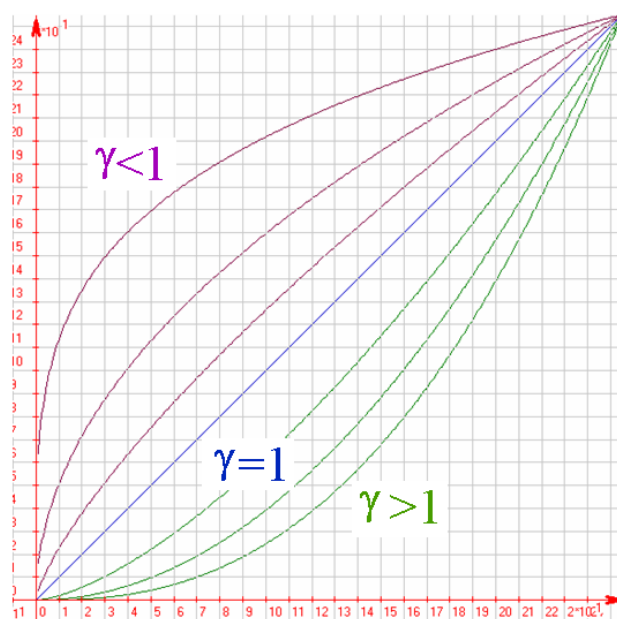


**Obr. 2.4.3:** Úprava lineární transformační funkce. Vlevo úprava jasu při stálém kontrastu, vpravo úprava kontrastu při stálém jasu. Přesáhnou-li paty transformační funkce paty histogramu, dochází ke ztrátě informace a ke změnám jasu i kontrastu současně.

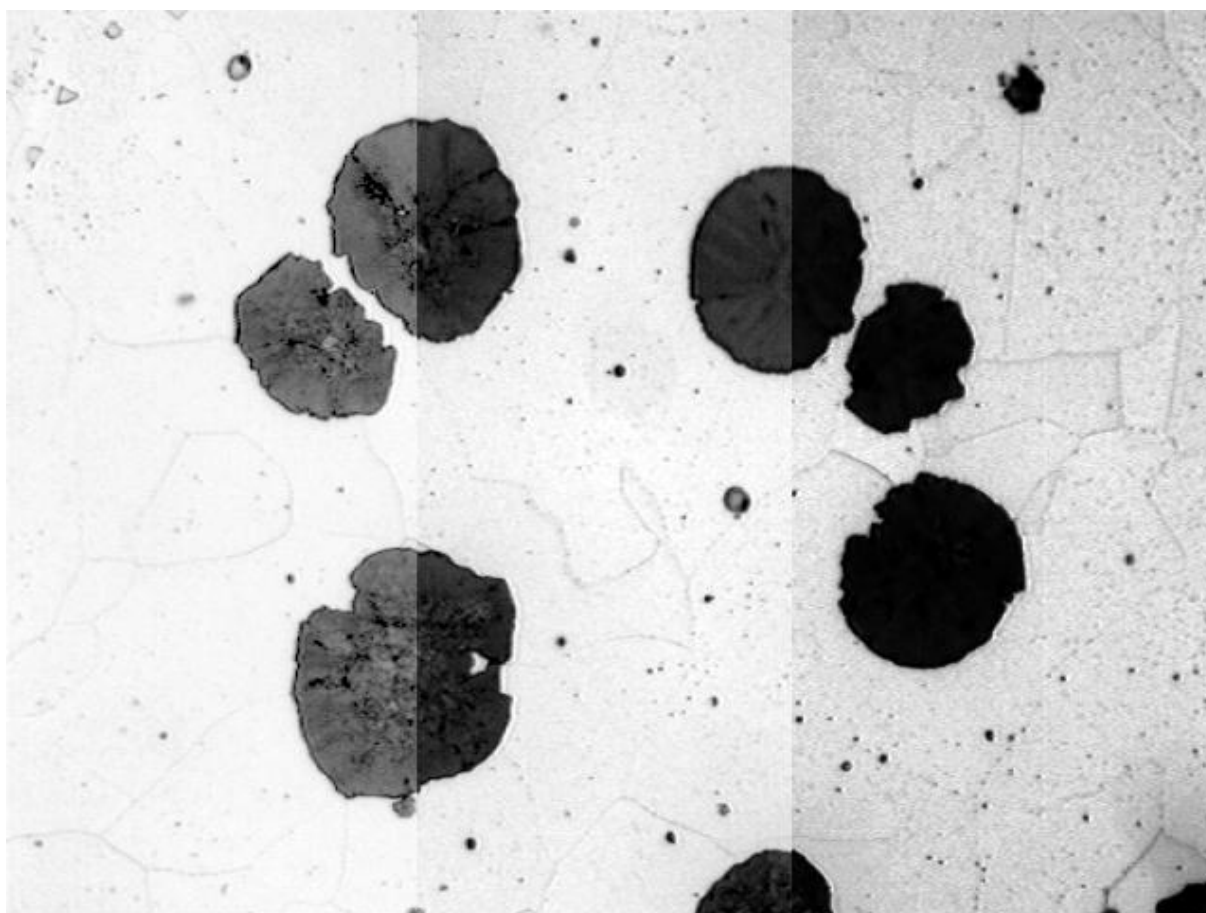
**Gamma korekce:** je transformace histogramu pomocí funkce  $f(x) = x^{\gamma D^{-1}}$ , jejíž proběh je pro různé hodnoty  $\gamma$  na obr. 2.4.5. Z něho je zřejmé, že funkce zachová nosič i jádro histogramu, tj. transformací pomocí gama funkce nedojde ke ztrátě informace (až na případné zaokrouhlování hodnot pixelů). Z ploch, které funkce pro různé hodnoty  $\gamma$  omezuje, je zřejmé, že dochází k celkové změně jasu obrazu – pro  $\gamma > 1$  se snižuje, pro  $\gamma < 1$  zvyšuje. Gama funkce mění lokálně kontrast i jas – při  $\gamma < 1$  se zvyšuje kontrast a jas tmavých částí a snižuje se kontrast a mírně zvyšuje jas částí světlých, při  $\gamma > 1$  se snižuje kontrast a jas tmavých částí a zvyšuje se kontrast a mírně se snižuje jas částí světlých. Účinek gama korekce ilustruje obr. 2.4.6.



**Obr. 2.4.4:** Úprava kontrastu a jasu snímku feritu. Vlevo původní snímek, uprostřed - odstraněním černé paty histogramu se zvýšil celkový kontrast (zvýšila se směrnice transformační přímky) a snížil se celkový jas (snížil se obsah plochy pod přímkou). Vpravo další zvyšování kontrastu za cenu ztráty informace v úrovních vysokého jasu – s ořezáváním transformační funkce dochází zároveň k ořezávání histogramu.



**Obr. 2.4.5:** Transformační funkce  $f(x) = x^{\gamma D^{-1}}$  pro různé hodnoty  $\gamma$ .



**Obr. 2.4.6:** Mikrosnímek litiny upravovaný gama korekcí. Uprostřed původní snímek ( $\gamma = 1$ ), vlevo  $\gamma = 0,5$ ; vpravo  $\gamma = 2$ ,

## 2. 5 Ekvalizace histogramu

Jak bylo řečeno v úvodu minulé kapitoly, jas ideálního obrazu by měl mít rovnoměrné rozložení, sloupce jeho histogramu by tedy měly být v ideálním případě v celém dynamickém rozsahu stejně vysoké. Tohoto ideálu sice prakticky není možné dosáhnout kvůli diskrétní povaze obrazu, lze se však k němu značně přiblížit pomocí techniky ekvalizace (vyrovnání) histogramu.

Úkolem ekvalizace je obraz jasově vyvážit. Tohoto jasového vyvážení docílíme tak, že zpracováváný pixel nahradíme pixelem o hodnotě  $f_i = \text{round}(D \cdot c_i)$ , kde  $c_i$  je relativní kumulativní četnost zpracováváného pixelu a  $D$  je dynamický rozsah obrazu. Po ekvalizaci je program podstatně lépe jasově vyvážen, což je většinou jasně patrné pouhým pohledem. Histogram ekvalizovaného obrazu je podstatně širší než původní histogram, v ideálním případě by měl být konstantní.

O kvalitě ekvalizace většinou lépe vypovídá tzv. kumulativní histogram, který zachycuje **kumulativní četnosti** – součty četností pixelů s jasnem menším nebo rovným jasem zpracováváného pixelu. Tento histogram je neklesající a u jasově nevyváženého obrazu se znatelně odlišuje od přímky (viz obr.2.5.1 vlevo). Po ekvalizaci by se měl přiblížit přímce  $y = x$ . K tomuto ideálu je ve většině praktických situací podstatně blíže než histogram ke své ideální konstantě (viz obr.2.5.1 vpravo).





**Obr. 2.5.1:** Mikrosnímek feritu upravovaný ekvalizací histogramu. Vlevo původní snímek s původním histogramem a kumulovaným histogramem, vpravo ekvalizovaný snímek s ekvalizovanými histogramy.

## 3 Aritmetické a logické operace s obrazy

### 3.1 Aritmetické operace

Hodnoty pixelů obrazu jsou přirozená čísla, se kterými je možné provádět základní aritmetické operace – sčítání, odčítání, násobení a dělení a dokonce i operace logické – konjinkci, disjunkci, implikaci a negaci. Protože v „klasické“ výrokové logice máme k dispozici jen dvě pravdivostní hodnoty – 0;1 – vyžadují si logické operace s obrazy zobecnění této dvojhodnotové logiky na logiku vícehodnotovou. Výsledky logických operací s obrazy ve vícehodnotové logice jakož i výsledky aritmetických operací s obrazy mají v analýze obtazu zajímavá využití. Před dalším čtením je třeba si uvědomit, že i když běžně

mluvíme o obrazových maticích, následující operace (zejména součin a podíl) nelze chápat jako operace maticové. Jsou to operace definované „prvek po prvku“. Budeme je pro jednoduchost definovat jen pro obrazy stejného rozlišení. V celém dalším textu značí  $a_{ij}$ ,  $b_{ij}$ ;  $c_{ij}$  pořadě hodnoty pixelů obrazů  $A; B; C$  a předpokládáme, že  $i = 0; 1; \dots; W-1$ ;  $j = 0; 1; \dots; H-1$ .

Specifikem aritmetických operací s obrazy je skutečnost, že jejich (nemodifikovaný) výsledek nemusí být v intervalu dynamického rozsahu obrazu  $\langle 0; D \rangle$ . Tyto operace je tedy třeba modifikovat

#### **Součet obrazů:**

**Kompresní:** Obraz  $C$  je kompresním součtem obrazů  $A; B$  právě tehdy, když

$$c_{ij} = \text{round} \left[ \frac{a_{ij} + b_{ij}}{2} \right]$$

**Ořezaný:** Obraz  $C$  je ořezaným součtem obrazů  $A; B$  právě tehdy, když

$$c_{ij} = \max \left\{ 0; \min \left\{ D; a_{ij} + b_{ij} - d \right\} \right\}$$

**Lineární kombinace obrazů:** Obraz  $C$  je lineární kombinací obrazů  $A; B$  právě tehdy, když

$$c_{ij} = \text{round} (r \cdot a_{ij} + s \cdot b_{ij}); \quad r + s = 1$$

Kompresní součet je tedy vlastně aritmetickým průměrem obrazů a je speciálním případem lineární kombinace obrazů. Ořezaný součet spočte hodnotu  $a_{ij} + b_{ij} - d$ , kde konstantu  $d$  (displacement) zadává uživatel, tuto hodnotu v případě potřeby ořízne zdola nulou a shora dynamickým rozsahem obrazu. Displacement by měl být volen tak, aby tento ořez minimalizoval. Kompresní součet se používá především k omezení aditivního šumu tak, jak o tom bude pojednáno níže, dále ke skládání obrazů z různých částí spektra, případně může simulovat průhlednost obrazů (viz obr. 3.1.1). Lineární kombinace umožňuje dávat sčítancům různou váhu, a tak simulovat např. různou průhlednost jednotlivých obrazů. Ořezaný součet může sloužit ke zvýšení kontrastu u málo kontrastních obrazů (3.1.2).

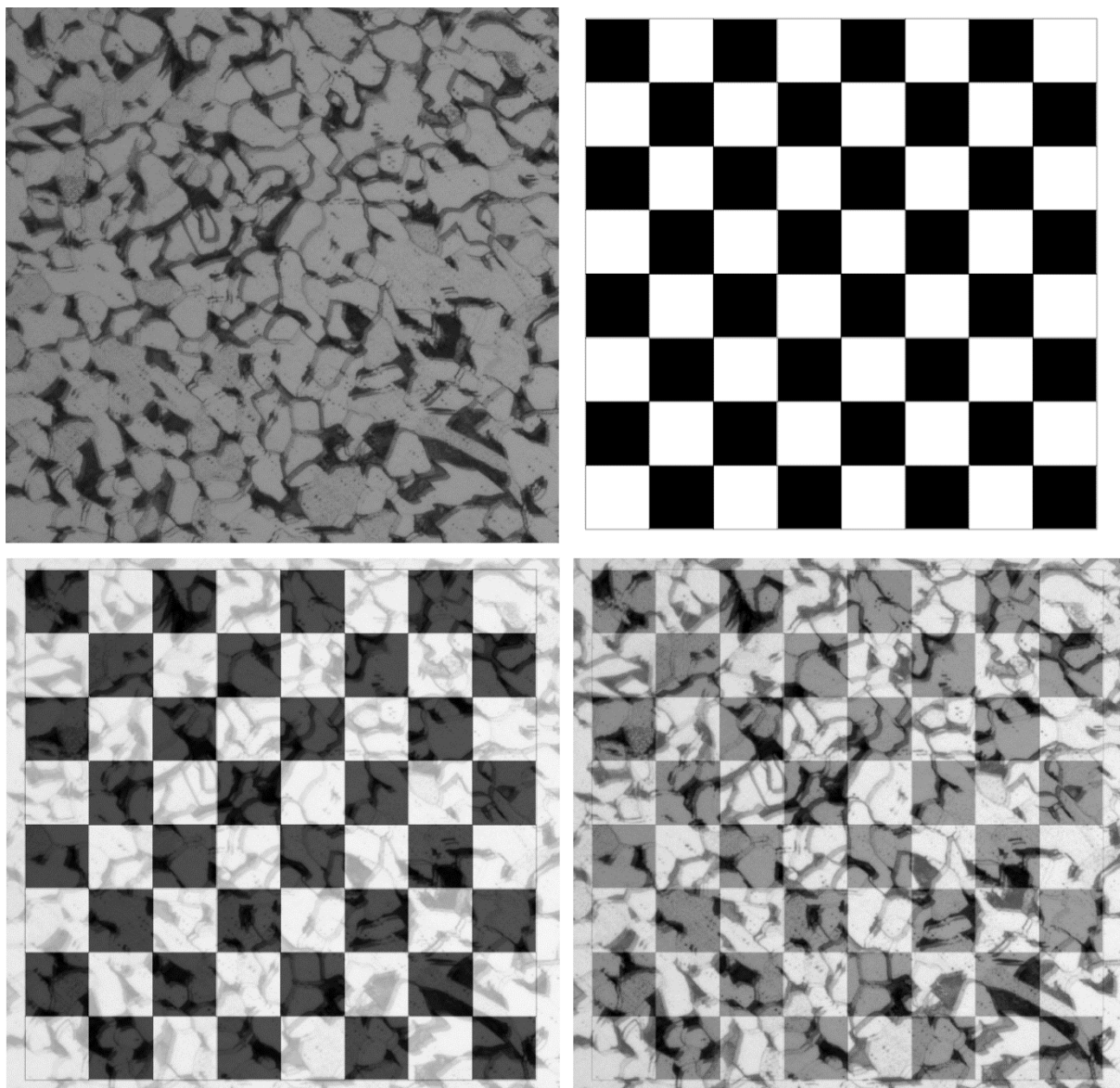
#### **Rozdíl obrazů:**

**Kompresní:** Obraz  $C$  je kompresním rozdílem obrazů  $A; B$  právě tehdy, když

$$c_{ij} = \text{round} \left[ \frac{a_{ij} - b_{ij} + D}{2} \right]$$

**Ořezaný:** Obraz  $C$  je ořezaným rozdílem obrazů  $A; B$  právě tehdy, když

$$c_{ij} = \max \left\{ 0; \min \left\{ D; a_{ij} - b_{ij} + d \right\} \right\}$$

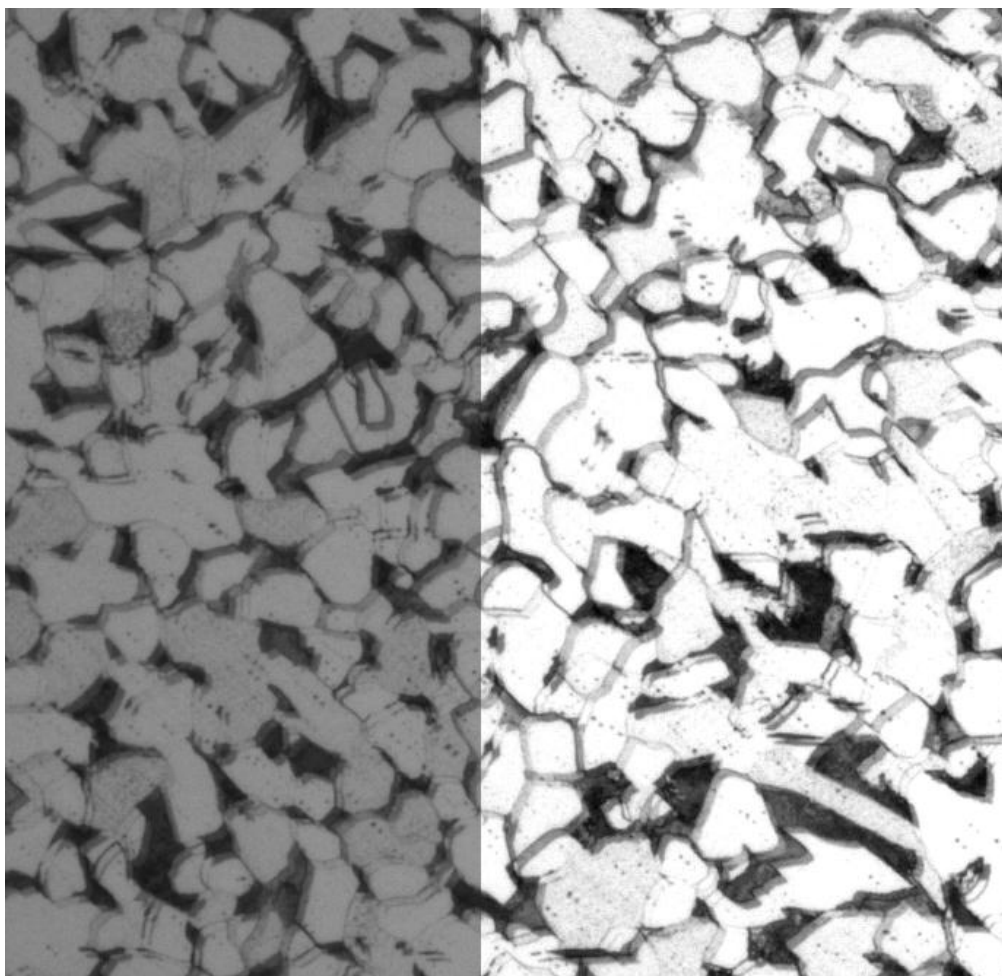


**Obr. 3.1.1:** Rozdělení snímku oceli (vlevo nahoře) na oblasti různého jasu poloprůhlednou šachovnicí (vpravo nahoře), simulovanou kompresním součtem – obrazy jsou stejně průhledné (vlevo dole) a lineární kombinací pro  $r = 0.8$   $s = 0.2$  (vpravo dole).

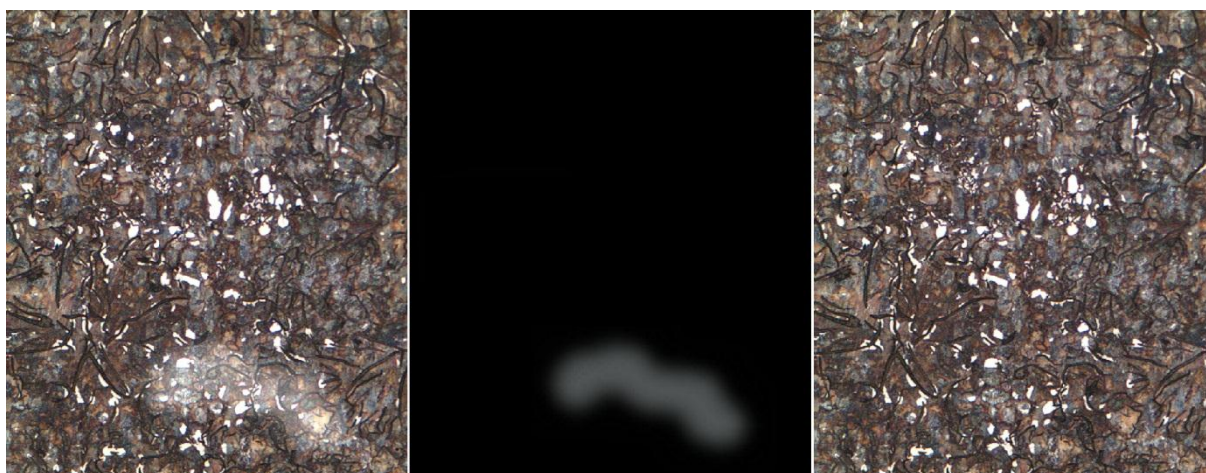
Ořezaný rozdíl se používá ke korekci vad v obraze zaviněných různou citlivostí čipů. Na snímku se to projeví oblastmi se zvýšeným jasnem. Ke korekci této vady je třeba pořídit ke snímku ještě snímek černé předlohy (Flat field), na které se tato vada projeví stejně. Odečtením snímku tmavého pozadí od snímku určeného ke korekci je vada korigována.

Právě k identifikaci „neviditelných“ vad je vhodný kompresní rozdíl tak, jak ilustruje obr. 3.1.4. Zde vidíme dva zdánlivě zcela identické snímky pořízené dvěma různými zařízeními (nahore) a chceme zjistit, zda jsou snímky skutečně identické, a pokud ne, kde a jak se liší. Výsledkem kompresního rozdílu těchto dvou snímků je šedá plocha se (zdánlivě) konstantním jasnem  $\frac{1}{2}$  (vlevo dole). Pokud se však snímky liší, má tato plocha nepatrně tmavší části tam, kde je první snímek tmavší než druhý a nepatrně světlejší místa tam, kde je tomu naopak. Tyto rozdíly můžeme následně zviditelnit zvýšením kontrastu (vpravo dole).





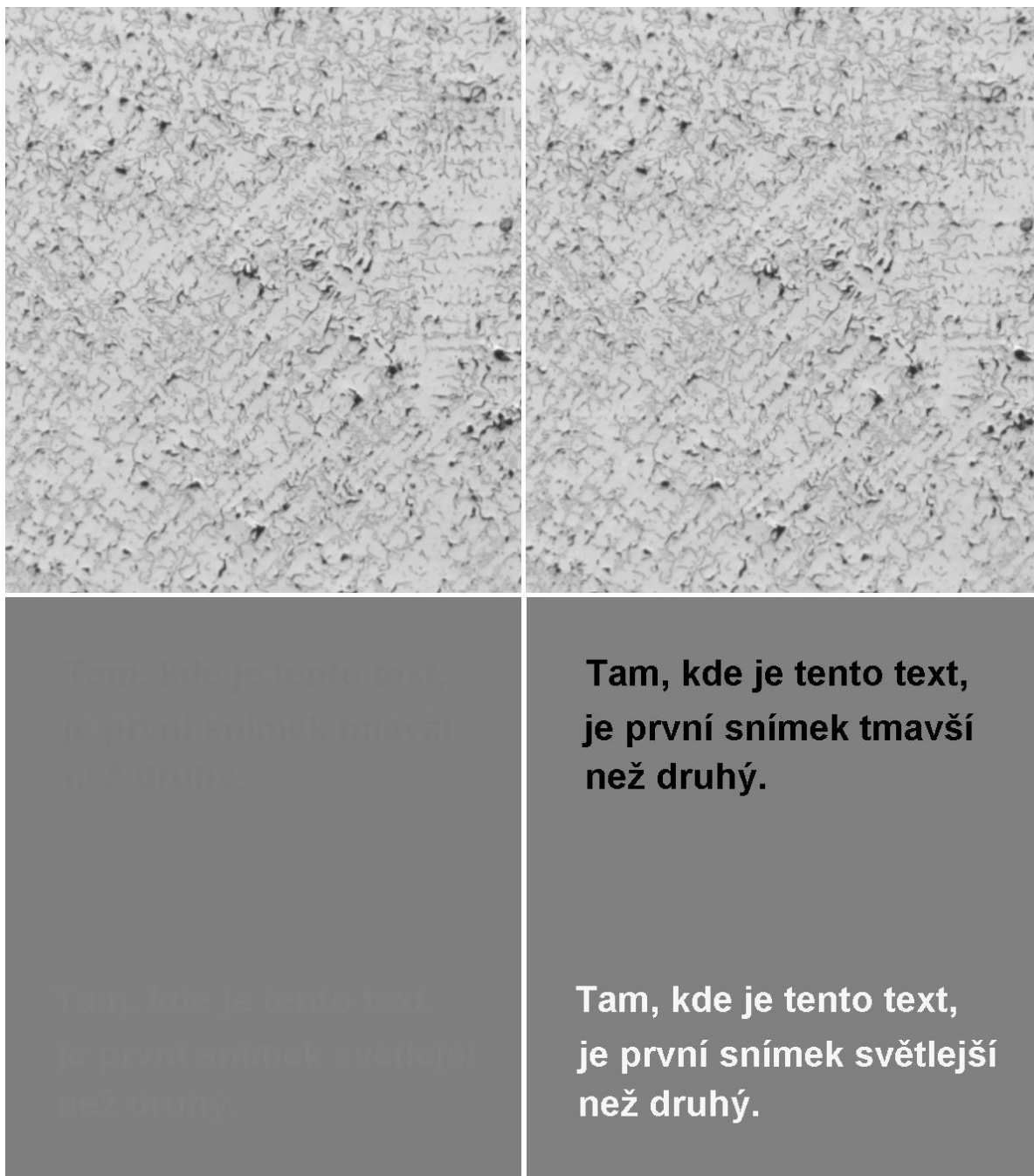
**Obr. 3.1.2:** Zvýšení kontrastu snímku oceli ořezaným součtem. Obraz A (vlevo) byl sečten sám se sebou (vpravo). Použit displacement  $d = 40$ .



**Obr. 3.1.3:** Korekce různé citlivosti čipů pomocí rozdílu korigovaného snímku feritu (vlevo) a snímku černého pozadí (uprostřed). Výsledkem je zkorigovaný snímek (vpravo).

V praxi se sice tato vada neprojevuje tak drasticky, jak je z ilustračních důvodů vymodelováno na obr. 3.1.3, pokud jsou však snímky pořizovány k fotometrickým účelům, je třeba tyto korekce provádět prakticky vždy, i když jsou tyto vady pouhým okem neviditelné.





**Obr. 3.1.4:** Korekce různé citlivosti čipů pomocí rozdílu korigovaného snímku feritu (vlevo) a snímku černého

**Součin obrazů:**

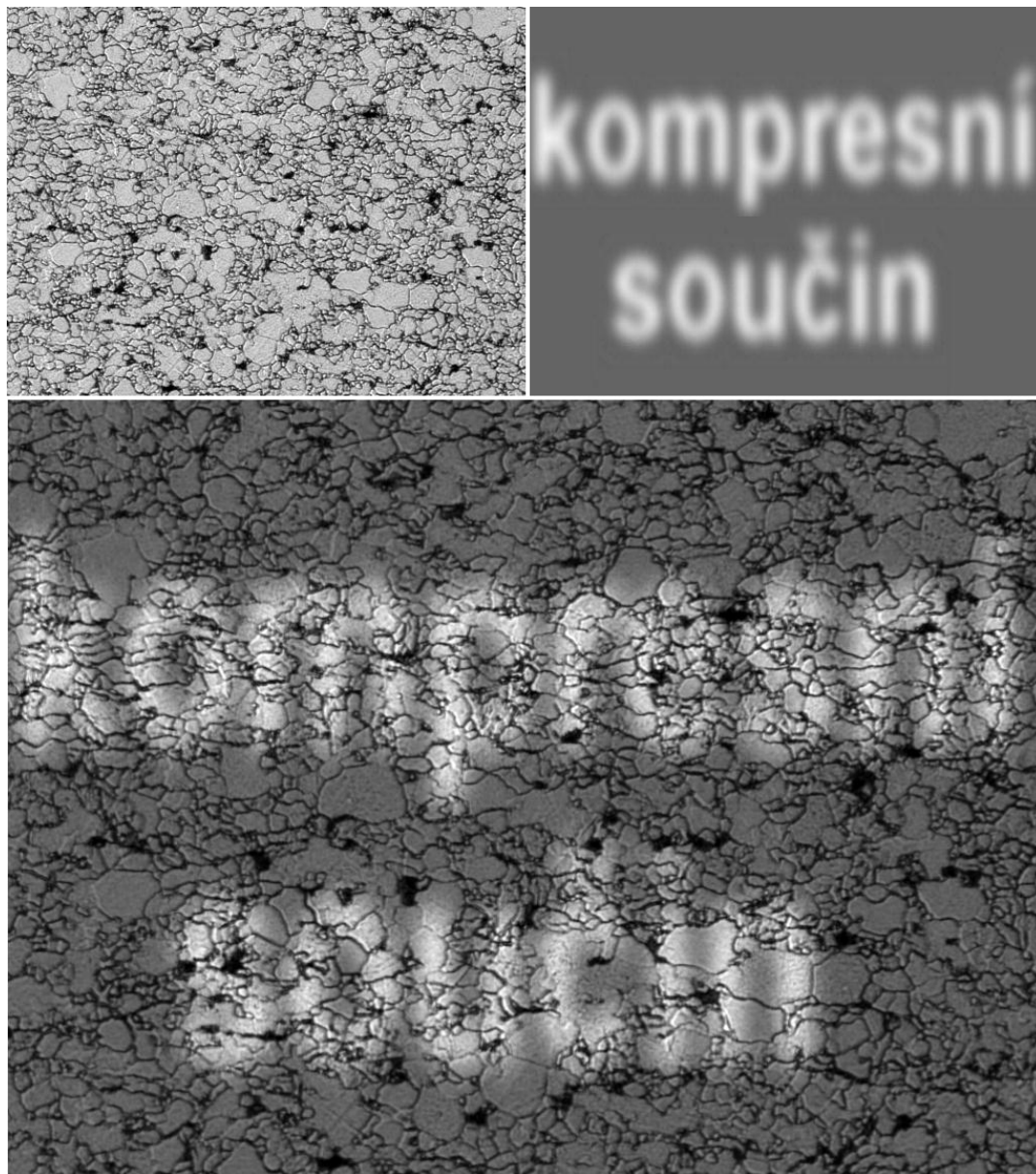
**Kompresní:** Obraz  $C$  je kompresním součinem obrazů  $A; B$  právě tehdy, když

$$c_{ij} = \text{round} \left[ \frac{a_{ij} \cdot b_{ij}}{D} \right]$$

**Ořezaný:** Obraz  $C$  je ořezaným součinem obrazů  $A; B$  právě tehdy, když

$$c_{ij} = \max \{ D; a_{ij} \cdot b_{ij} \}$$

Součin obrazů je vhodný k odstranění (odmaskování) nežádoucích částí snímku. Obraz *A* obsahuje zpracovávaný snímek, obraz *B* (maska) obdahuje informace o tom, které body ponechat a které odstranit. Vhodnou maskou lze docílit i lokální změny kontrastu tak, jak ilustruje obr. 3.1.5.



**Obr. 3.1.6:** Lokální změny kontrastu dosažené vhodnou maskou.

### 3.2 Základy vícehodnotové logiky

Kromě dvouhodnotové výrokové logiky, jejíž základy známe ze střední školy, nachází v řadě aplikací uplatnění i logika vícehodnotová. Jejím základním rysem je skutečnost, že výrokům nepřisuzuje jen dvě pravdivostní hodnoty (pravda a nepravda, resp. 0 a 1), ale těchto pravdivostních hodnot je více (může jich být dokonce nekonečně mnoho). Vícehodnotových konstrukcí existuje celá řada, jejich základním společným rysem však je, že dvouhodnotová logika se svými spojkami je jejich speciálním případem. Nejrozšířenějším vícehodnotovým systémem je logika, jejímž autorem je polský matematik **Jan Lukasiewicz** (1878 - 1956).

Lukasiewiczova  $n$ -hodnotová logika přiřazuje každému výroku  $p$  právě jedno z celkem  $n$  možných pravdivostních ohodnocení  $V(p) \in \langle 0; 1 \rangle$ . Výroky  $p, q$  můžeme vzájemně spojovat logickými spojkami  $\wedge; \vee; \Rightarrow; \neg$  známými z logiky dvouhodnotové. Navíc přibývá spojka  $\&$  (tzv. odvážná konjunkce), která ve dvouhodnotovém případě splývá s konjunkcí. Pravdivostní ohodnocení složených výroků je definováno takto:

Konjunkce	$V(p \wedge q) = \min\{V(p); V(q)\}$
Disjunkce	$V(p \vee q) = \max\{V(p); V(q)\}$
Implikace	$V(p \Rightarrow q) = \min\{1; 1 - V(p) + V(q)\}$
Odvážná konjunkce	$V(p \& q) = \max\{0; V(p) + V(q) - 1\}$
Negace	$V(\neg p) = 1 - V(p)$

### 3.3 Logické operace s obrazy

Každý pixel obrazu můžeme chápat jako výrok ve smyslu  $D$ -hodnotové logiky. Pixelu s hodnotou  $a_{ij}$  můžeme přisoudit pravdivost  $a_{ij}/D$ . Dva obrazy  $A = (a_{ij}); B = (b_{ij})$  se stejným rozlišením  $W \times H$  tak můžeme chápat jako množinu  $W \times H$  výroků s pravdivostmi  $a_{ij}/D$  resp.  $b_{ij}/D$ , logické operace s těmito obrazy pak můžeme chápat přesně ve smyslu předchozích definic. Abychom však mohli výsledek vizualizovat, je třeba interval  $\langle 0; 1 \rangle$  následně transformovat na interval  $\langle 0; 255 \rangle$ , tj. pravdivostní hodnotu násobit číslem  $D - 1$ . Logické operace s obrazy lze tedy jednoduše definovat takto:

Pro obrazy se stejným rozlišením a stejným dynamickým rozsahem definujeme tyto logické operace:

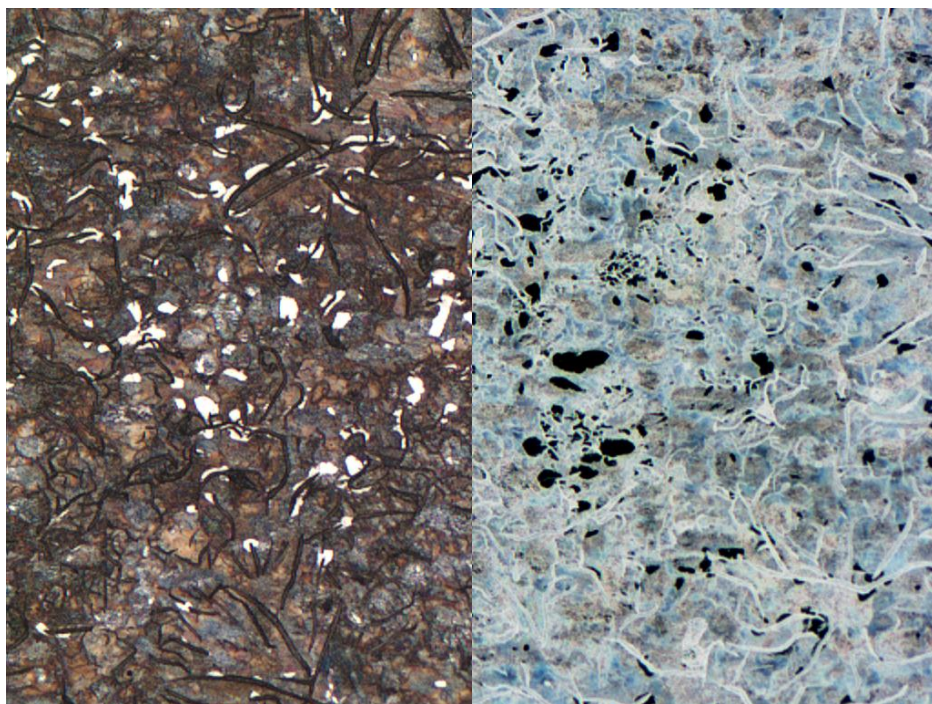
Obraz  $A' = (a'_{ij})$  je **negací** obrazu  $A = (a_{ij})$  právě tehdy, když pro každý pixel  $[i; j]$  s hodnotou  $a'_{ij}$  platí

$$a'_{ij} = D - 1 - a_{ij}$$

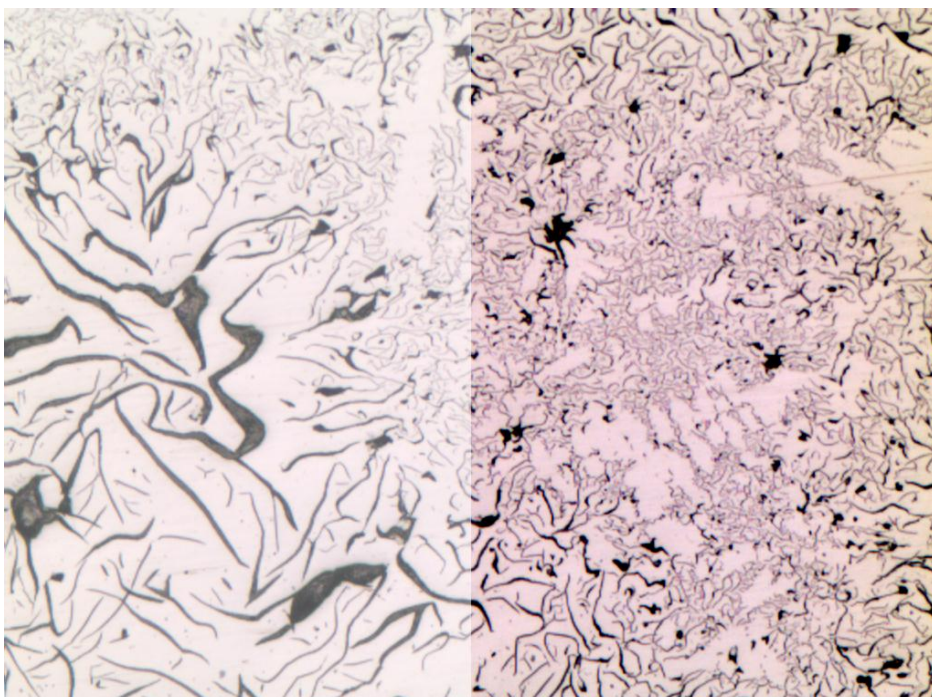
Necht'  $A = (a_{ij}); B = (b_{ij})$  jsou obrazy se stejným rozlišením a dynamickým rozsahem  $D$ . Pak obraz  $C = (c_{ij})$  nazýváme

<b>konjunkcí</b> $A; B$ právě tehdy, když	$c_{ij} = \min\{a_{ij}; b_{ij}\}$
<b>disjunkcí</b> $A; B$ právě tehdy, když	$c_{ij} = \max\{a_{ij}; b_{ij}\}$
<b>implikací</b> $A; B$ právě tehdy, když	$c_{ij} = \min\{D; D - a_{ij} + b_{ij}\}$
<b>odvážnou konjunkcí</b> $A; B$ právě tehdy, když	$c_{ij} = \max\{0; a_{ij} + b_{ij} - D\}$

Negací obrazu můžeme využít k sestrojení negace obrazu – viz obr. 3.3.1. Konjunkce a disjunkce umožňuje vyrovnat případnou jasovou nevyváženost dvou snímků téže scény. Odvážnou konjunkci můžeme využít ke zvýšení kontrastu obrazu s vysokým jasnem (jako obrazy  $A; B$  použijeme tentýž obraz)



**Obr. 3.3.1:** Negativ obrazu sestrojený jako logická negace



**Obr. 3.3.2:** Zvýšení kontrastu obrazu odvážnou konjunkcí



## 4 Geometrické transformace obrazu

V kapitole 1.1 jsme se zabývali zobrazením v euklidovském a projektivním prostoru. V této kapitole se budeme podrobněji zajímat o popis zobrazení v rovině, a to nejen zobrazení spojitého geometrického útvaru, ale především o zobrazení útvaru diskrétního, kterým je i digitální obraz. Nejdříve ovšem popíšeme zobrazení bodu v projektivní rovině.

### 4.1 Zobrazení v projektivní rovině

V kpt. 1. 1 jsme uvedli obecnou rovnici (1) promítání v projektivním prostoru a dále rovnici (2) zobrazení, které zachovává vlastní a nevlastní body. V této kapitole nás budou zajímat zobrazení, se kterými se běžně setkáváme při manipulaci s obrazy – posouvání, otáčení, změna měřítka, osová souměrnosti. Všechna tato zobrazení zachovávají vlastní a nevlastní body. Každý digitální obraz obsahuje jen vlastní body a jeho vlastní bod  $(x \ y \ 1)^T$  se vždy zobrazí na vlastní bod  $(x' \ y' \ 1)^T$ . Takové zobrazení je tvaru

$$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \quad (4.1.1)$$

rozeepsáno do souřadnic

$$\begin{aligned} x' &= a_{11}x + a_{12}y + a_{13} \\ y' &= a_{21}x + a_{22}y + a_{23} \end{aligned} \quad (4.1.2)$$

Tvar (4.1.2) bývá většinou „průhlednější“, tvar (4.1.1) je výhodnější v případě skládání zobrazení. Rovnice nejznámějších zobrazení v rovině jsou

Zobrazení	Tvar (4.1.2)	Tvar (4.1.1)	
Osová souměrnost podle osy x	$\begin{aligned} x' &= x \\ y' &= -y \end{aligned}$	$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$	(4.1.3)

Osová souměrnost podle osy y	$\begin{aligned} x' &= -x \\ y' &= y \end{aligned}$	$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$	(4.1.4)
------------------------------	---	--	---------

Otočení kolem počátku soustavy

$$\begin{aligned} x' &= x \cdot \cos \alpha - y \cdot \sin \alpha \\ y' &= x \cdot \sin \alpha + y \cdot \cos \alpha \end{aligned} \quad \begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \begin{pmatrix} \cos \alpha & -\sin \alpha & 0 \\ \sin \alpha & \cos \alpha & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \quad (4.1.5)$$

Posunutí o vektor $\mathbf{v} = (v_1; v_2)$	$\begin{aligned} x' &= x + v_1 \\ y' &= y + v_2 \end{aligned}$	$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & v_1 \\ 0 & 1 & v_2 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$	(4.1.6)
---	--	---	---------

Stejnolehlost s koeficientem $\lambda$	$\begin{aligned} x' &= \lambda x \\ y' &= \lambda y \end{aligned}$	$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \begin{pmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$	(4.1.7)
--	--	---	---------

V programech na úpravu obrazů se osově souměrnosti podle svislé či vodorovné osy většinou nazývají překlopení či převrácení (vertical, horizontal flip), rotace o  $\pm \frac{\pi}{2}$  se skrývají pod příkazy otočit vlevo resp. vpravo (rotate left, right) a stejnolehlost s kladným koeficientem je označována jako změna velikosti (resize). Protože však indexy udávající pozici pixelu v obraze slouží současně jako souřadnice, výše uvedené rovnice k těmto transformacím nestačí. „Překlápění“ obrazu nepotřebuje souřadnicové osy, ale vodorovnou resp. svislou přímku, procházející středem obrazu, tj. bodem  $S = [\frac{w}{2}; \frac{h}{2}]$ . Otáčení obrazu musí proběhnout tak, aby souřadnice žádného bodu nebyly po otočení záporné. Zobrazení (4.1.3) – (4.1.7) je tedy třeba skládat. Matici složeného zobrazení dostaneme tak, že vynásobíme matice všech zobrazení, která ke skládání použijeme, a to v obráceném pořadí.

**Příklad – horizontální překlopení** je zobrazení osově souměrnosti s vertikální osou procházející středem obrazu, osa souměrnosti má tedy rovnici  $x = \frac{w}{2}$ . K jeho realizaci potřebujeme nejdříve tuto osu ztotožnit s levým okrajem obrazu, posuneme tedy celý obraz o vektor  $\mathbf{v} = (-\frac{w}{2}; 0)$ . Matice tohoto posunutí je dle (4.1.6) tvaru

$$\mathbf{T} = \begin{pmatrix} 1 & 0 & -\frac{w}{2} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Dále použijeme osovou souměrnost s osou v levém okraji obrazu, tj. osou y. Matice této souměrnosti je – viz (4.1.4)

$$\mathbf{S} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Konečně je třeba obraz posunout „zpět“, tj. o vektor  $-\mathbf{v} = (\frac{w}{2}; 0)$  Matice tohoto posunutí je

$$\mathbf{T}^{-1} = \begin{pmatrix} 1 & 0 & \frac{w}{2} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Označení této matice není náhodné; platí totiž

$$\mathbf{T}^{-1} \cdot \mathbf{T} = \begin{pmatrix} 1 & 0 & \frac{w}{2} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & -\frac{w}{2} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

takže matice  $\mathbf{T}^{-1}$  je vskutku maticí inverzní k  $\mathbf{T}$ . Matici  $\mathbf{P}$  překlopení pak dostaneme vynásobením matic  $\mathbf{T}$ ;  $\mathbf{S}$ ;  $\mathbf{T}^{-1}$ , a to v obráceném pořadí, tj.

$$\mathbf{P} = \mathbf{T}^{-1} \cdot \mathbf{S} \cdot \mathbf{T} = \begin{pmatrix} 1 & 0 & \frac{w}{2} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & -\frac{w}{2} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} =$$

$$= \begin{pmatrix} -1 & 0 & \frac{W}{2} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & -\frac{W}{2} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} -1 & 0 & W \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Bod  $X = (x \ y \ 1)^T$  je tedy tímto překlopením zobrazen na bod  $X' = (x' \ y' \ 1)^T$ , pro který platí

$$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \begin{pmatrix} -1 & 0 & W \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \begin{pmatrix} W - x \\ y \\ 1 \end{pmatrix}$$

**Příklad – otočení:** Otočme obraz o úhel  $\alpha = 30^\circ$  proti směru chodu hodinových ručiček. Pravý horní vrchol obrazu má souřadnice  $(W \ 0 \ 1)^T$ . Při otočení kolem počátku (levý horní vrchol obrazu) o úhel  $\alpha = 30^\circ$  proti směru chodu hodinových ručiček (tedy v kladném smyslu) bude mít pravý horní roh obrazu dle (4.1.5) souřadnice

$$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \begin{pmatrix} \cos \frac{\pi}{6} & -\sin \frac{\pi}{6} & 0 \\ \sin \frac{\pi}{6} & \cos \frac{\pi}{6} & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} W \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} W \cos \frac{\pi}{6} \\ W \sin \frac{\pi}{6} \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{\sqrt{3}}{2} W \\ \frac{1}{2} W \\ 1 \end{pmatrix} \quad (1.4.8)$$

Pozornému čtenáři asi neunikla chyba. Obě souřadnice  $[x'; y'] = [W \cos \alpha; W \sin \alpha]$  otočeného levého horního rohu by podle (1.4.8) měly být kladné. Jestliže si však celou situaci představíme, je však zřejmé, že po otočení bude souřadnice  $y'$  záporná. Problém je v tom, že na střední škole jsme používali výhradně pravotočivou souřadnicovou soustavu. Soustava spojená s obrazem má však osu  $y$  orientovanou opačně a je tedy levotočivá. V této soustavě má orientovaný úhel opačnou orientaci, než na jakou jsme zvyklí. Otáčíme tedy nikoli o úhel  $+30^\circ$ , ale  $-30^\circ$ , tedy

$$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \begin{pmatrix} \cos \left(-\frac{\pi}{6}\right) & -\sin \left(-\frac{\pi}{6}\right) & 0 \\ \sin \left(-\frac{\pi}{6}\right) & \cos \left(-\frac{\pi}{6}\right) & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} W \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} W \cos \left(-\frac{\pi}{6}\right) \\ W \sin \left(-\frac{\pi}{6}\right) \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{\sqrt{3}}{2} W \\ -\frac{1}{2} W \\ 1 \end{pmatrix}$$

Souřadnice  $y'$  v je tedy záporná. Protože záporné souřadnice v obraze nepřipouštíme, musí po této rotaci následovat posunutí o vektor  $\mathbf{v} = (0; -W \sin \alpha)$  (ve směru osy  $x$  není potřeba posouvat). Matice takto složeného zobrazení tedy bude

$$\mathbf{M} = \mathbf{T} \cdot \mathbf{R} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & W \sin \alpha \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} \cos \alpha & -\sin \alpha & 0 \\ \sin \alpha & \cos \alpha & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} \cos \alpha & -\sin \alpha & 0 \\ \sin \alpha & \cos \alpha & W \sin \alpha \\ 0 & 0 & 1 \end{pmatrix}$$

Například bod  $X = (100 \ 100 \ 1)^T$  bude mít v otočeném obraze souřadnice



$$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \begin{pmatrix} \cos \alpha & -\sin \alpha & 0 \\ \sin \alpha & \cos \alpha & W \sin \alpha \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 100 \\ 100 \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{\sqrt{3}}{2} & \frac{1}{2} & 0 \\ -\frac{1}{2} & \frac{\sqrt{3}}{2} & W \cdot \frac{1}{2} \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 100 \\ 100 \\ 1 \end{pmatrix} = \begin{pmatrix} 50 \cdot (\sqrt{3} + 1) \\ 50 \cdot (\sqrt{3} - 1 + W) \\ 1 \end{pmatrix}$$

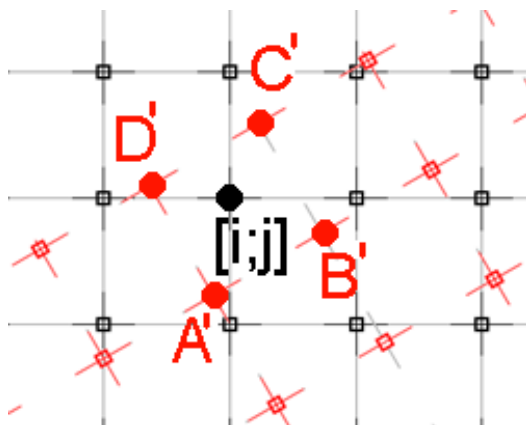
Čtenář teď možná opět protestuje a tentokrát oprávněně. Na obraze jsou přece jen body, které mají celočíselné souřadnice a tyto souřadnice zcela evidentně celočíselné nejsou. Tato kapitola se zabývala zobrazeními v projektivní rovině. Obraz však za rovinu považovat nemůžeme, protože není spojitý, ale diskrétní. O tom, jak budou tyto transformace vypadat v diskrétním případě, pojedná další kapitola.

## 4.2 Transformace diskrétního obrazu

Jak jsme viděli v posledním příkladě, původně celočíselné souřadnice pixelu se mohou geometrickými transformacemi popsanými v kpt. 4. 1. stát neceločíselnými. Ovšem souřadnice bodů transformovaného diskrétního obrazu musejí být opět celočíselné. Proto jsou transformace diskrétního obrazu poněkud složitější, než transformace spojitých geometrických útvarů. Transformace popsané v předchozí kapitole, jsou při transformaci reálného obrazu pouze prvním krokem.

Předpokládejme, že obraz byl podroben transformaci  $T$ . Po tomto prvním kroku známe hodnoty, které by obraz měl mít v bodech  $A' = T(A)$ ;  $B' = T(B)$ ;  $C' = T(C)$ ;  $D' = T(D)$  s neceločíselnými souřadnicemi. My však můžeme přiřadit hodnoty pouze bodům se souřadnicemi celočíselnými. Možnosti, jak to udělat, jsou v podstatě dvě:

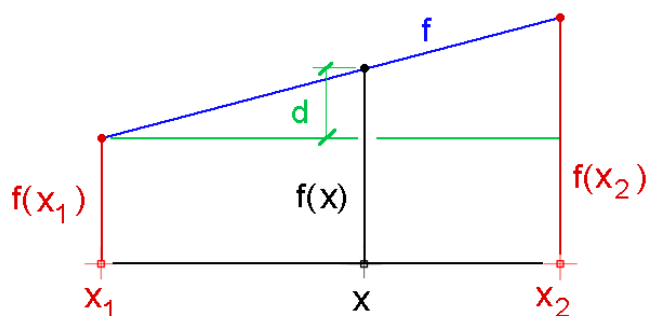
**Metoda nejbližšího souseda:** Pixelu  $[i; j]$  se přiřadí hodnota toho z bodů  $A'; B'; C'; D'$ , který je pixelu  $[i; j]$  nejbližší.



**Obr. 4.2.1:** K problému transformace diskrétního obrazu

Tato metoda je sice nejjednodušší, ale nepřesná a zcela nepoužitelná např. při zvětšování obrazů. Proto se častěji používá tzv. bilineární interpolace. Abychom ji vysvětlili, musíme nejdříve uvést jednodušší interpolaci lineární.

**Lineární interpolace:** Předpokládejme, že známe hodnoty  $f(x_1)$ ;  $f(x_2)$  lineární funkce v bodech  $x_1$ ;  $x_2$  a potřebujeme znát hodnotu  $f(x)$  v bodě  $x$ ;  $x_1 < x < x_2$  - viz obr 4.2.2.



**Obr. 4.2.2:** K lineární interpolaci

Z podobnosti trojúhelníků plyne, že

$$\frac{d}{x - x_1} = \frac{f(x_2) - f(x_1)}{x_2 - x_1} \Rightarrow d = \frac{x - x_1}{x_2 - x_1} \cdot [f(x_2) - f(x_1)]$$

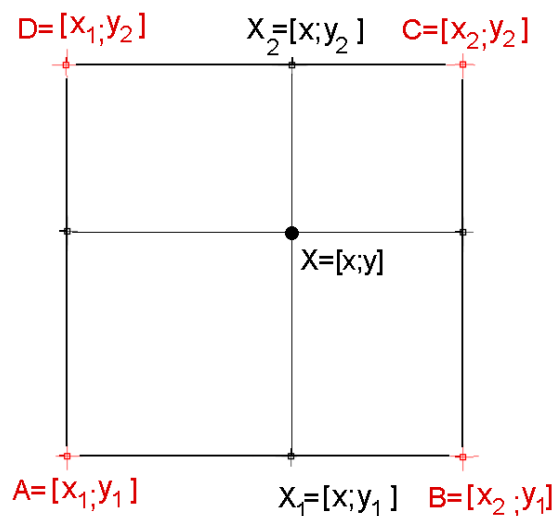
Protože  $f(x) = f(x_1) + d$ , je

$$f(x) = f(x_1) + \frac{x - x_1}{x_2 - x_1} \cdot [f(x_2) - f(x_1)]$$

což po úpravě dává

$$f(x) = \frac{x - x_1}{x_2 - x_1} \cdot f(x_2) + \frac{x_2 - x}{x_2 - x_1} \cdot f(x_1) \quad (4.2.1)$$

**Bilineární interpolace:** řeší na podobném principu obecný případ, kdy se interpolovaný pixel nachází uvnitř čtverce, jehož vrcholy tvoří body  $A'; B'; C'; D'$  se známými hodnotami. V tom případě musíme provést lineární interpolaci ve dvou směrech (odtud název bilineární interpolace), ale lineární interpolaci je třeba provést celkem třikrát. Čtverec  $A'; B'; C'; D'$  včetně vnitřního pixelu  $[i; j]$  nejprve zobrazíme tak, aby jeho strany byly vodorovné resp. svislé. Nejlépe tedy transformací  $T^{-1}$  inverzní k  $T$ . Body  $A'; B'; C'; D'$  přejdou v původní  $A; B; C; D$  a bod  $T^{-1}[i; j]$  označme  $X = [x; y]$  (viz obr. 4.2.3)



**Obr. 4.2.3:** K bilineární interpolaci

Hodnoty obrazu v bodech  $X_1; X_2$  interpolujeme dle (4.2.1) a obdržíme

$$f(x; y_1) = \frac{x - x_1}{x_2 - x_1} \cdot f(x_2; y_1) + \frac{x_2 - x}{x_2 - x_1} \cdot f(x_1; y_1) \quad (4.2.2)$$

$$f(x; y_2) = \frac{x - x_1}{x_2 - x_1} \cdot f(x_2; y_2) + \frac{x_2 - x}{x_2 - x_1} \cdot f(x_1; y_2) \quad (4.2.3)$$

Hledanou hodnotu  $f(X) = f(x; y)$  v bodě  $X = [x; y]$  dostaneme opět podle (4.2.1) interpolací bodů  $X_1; X_2$ :

$$f(x; y) = \frac{y - y_1}{y_2 - y_1} \cdot f(x; y_2) + \frac{y_2 - y}{y_2 - y_1} \cdot f(x; y_1)$$

Po dosazení ze (4.2.2), (4.2.3) a úpravě dostaneme

$$\begin{aligned} f(x; y) = & \frac{(x_2 - x)(y_2 - y)}{(x_2 - x_1)(y_2 - y_1)} \cdot f(x_1; y_1) + \frac{(x - x_1)(y_2 - y)}{(x_2 - x_1)(y_2 - y_1)} \cdot f(x_2; y_1) + \\ & + \frac{(x_2 - x)(y - y_1)}{(x_2 - x_1)(y_2 - y_1)} \cdot f(x_1; y_2) + \frac{(x - x_1)(y - y_1)}{(x_2 - x_1)(y_2 - y_1)} \cdot f(x_2; y_2) \end{aligned} \quad (4.2.4)$$

Jestliže interpolujeme hodnotu pixelu, je  $x_2 - x_1 = y_2 - y_1 = 1$  a vzorec (4.2.4) se zjednoduší na

$$\begin{aligned} f(x; y) = & (x_2 - x)(y_2 - y) \cdot f(x_1; y_1) + (x - x_1)(y_2 - y) \cdot f(x_2; y_1) + \\ & + (x_2 - x)(y - y_1) \cdot f(x_1; y_2) + (x - x_1)(y - y_1) \cdot f(x_2; y_2) \end{aligned}$$

nebo v maticové podobě

$$f(x; y) = \begin{pmatrix} x_2 - x & x - x_1 \end{pmatrix} \begin{pmatrix} f(x_1; y_1) & f(x_1; y_2) \\ f(x_2; y_1) & f(x_2; y_2) \end{pmatrix} \begin{pmatrix} y_2 - y \\ y - y_1 \end{pmatrix} \quad (4.2.5)$$

## 5 Fourierova transformace

Fourierova transformace je matematický nástroj, který umožňuje vyjádřit obraz pomocí funkcí sinus a kosinus. Přitom frekvence a amplitudy těchto funkcí rozhodují o řadě významných vlastností obrazu. Na Fourierově transformaci je proto založena řada důležitých metod analýzy obrazu. Uvedme nejdříve Fourierovu řadu, která s Fourierovou transformací úzce souvisí.

### 5.1 Fourierova řada funkcí reálné proměnné

Fourierovy řady umožňují rozložit danou reálnou funkci definovanou na jistém intervalu  $\langle a; b \rangle$  na součet (obecně nekonečně mnoha) sinů a kosinů s reálným argumentem. S těmito řadami se podrobněji seznámíte v matematice, zde uvedme na ukázkou Fourierovu řadu na intervalu  $\langle -\pi; \pi \rangle$ . Tuto řadu vypočteme podle následujících vztahů:

$$f(x) = \frac{1}{2} a_0 + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx) \quad (5.1.1)$$

$$a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos nx dx \quad (5.1.2)$$

$$b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin nx dx \quad (5.1.3)$$

**Příklad:** Vypočtěme Fourierovu řadu funkce  $f(x) = \operatorname{sgn} x$  na intervalu  $\langle -\pi; \pi \rangle$ .

**Řešení:** Podle (5.1.2) máme:

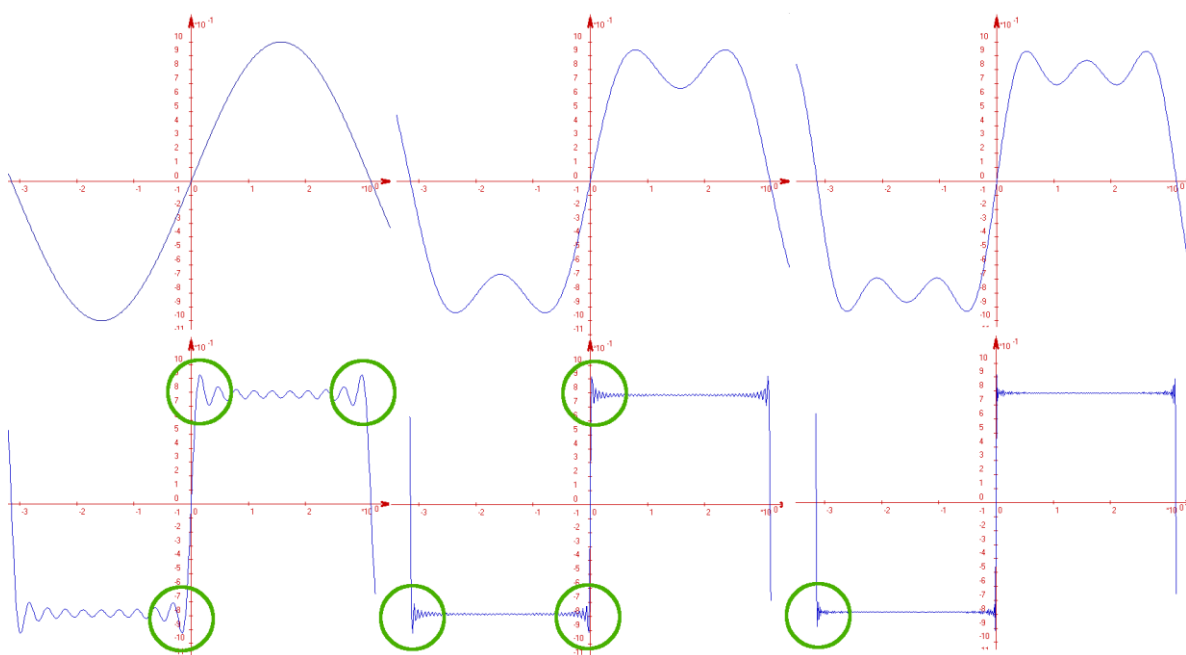
$$a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos nx dx = \frac{1}{\pi} \int_{-\pi}^{\pi} \operatorname{sgn}(x) \cos nx dx = \frac{1}{\pi} \left[ \int_{-\pi}^0 -\cos nx dx + \int_0^{\pi} \cos nx dx \right] = 0$$

$$b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin nx dx = \frac{1}{\pi} \int_{-\pi}^{\pi} \operatorname{sgn}(x) \sin nx dx = \frac{2}{\pi} \int_0^{\pi} \sin nx dx = -\frac{2}{\pi} \left[ \frac{\cos nx}{n} \right]_0^{\pi} = -\frac{2}{\pi n} (\cos n\pi - 1)$$

pro lichá  $-2$ , pro sudá  $0$ .  $\Rightarrow b_{2n+1} = \frac{4}{(2n+1)\pi}$ ;  $n = 0; 1; \dots$ . Tedy

$$\operatorname{sgn}(x) = \frac{4}{\pi} \sum_{n=0}^{\infty} \frac{\sin nx}{2n+1}$$

Několik prvních částečných součtů této řady je ilustrováno na obr. 5.1.1. Všimněte si, že při vyšším počtu sčítanců, kdy jsou v součtu přítomny funkce vyšších frekvencí, dochází k jejich významnému zesílení v místech, kde má prokládaná funkce velký skok ve funkčních hodnotách (zde v okolí nuly), anebo v blízkosti hranic definičního oboru (zde  $\pm\pi$ ). Na obr. 5.1.1. jsou některé situace zakroužkovány. Tento efekt (tzv. Gibbsův jev) je při zpracovávání signálu či obrazu velmi nežádoucí většinou je třeba ho významně potlačit. Setkáme s s ním i v dalším textu.



**Obr. 5.1.1:** Částečné součty Fourierovy řady funkce  $f(x) = \operatorname{sgn} x$  na intervalu  $\langle -\pi; \pi \rangle$

Obraz je dvojrozměrný, uveďme tedy vzorce (5.1.1), (5.1.2) a (5.1.3) pro případ funkce dvou proměnných. Pro funkci definovanou a integrovatelnou na obdélníku  $\langle -p; p \rangle \times \langle -q; q \rangle$  platí

$$f(x) = \sum_{k,n=0}^{\infty} \lambda_{kn} \left[ a_{kn} \cos \frac{k\pi x}{p} \cdot \cos \frac{n\pi y}{q} + b_{kn} \sin \frac{k\pi x}{p} \cdot \cos \frac{n\pi y}{q} + c_{kn} \cos \frac{k\pi x}{p} \cdot \sin \frac{n\pi y}{q} + d_{kn} \sin \frac{k\pi x}{p} \cdot \sin \frac{n\pi y}{q} \right]$$

$$a_{k;n} = \frac{1}{pq} \iint_{\langle -p;p \rangle \times \langle -q;q \rangle} f(x,y) \cdot \cos \frac{k\pi x}{p} \cdot \cos \frac{n\pi y}{q} dx dy$$

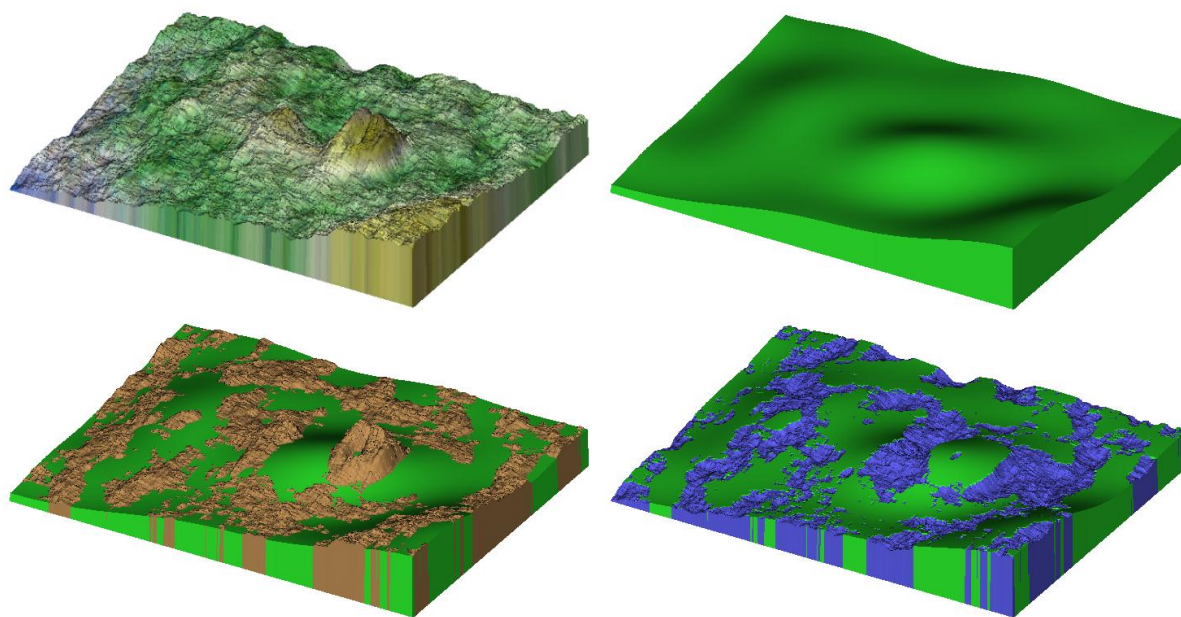
$$b_{k;n} = \frac{1}{pq} \iint_{\langle -p;p \rangle \times \langle -q;q \rangle} f(x,y) \cdot \sin \frac{k\pi x}{p} \cdot \cos \frac{n\pi y}{q} dx dy$$

$$c_{k;n} = \frac{1}{pq} \iint_{\langle -p;p \rangle \times \langle -q;q \rangle} f(x,y) \cdot \cos \frac{k\pi x}{p} \cdot \sin \frac{n\pi y}{q} dx dy$$

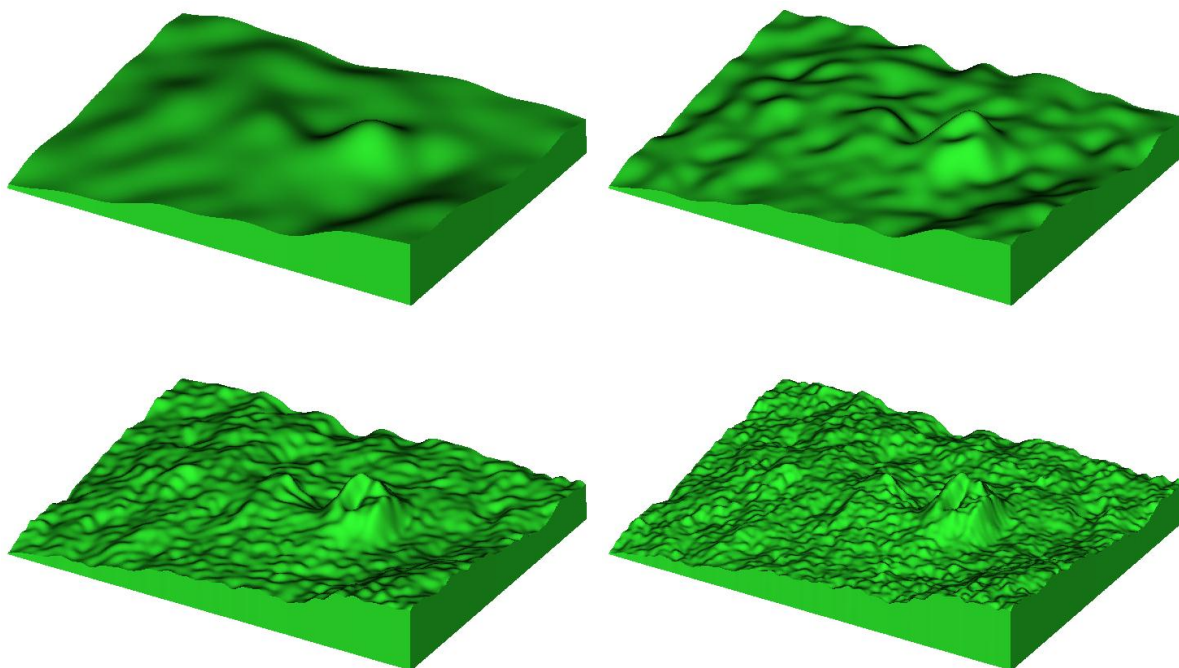
$$d_{k;n} = \frac{1}{pq} \iint_{\langle -p;p \rangle \times \langle -q;q \rangle} f(x,y) \cdot \sin \frac{k\pi x}{p} \cdot \sin \frac{n\pi y}{q} dx dy$$

$$\lambda_{k;n} = \begin{cases} 1 & \Leftrightarrow k, n > 0; k \neq n \\ 1/2 & \Leftrightarrow k + n \neq 0; k \cdot n = 0 \\ 1/4 & \Leftrightarrow k = n = 0 \end{cases} \quad (5.1.4)$$

Tyto Fourierovy řady lze použít např. při aproximaci lomových ploch nejrůznějších materiálů tak, jak je ilustrováno na obr. 5.1.2. Vlevo nahoře lomová plocha obarvená pseudobarvami ilustrujícími výšku profilu, vpravo nahoře částečný součet šesti členů Fourierovy řady (5.1.4). Dole tzv. high-surface – části plochy, které mají větší výšku než plocha referenční (vlevo) a konečně low-surface (vpravo.). Integrály (5.1.4) se řeší numericky. I ve dvojrozměrných případech se setkáváme s nepříjemným Gibbsovým jevem.



**Obr. 5.1.2:** Lomová plocha cementové pasty aproximovaná částečným součtem Fourierovy řady dvou proměnných dle (5.4.1)



**Obr. 5.1.3:** Částečné součty členů Fourierovy řady lomové plochy cementové pasty z obr. 5.1.2. postupně pro  $n = 10; 20; 50; 100$ .

## 5.2 Fourierova transformace

**Fourierova řada funkce jedné proměnné v komplexním oboru:** Uvažujme Eulerovy identity

$$e^{ix} = \cos x + i \sin x$$

$$e^{-ix} = \cos x - i \sin x$$

Jejich sečtením resp. odečtením obdržíme

$$\cos x = \frac{1}{2}(e^{ix} + e^{-ix}) \quad \text{resp.} \quad \sin x = \frac{1}{2i}(e^{ix} - e^{-ix})$$

tedy

$$\cos nx = \frac{1}{2}(e^{inx} + e^{-inx}) \quad \text{resp.} \quad \sin nx = -\frac{i}{2}(e^{inx} - e^{-inx}) \quad (5.2.1)$$

Vztahy (5.2.1) dosadíme do vzorců (5.1.1), (5.1.2), (5.1.3). Obdržíme

$$\begin{aligned} f(x) &= \frac{1}{2}a_0 + \frac{1}{2} \sum_{n=1}^{\infty} [a_n(e^{inx} + e^{-inx}) - i \cdot b_n(e^{inx} - e^{-inx})] \\ f(x) &= \frac{1}{2}a_0 + \frac{1}{2} \sum_{n=1}^{\infty} [e^{inx}(a_n - i \cdot b_n) + e^{-inx}(a_n + i \cdot b_n)] \end{aligned} \quad (5.2.2)$$

Označme

$$c_0 = \frac{1}{2}a_0; \quad c_n = \frac{1}{2}(a_n - i \cdot b_n); \quad c_{-n} = \frac{1}{2}(a_n + i \cdot b_n) \quad (5.2.3)$$

Protože

$$a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos nx dx; \quad b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin nx dx$$

- viz (5.1.2); (5.1.3), můžeme zřejmě psát

$$a_{-n} = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos(-nx) dx = a_n; \quad b_{-n} = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin(-nx) dx = -b_n \quad (5.2.4)$$

Dosadíme-li z (5.2.4) do (5.2.3), máme

$$c_n = \frac{1}{2}(a_n - i \cdot b_n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) \cdot (\cos nx - i \cdot \sin nx) dx = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) \cdot e^{-inx} dx$$

$$c_{-n} = \frac{1}{2}(a_n + i \cdot b_n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) \cdot (\cos nx + i \cdot \sin nx) dx = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) \cdot e^{inx} dx$$

$$c_0 = \frac{1}{2}a_0 = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) \cdot (\cos 0x - i \cdot \sin 0x) dx = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) \cdot e^{-i0x} dx$$

Fourierovu řadu (5.1.1) můžeme tedy psát ve tvaru

$$f(x) = c_0 e^{i0x} + \sum_{n=1}^{\infty} (c_n \cdot e^{inx} + c_{-n} \cdot e^{-inx})$$

tedy

$$f(x) = \sum_{n=-\infty}^{\infty} c_n \cdot e^{inx}; \quad c_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) \cdot e^{-inx} dx \quad (5.2.5)$$

Nekonečnou řadu (5.2.5) nazýváme Fourierovou řadou funkce  $f(x)$  na intervalu  $\langle -\pi; \pi \rangle$ .

Zopakujeme-li celý postup na intervalu  $\langle -p; p \rangle$ , dostaneme koeficienty  $c_n$  ve tvaru

$$c_n = \frac{1}{2p} \int_{-p}^p f(x) \cdot e^{-i \frac{n\pi}{p} x} dx; \quad n \in \mathbb{Z} \quad (5.2.6)$$

**Fourierova transformace funkce jedné reálné proměnné:** Ve vztahu (5.2.6) označme

$$\frac{n\pi}{p} = u_n, \quad (5.2.7)$$

dostaneme

$$c_n = \frac{1}{2p} \int_{-p}^p f(x) \cdot e^{-iu_n x} dx; \quad n \in \mathbb{Z}$$

Rozšíříme-li dále definiční obor funkce  $f(x)$  tak, že  $x \notin \langle -p; p \rangle \Rightarrow f(x) = 0$ , můžeme psát

$$c_n = \frac{1}{2p} \int_{-\infty}^{\infty} f(x) \cdot e^{-iu_n x} dx; \quad n \in \mathbb{Z}$$

Označíme-li



$$\int_{-p}^p f(x) \cdot e^{-iu_n x} dx = F(u_n)$$

dostáváme koeficienty Fourierovy řady (5.2.6) ve tvaru

$$c_n = \frac{1}{2p} F(u_n); \quad n \in \mathbb{Z}$$

Rozšíříme-li definiční obor funkce  $F$  z posloupnosti  $\{u_n\}$  čísel tvaru (5.2.7) na množinu dostaneme funkci  $F$  tvaru

$$F(u) = \int_{-\infty}^{\infty} f(x) \cdot e^{-iux} dx \quad (5.2.8)$$

Lze ukázat, že tato funkce je prostá, existuje k ní tedy funkce inverzní. Ta je tvaru

$$f(x) = \frac{1}{2p} \int_{-\infty}^{\infty} F(u) \cdot e^{iux} du \quad (5.2.9)$$

Funkci  $F(u)$  tvaru (5.2.8) nazýváme **(přímou) Fourierovou transformací** funkce  $f(x)$ . Funkci  $f(x)$  tvaru (5.2.9) nazýváme **zpětnou Fourierovou transformací** funkce  $F(u)$ .

**Diskrétní Fourierova transformace funkce jedné reálné proměnné:** Je-li funkce  $f(x)$  známa jen v konečném počtu hodnot

$$\{x_0; x_1; \dots; x_{N-1}\} = \{x_r\}_{r=0}^{N-1}$$

může být její Fourierova transformace  $F(u)$  známa rovněž jen v konečném počtu hodnot

$$\{X_0; X_1; \dots; X_{N-1}\} = \{X_s\}_{s=0}^{N-1}$$

Integrály v (5.2.8) a (5.2.9) je třeba nahradit příslušnými sumami.

Posloupnost

$$X_s = \sum_{r=0}^{N-1} x_r \cdot e^{-irs} \quad r; s = 0; 1; \dots; N-1$$

nazýváme **(přímou) diskrétní Fourierovou transformací** posloupnosti  $\{x_r\}_{r=0}^{N-1}$ .

Posloupnost

$$x_r = \frac{1}{N} \sum_{s=0}^{N-1} X_s \cdot e^{irs} \quad r; s = 0; 1; \dots; N-1$$

nazýváme **zpětnou diskrétní Fourierovou transformací** posloupnosti  $\{X_s\}_{s=0}^{N-1}$

**Diskrétní Fourierova transformace funkce dvou reálných proměnných:** Obrazy lze chápat jako dvojrozměrné posloupnosti. Proto nyní zobecníme výsledky předchozího odstavce pro funkce dvou proměnných.

Je-li funkce  $f(x; y)$  známa jen v konečném počtu hodnot, tvoří její známé hodnoty matici

$$\begin{pmatrix} x_{00} & x_{01} & \dots & x_{0;N-1} \\ x_{10} & x_{11} & \dots & x_{1;N-1} \\ \dots & \dots & \dots & \dots \\ x_{M-1;0} & x_{M-1;1} & \dots & x_{M-1;N-1} \end{pmatrix} = (x_{m;n}); \quad m = 0; 1; \dots; M-1; \quad n = 0; 1; \dots; N-1;$$

a pro Fourierovu transformaci dostáváme rovněž jen konečný počet hodnot  $F(u)$

$$\begin{pmatrix} X_{00} & X_{01} & \dots & X_{0;N-1} \\ X_{10} & X_{11} & \dots & X_{1;N-1} \\ \dots & \dots & \dots & \dots \\ X_{M-1;0} & X_{M-1;1} & \dots & X_{M-1;N-1} \end{pmatrix} = (X_{r;s}); \quad r = 0; 1; \dots; M-1; \quad s = 0; 1; \dots; N-1;$$

Fourierovy transformace mají pak následující tvar:

**Přímá:**

$$X_{rs} = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} x_{m;n} \cdot e^{-imr} \cdot e^{-ins}; \quad r; m = 0; 1; \dots; M-1; \quad s; n = 0; 1; \dots; N-1; \quad (5.2.10)$$

**Zpětná**

$$x_{m;n} = \sum_{r=0}^{M-1} \sum_{s=0}^{N-1} X_{r;s} \cdot e^{imr} \cdot e^{ins}; \quad r; m = 0; 1; \dots; M-1; \quad s; n = 0; 1; \dots; N-1; \quad (5.2.11)$$

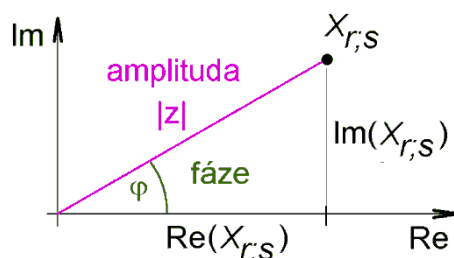
**Amplitudové a fázové spektrum:** Jak plyne z předchozího textu, Fourierovou transformací reálné obrazové matice je matice komplexních čísel. Každé toto komplexní číslo má svoji amplitudu

$$|X_{r;s}| = \sqrt{\text{Re}^2(X_{r;s}) + \text{Im}^2(X_{r;s})}$$

a svoji fázi

$$\text{tg } \varphi_{r;s} = \frac{\text{Im}(X_{r;s})}{\text{Re}(X_{r;s})} \Rightarrow \varphi_{r;s} = \text{arctg} \frac{\text{Im}(X_{r;s})}{\text{Re}(X_{r;s})}$$

(viz obr. 5.2.1).



**Obr. 5.2.1:** Amplituda a fáze komplexního čísla

Matici amplitud nazýváme **amplitudovým spektrem**, matici fází nazýváme **fázovým spektrem**,

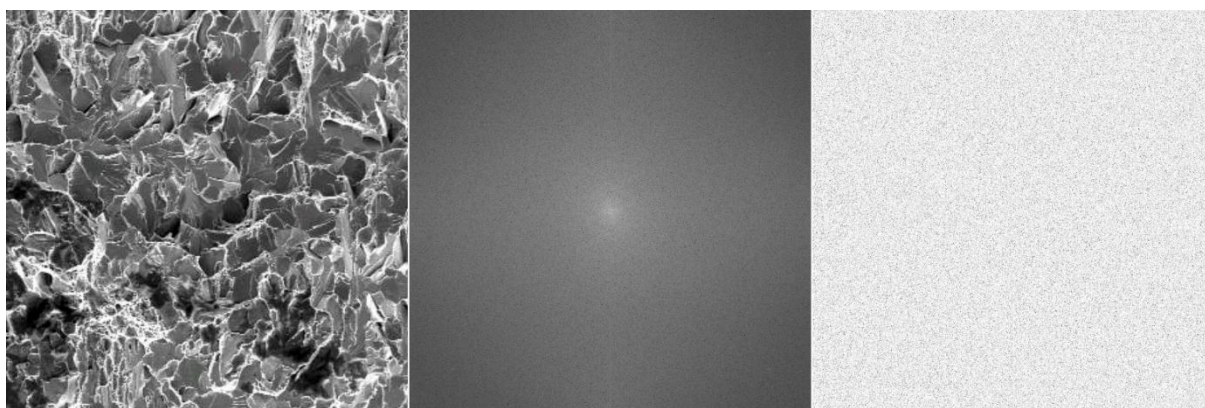
**Rychlá Fourierova transformace (Fast Fourier Transform - FFT):** je efektivní algoritmus výpočtu Fourierovy transformace. Využívá skutečnosti, že je-li matice  $(x_{m;n})$  vstupních hodnot reálná, je Fourierova transformace souměrná podle svého středu. Typ vstupní matice („délky stran“) tak lze rozdělit na poloviny a počítat jen čtvrtinu matice. Hodnota  $X_{r;s}$  navíc vzniká z „předchozích“  $X_{r-1;s}$  ;  $X_{r;s-1}$  , kterých lze při dalším půlení s výhodou použít. Jsou-li „délky stran“ mocniny dvou, počítají se „přímo“ pouze první čtyři prvky „do kříže“ (tzv. motýlek). Amplitudové a fázové spektrum se pak zobrazuje jako středově souměrné.

**Poznámka:** V analýze obrazu značíme obrazy většinou velkými písmeny z počátku abecedy, hodnoty jednotlivých pixelů příslušnými písmeny malými, tedy např.

$$A = (a_{ij}); i = 0; 1; \dots; H - 1; j = 0; 1; \dots; W - 1$$

Komplexní matici, kterou obdržíme z obrazu  $A$  Fourierovou transformací, budeme značit  $F(A)$ ; její jednotlivé prvky pak  $F(a_{ij})$ .

Na obr 5.2.2 si můžeme prohlédnout snímek křehkého lomu oceli a jeho amplitudové resp. fázové spektrum.



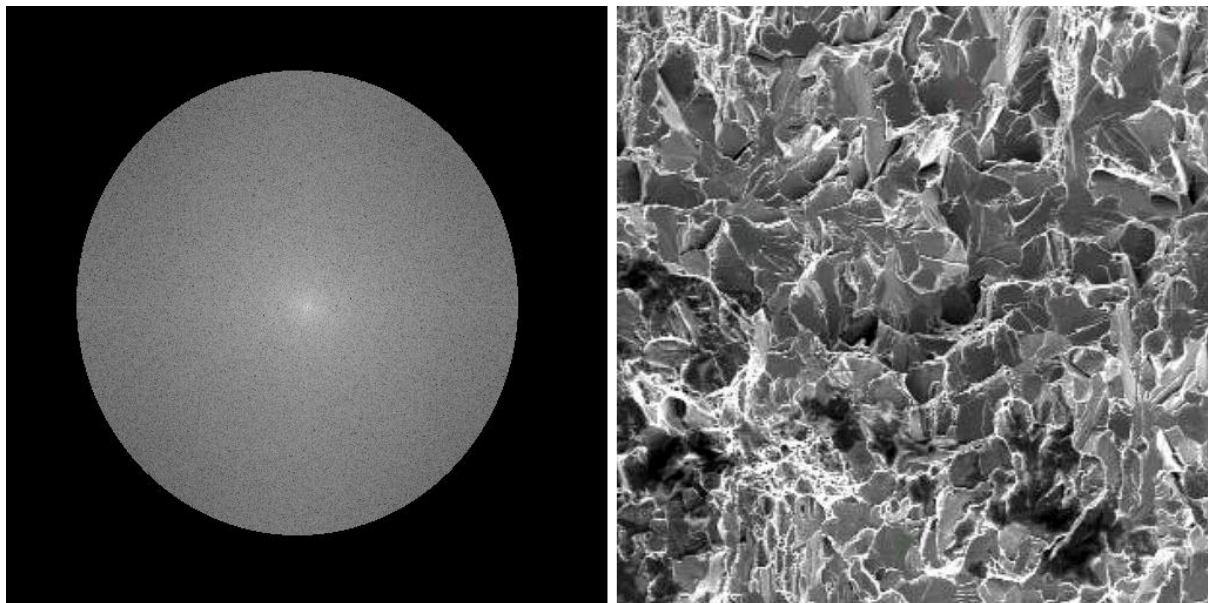
**Obr. 5.2.2:** Snímek křehkého lomu oceli (vlevo), jeho amplitudové spektrum (uprostřed) a fázové spektrum (vpravo)

**Zaostřování a rozostřování obrazu pomocí amplitudového spektra:** Fourierova transformace rozkládá obraz na periodické signály předem daných frekvencí. Hodnoty pixelů amplitudového spektra vypovídají o amplitudách jednotlivých signálů, jejich vzdálenost od středu pak o frekvenci příslušné vlny. Čím je bod dále od středu, tím má příslušná vlna vyšší frekvenci. Vysoké frekvence vnímá oko jako prudké změny jasu. Přítomnost vysokých frekvencí v amplitudovém spektru tedy signalizuje velké zaostření obrazu, může však také prozrazovat přítomnost šumu, který má většinou ještě vyšší frekvence než ostrá užitečná informace.

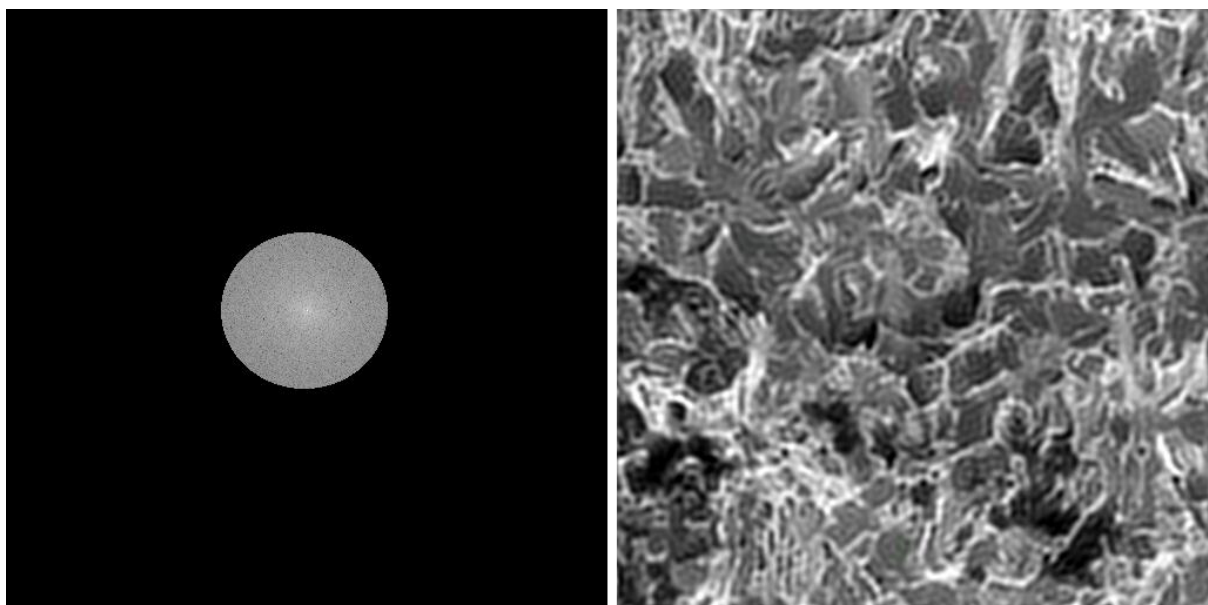
Tyto vlastnosti amplitudového spektra umožňují nejjednodušší úpravy zaostření obrazu a jeho šumu, a to vymaskováním (odstraněním) vyšších či naopak nižších frekvencí a aplikací zpětné Fourierovy transformace na takto upravené spektrum (zpětnou transformací nijak neupraveného spektra dostaneme původní obraz). Maskování amplitudového spektra by teoreticky mělo být zcela symetrické. Pokud není, dostáváme zpětnou transformací obraz, který není reálný, ale komplexní. Nicméně obrazové analyzátory, které umožňují ruční editaci

amplitudového spektra, kde k těmto nesymetrickým zásahům může docházet, po zpětné transformaci imaginární část obrazu ignorují.

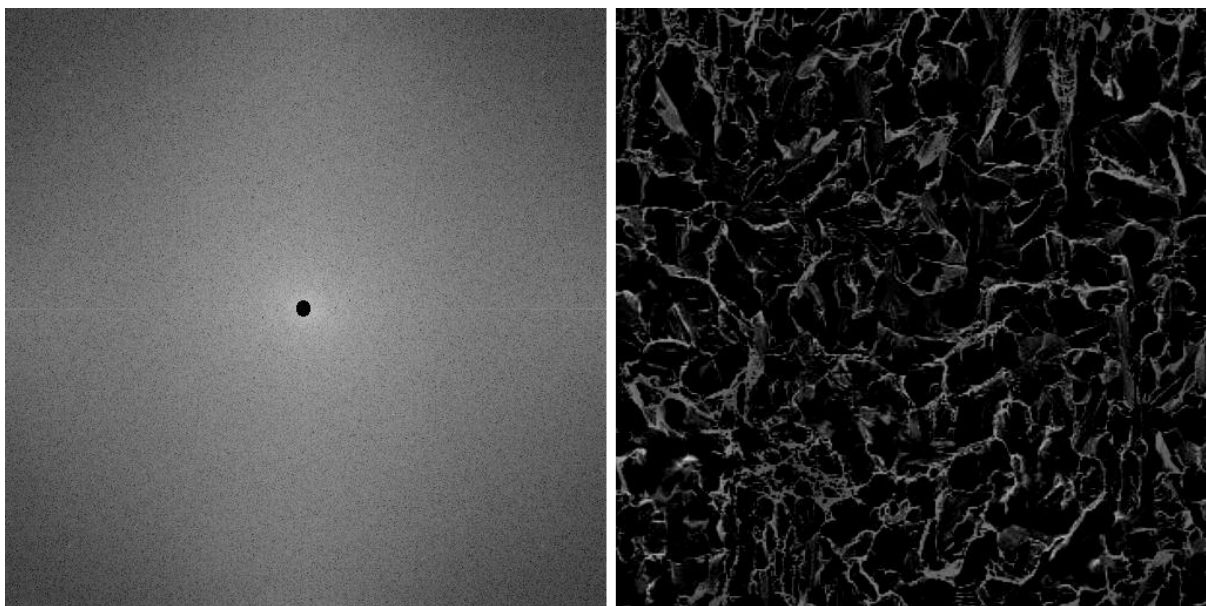
Na několika dalších obrázcích si můžeme prohlédnout účinky některých jednoduchých úprav amplitudového spektra snímku křehkého lomu oceli. Na obr 5.2.3 je ilustrována skutečnost, že ani odstranění poměrně velké vysokofrekvenční části amplitudového spektra nemusí znamenat rozostření obrazu. Odstraněním vysokých frekvencí jsme v tomto případě odstranili jen šum.



**Obr. 5.2.3:** Vymaskování vysokofrekvenční části amplitudového spektra a jeho zpětná transformace. Snímek není rozostřen, vymaskováním byl odstraněn pouze šum.



**Obr. 5.2.4:** Vymaskování vysokofrekvenční části amplitudového spektra a jeho zpětná transformace. Snímek už je silně rozostřen, vymaskování zasáhlo již i vysoké frekvence obrazu.



**Obr. 5.2.5:** Vymaskování nízkofrekvenční části amplitudového spektra a jeho zpětná transformace. Na snímku zůstávají jen ostré kontrastní přechody.

### 5.3 Konvoluce

**Konvoluce** je matematická operace „zpracovávající“ dvě funkce, výsledkem je funkce třetí. Konvoluce může být aplikována jak na funkce jedné, tak na funkce více proměnných, a to jak na spojité, tak diskrétní. Definujme jednotlivé případy.

**Konvoluce funkcí jedné reálné proměnné:** Konvolucí dvou funkcí  $f(x)$ ;  $g(x)$  rozumíme integrál (pokud existuje).

$$(f \otimes g)(x) = \int_{-\infty}^{\infty} f(t) g(x-t) dt$$

Funkci  $g(x)$  nazýváme konvolučním jádrem

**Konvoluce funkcí dvou reálných proměnných:** Konvolucí dvou funkcí  $f(x; y)$ ;  $g(x; y)$  rozumíme integrál (pokud existuje).

$$(f \otimes g)(x; y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(u; v) g(x-u; y-v) du dv$$

Funkci  $g(x; y)$  nazýváme konvolučním jádrem

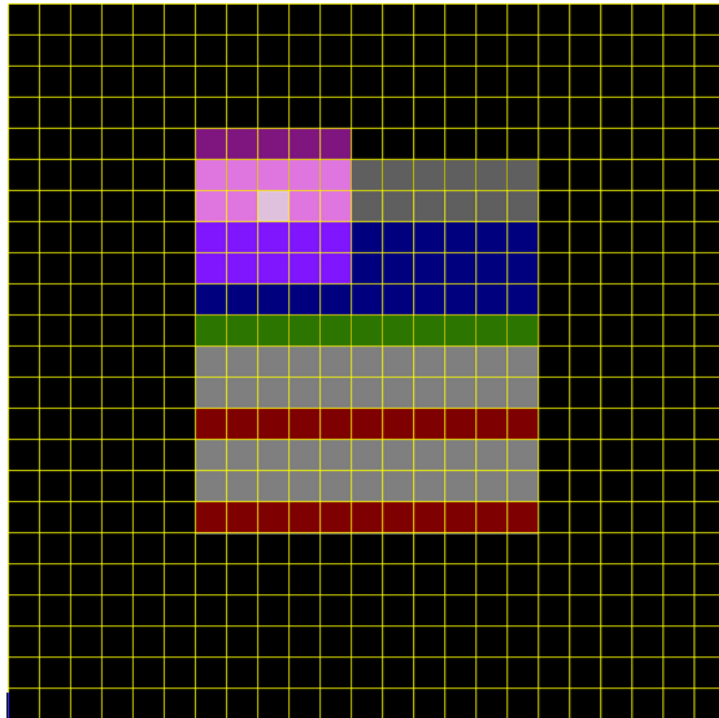
**Diskrétní konvoluce funkcí dvou reálných proměnných:** Konvolucí dvou funkcí  $f(m; n)$ ;  $g(m; n)$ ;  $m; n \in \mathbb{Z}$  rozumíme dvojrozměrnou nekonečnou posloupnost součtů

$$(f \otimes g)(i; j) = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} f(m; n) g(i-m; j-n)$$

(pokud každý součet existuje).

Použití diskrétní konvoluce v analýze obrazu ilustruje obr. 5.3.1. Obraz  $A$  s rozlišením  $W \times H$  doplníme na nekonečnou dvojrozměrnou posloupnost tak, že pro  $m \notin \{0; 1; \dots; H-1\}$  nebo  $n \notin \{0; 1; \dots; W-1\}$  je  $A(m; n) = 0$ . Konvoluční matici, která je v praxi samozřejmě rovněž konečného typu  $I \times J$  (a kterou můžeme chápat rovněž jako obraz), doplníme stejným způsobem. Rozlišení původního obrazu na obr. 5.3.1 je  $W \times H = 11 \times 12$ , původní typ konvoluční matice je  $I \times J = 5 \times 5$ . Pixel  $[1; 2]$  na zpracovávaném obraze  $A$  je nahrazován lineární kombinací pixelů v růžovém okolí. Koefficienty lineární kombinace jsou příslušné prvky konvoluční matice. Hodnotu pixelu v novém obraze tak obdržíme takto:

$$A'(i, j) = (A \otimes C)(i, j) = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} A(m; n) C(i-m; j-n) \quad (5.3.1)$$



**Obr. 5.3.1.** Schématické znázornění konvoluce

**Konvoluční teorém:** Necht'  $A$ ;  $C$  jsou dva obrazy. Pro jejich konvoluci platí

$$F(A \otimes C) = F(A) \cdot F(C) \quad (5.3.2)$$

kde **součin** na pravé straně se provádí **prvek po prvku**. **Pozor! Není to tedy součin matic.**

Jestliže na levou i pravou stranu vztahu (5.3.1) aplikujeme zpětnou Fourierovu transformaci, obdržíme

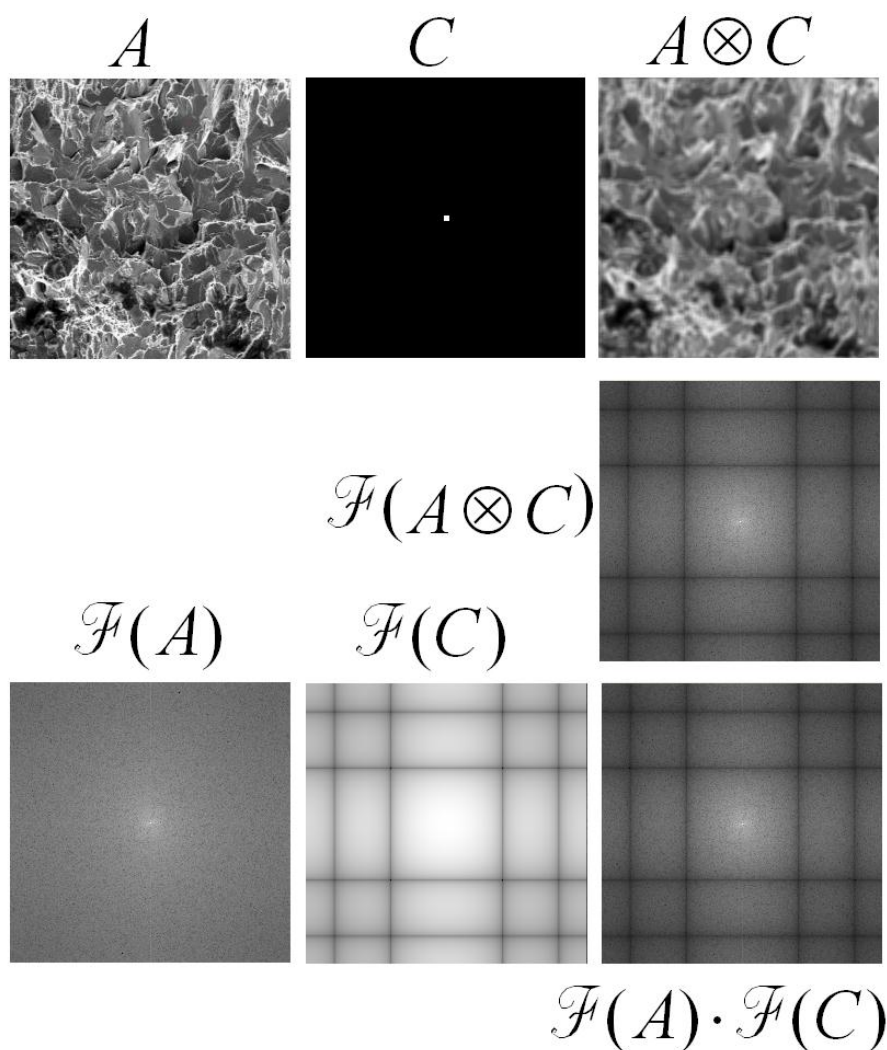
$$F^{-1}[F(A \otimes C)] = F^{-1}[F(A) \cdot F(C)]$$

Levá strana tohoto vztahu je však zřejmě rovna konvoluci  $A \otimes C$ . Z konvolučního teorému tak vyplývá další tvrzení

$$A \otimes C = F^{-1}[F(A) \cdot F(C)] \quad (5.3.3)$$



Konvoluční teorém je ilustrován na obr. (5.3.2).

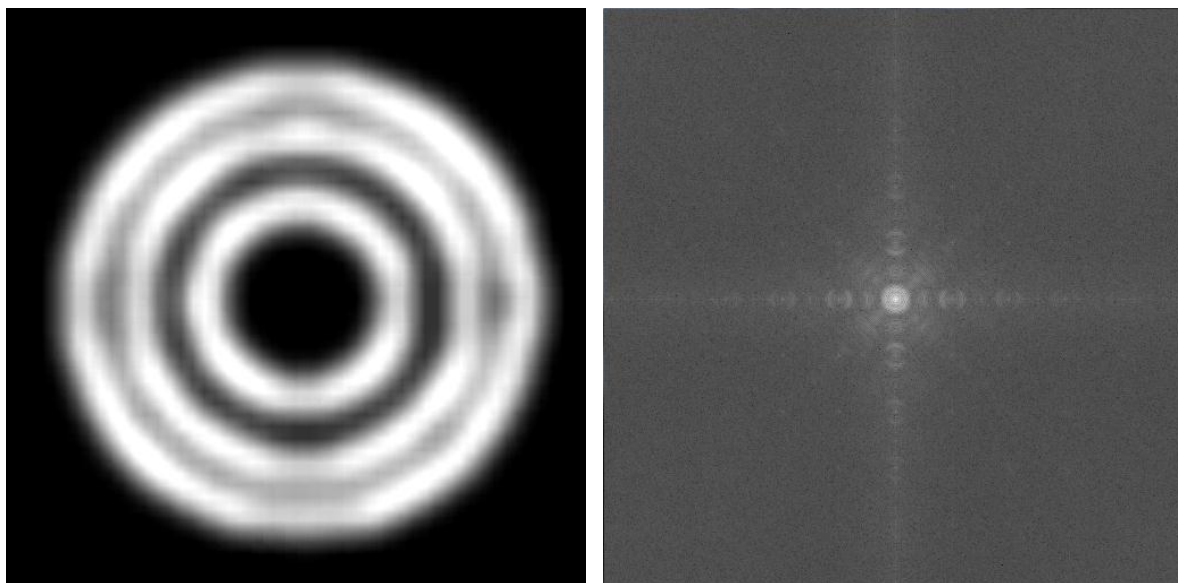


Obr. 5.3.2. Konvoluční teorém

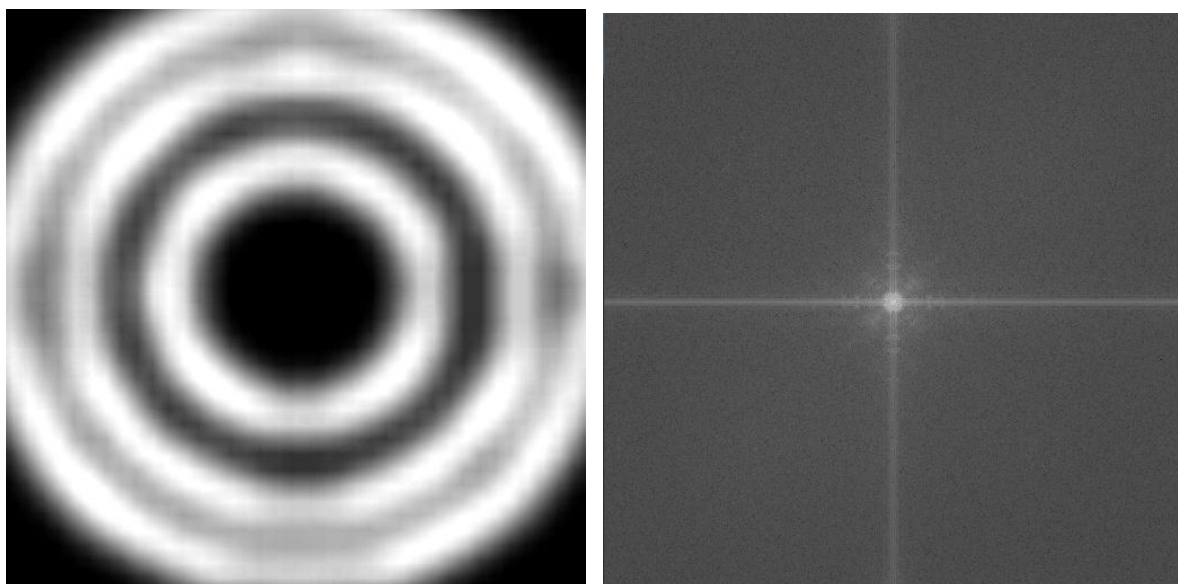
## 5. 4 Okrajový efekt

V kpt. 5.1. jsme si mohli všimnout, že v bodech nespojitosti funkce trpí částečné součty její Fourierovy řady tzv. Gibbsovým jevem – zesílením vysokých frekvencí. Tentýž jev je patrný i v blízkém okolí hraničních bodů definičního oboru. Podobná situace nastává i v případě Fourierovy transformace obrazu. Není-li na okrajích obrazu jeho hodnota konstantní, pak okraj obrazu představuje pro Fourierovu transformaci vysokofrekvenční informaci, která se projeví v jeho spektru. Situace je ilustrována na obr. 5.4.1 a 5.4.2.

Tento jev je velmi nepříjemný a v řadě situací může působit značné potíže. Proto se před nasazením Fourierovy transformace obraz často násobí vhodnou funkcí tak, aby jeho hodnoty směrem k okrajům „plynule“ klesaly k nule. Funkce, které se k tomuto účelu používají, se nazývají **window funkce** a nejpoužívanější jsou dvě z nich.



**Obr. 5.4.1:** Fourierovo spektrum obrazu s konstantním (zde nulovým) okrajem



**Obr. 5.4.2:** Fourierovo spektrum obrazu s proměnným okrajem. Ve srovnání s obr. 5.4.1. vidíme zřetelný nárůst amplitud ve vodorovném a svislém směru způsobený vysokofrekvenční okrajovou informací.

**Gaussovské okno:** Gaussova funkce určuje hustotu pravděpodobnosti normálního rozdělení a již jsme se s ní setkali v kpt. 2.1. – viz (2.1.2). Tato funkce je rovněž vhodná pro potlačení okrajového efektu při Fourierově transformaci (dvojměrného) obrazu. Zde ovšem potřebujeme Gaussovu funkci dvou proměnných, která má tvar

$$g(x; y) = \frac{1}{\sigma_1 \sigma_2 \cdot 2\pi} \cdot e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2} - \frac{(y-\mu_2)^2}{2\sigma_2^2}}$$

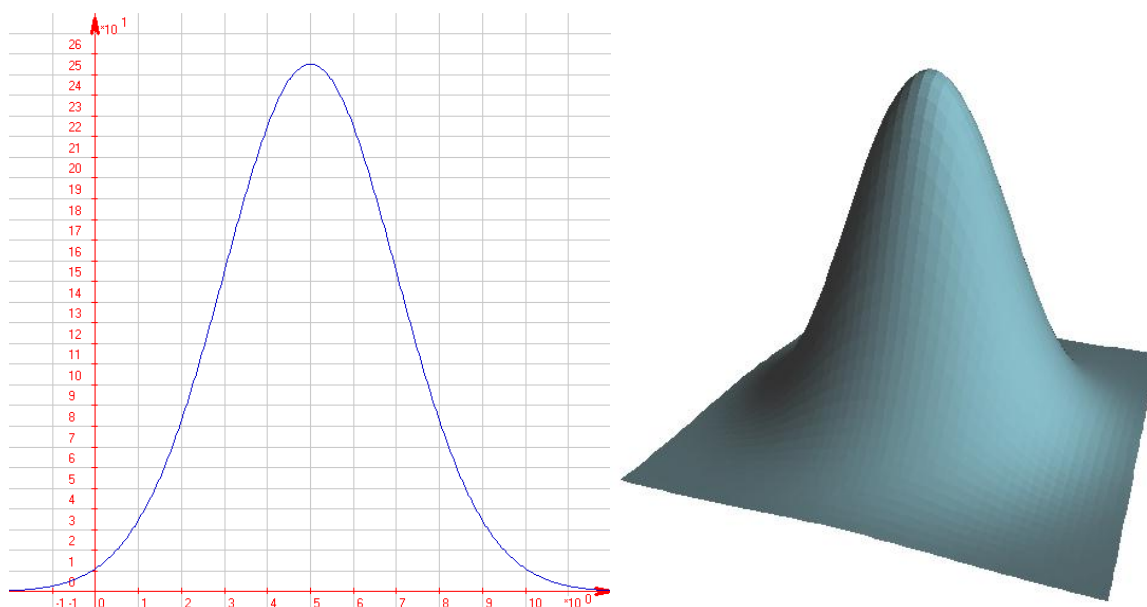
Multiplikativní konstanta je volena tak, aby platilo

$$\iint_{\mathbb{R}^2} g(x; y) \, dx dy = 1$$

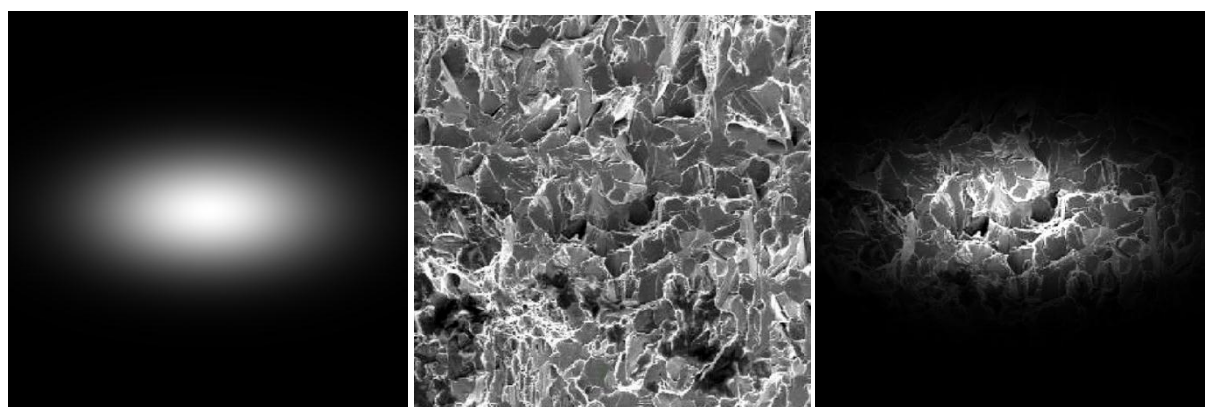
což je důležité jednak v teorii pravděpodobnosti a jednak i v analýze obrazu tak, jak se o tom zmíníme v kapitole o lineárních filtrech. Na tomto místě však potřebujeme, aby maximum funkce (střední hodnota)  $[\mu_1; \mu_2]$  byla ve středu obrazu a aby funkční hodnota zde byla rovna dynamickému rozsahu obrazu, tj.  $g(\mu_1; \mu_2) = D$ . Proto pokládáme

$$g(x; y) = D \cdot e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2} - \frac{(y-\mu_2)^2}{2\sigma_2^2}}; \quad \mu_1 = \frac{W}{2}; \quad \mu_2 = \frac{H}{2} \quad (5.4.1)$$

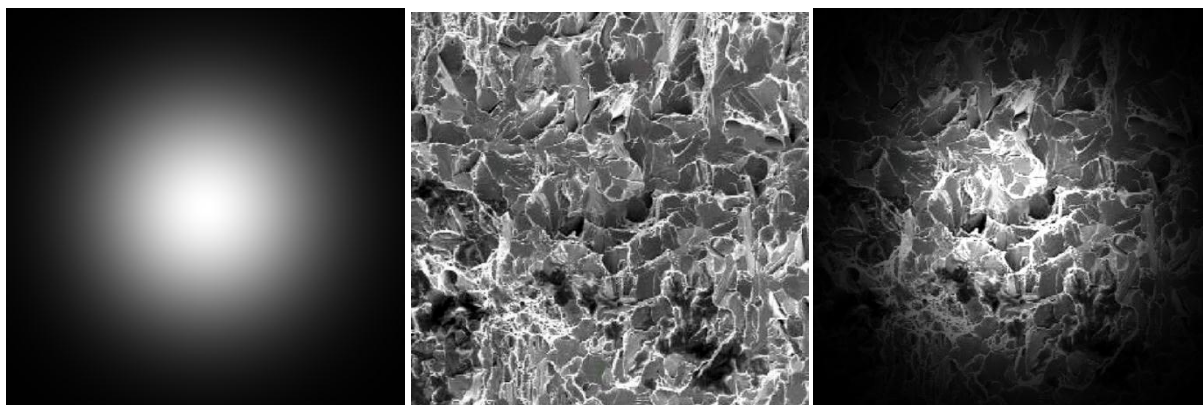
Gaussova funkce jedné a dvou proměnných je ilustrována na obr. 5.4.3, jako obraz pro různé rozptyly na obr. 5.4.4. a 5.4.5.



**Obr. 5.4.3:** Gaussova funkce jedné (vlevo) a dvou proměnných (vpravo).



**Obr. 5.4.4:** Gaussova funkce s rozptyly  $\sigma_1^2 = 0.3D$ ;  $\sigma_2^2 = 0.15D$  jako obraz (vlevo) použita jako window funkce na obraz křehkého lomu oceli (uprostřed). Vpravo kompresní součin.



**Obr. 5.4.5:** Gaussova funkce s rozptyly  $\sigma_1^2 = \sigma_2^2 = 0.3D$  jako obraz (vlevo) použita jako window funkce na obraz křehkého lomu oceli (uprostřed). Vpravo kompresní součin.

**Hanningovo okno:** je pro jednu proměnnou tvaru

$$h(x) = \frac{1}{2} \cos \frac{2\pi(x - \frac{P}{2})}{P} - \frac{1}{2}$$

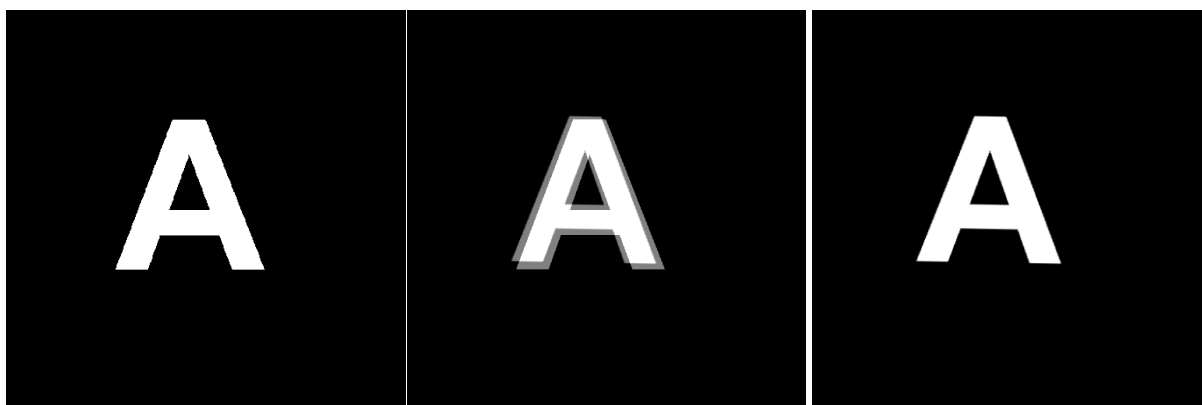
ve dvou proměnných má tvar

$$h(x; y) = \frac{1}{2} \cos \frac{2\pi(x - \frac{W}{2})}{W} \cos \frac{2\pi(y - \frac{H}{2})}{H} - \frac{1}{2}$$

K odstranění okrajového efektu při Fourierově transformaci je rovněž vhodné a hojně se používá.

## 5. 5 Registrace obrazů

**Registrací obrazů** rozumíme proces nalezení vhodné transformace jednoho obrazu, tak aby se v předem definovaném smyslu podobal jinému obrazu. Tyto obrazy většinou zobrazují stejnou scénu, ale jsou pořízeny z různých míst, různými přístroji, v různém čase atd. Na obr. 5.5.1 (vlevo a vpravo) vidíme dva „různé“ obrazy. Není těžké odhadnout, že v tomto případě se liší jen posunutím. V jednoduchých případech, jako je tento, můžeme odlišnosti obou obrazů odhalit jejich kompresním součtem (uprostřed).

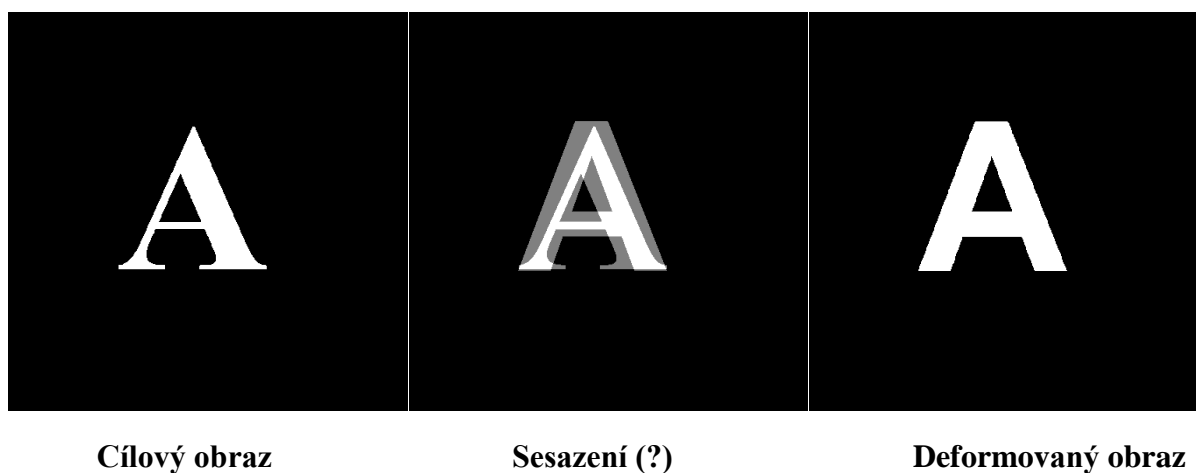


**Obr. 5.5.1:** Dva obrazy, které se liší jen vzájemným popsunutím (vlevo cílový obraz, vpravo transformovaný obraz). V jednoduchých případech můžeme odlišnosti posoudit vizuálně pomocí kompresního součtu (uprostřed).

Obraz, pro který hledáme transformaci, nazýváme transformovaný, druhý z nich, na který pomocí transformace oba obrazy „sesazujeme“, nazýváme cílový.

Celý problém je obecně značně složitý. Prvním zdrojem komplikací mohou být transformace, které je třeba použít. Můžeme uvažovat jednak transformace projektivní (je-li obrazem přímky vždycky přímka, anebo bod), tj. posunutí, otáčení, změnu měřítka, osovou afinitu a středovou kolineaci, jednak to mohou být transformace obecnější (deformace). V tomto textu se omezíme na posunutí, otočení a změnu měřítka.

Problém může dále komplikovat skutečnost, že objekty na snímku nejsou zcela identické a mohou se poněkud lišit. Situace je ilustrována na obr. 5.5.2, kde je písmeno A vysázeno jiným fontem. V tomto případě je nemožné sesadit obrazy na sebe tak, aby každý pixel obrazu  $A$  měl stejnou hodnotu jako odpovídající pixel na obrazu  $B$ . Zde vyvstává otázka, jak posoudit „nejlepší možné“ sesazení takových obrazů.



**Obr. 5.5.1:** Dva obrazy se liší objektem, který je na nich zobrazen. Otázkou je, jak posoudit „nejlepší možné“ sesazení takových obrazů.

K řešení tohoto problému je třeba se poněkud hlouběji zamyslet nad pojmem vzdálenost a zavést vzdálenost obrazů. Měříme-li vzdálenost dvou bodů (ať již metrem, anebo ji počítáme pomocí známého středoškolského vzorečku), přiřazujeme vždy dvěma bodům reálné číslo. Vzdálenost je tedy z matematického hlediska funkce (označme ji v tuto chvíli  $d$ ), která má následující vlastnosti:

Pro každé tři body  $A; B; C$  platí:

$$d(A; B) \geq 0, \text{ přičemž } d(A; B) = 0 \Leftrightarrow A = B \quad (5.5.1)$$

$$d(A; B) = d(B; A) \quad (5.5.2)$$

$$d(A; B) + d(B; C) \geq d(A; C) \quad (5.5.3)$$

Tyto tři základní vlastnosti, které běžně předpokládáme u vzdálenosti, může mít celá řada dalších funkcí, které v tomto smyslu můžeme tedy rovněž považovat za „vzdálenost“. Jsou-li např.  $A; B$  dva obrazy s pixely o hodnotách  $a_{ij}; b_{ij}$ , lze za jejich „vzdálenost“ považovat například funkce

$$f(A; B) = \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} |a_{ij} - b_{ij}|$$

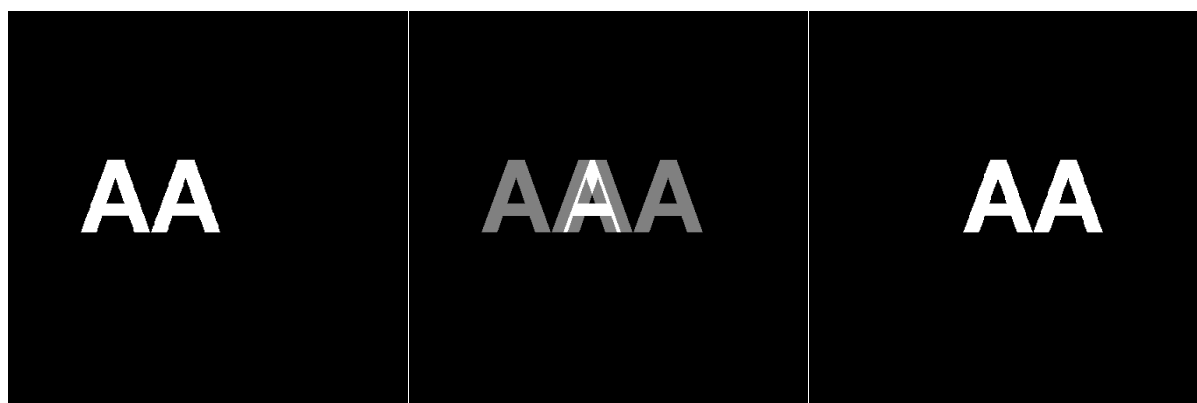
$$g(A; B) = \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} [a_{ij} - b_{ij}]^2 \quad (5.5.4)$$

pokud pixelům, které jsou mimo jednotlivé obrazy, definitoricky přiřadíme hodnotu nula. Naopak např. funkci

$$h(A; B) = \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} (a_{ij} - b_{ij})$$

za „vzdálenost“ prohlásit nemůžeme (proč?). Funkce, které mají vlastnosti (5.5.1); (5.5.2); (5.5.3), označujeme obecným názvem **metriky** a množiny, na kterých jsou definovány, pak **metrickými prostory**.

Hledání „nejlepšího sesazení“ obrazů pak můžeme chápat jako hledání minima nějaké vhodné metriky. Ovšem metody pracující na tomto principu mají některá svá úskalí. Na hledání extrémů v tomto případě není možné nasadit klasické metody matematické analýzy, které známe z kurzu matematiky. Minimum se hledá numericky metodami tzv. optimalizace, které ovšem většinou předpokládají, že na dané množině je jen jeden extrém. To však v případě registrace obrazů nemusí být pravda. Zvolená metrika může mít na množině všech možných transformací více minim tak, jak ilustruje obr. 5.5.2.



**Obr. 5.5.2:** Problém více lokálních minim. Sesazení „na nesprávné A“ vykazuje lokální minimum přesto, že obrazy nejsou správně sesazeneny.

Dále tedy popíšeme metodu registrace obrazů využívající Fourierovu transformaci. Ještě předtím však provedeme jednu přípravnou úvahu.

**Diracova distribuce:** uvažujme funkce tvaru (viz obr 5.5.3.)

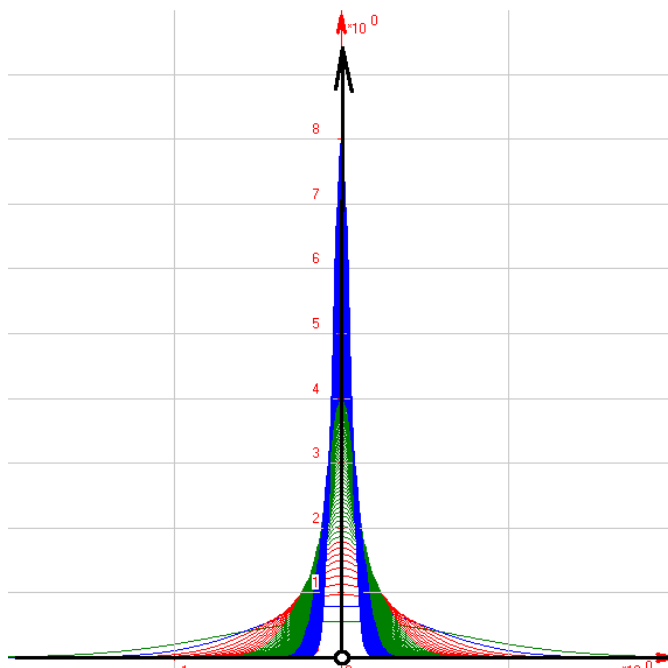
$$\delta_p(x) = \sqrt{\frac{p}{\pi}} \cdot e^{-px^2} \quad (5.5.5)$$

Grafem každé takovéto funkce je Gaussova křivka s maximem v nule. Dá se ukázat, že obsah plochy vymezené touto křivkou a osou x je stále stejný a je roven jedné, tj.

$$\int_{-\infty}^{\infty} \delta_p(x) dx = 1$$



Je zřejmé, že se vzrůstajícím  $p$  hodnota tohoto maxima roste a rozptyl klesá. Jestliže parametr  $p$  poroste nade všechny meze, maximum bude divergovat a rozptyl konvergovat k nule. Tuto limitu budeme označovat



**Obr. 5.5.3:** K Diracově distribuci

$$\delta(x) = \lim_{p \rightarrow \infty} \delta_p(x) = \lim_{p \rightarrow \infty} \sqrt{\frac{p}{\pi}} \cdot e^{-px^2} \quad (5.5.6)$$

a nazývat **Diracovou (jednorozměrnou) distribucí v bodě nula**. Pojem distribuce je tedy zobecněním pojmu funkce, kdy jako funkční hodnoty připouštíme hodnoty  $\pm\infty$ . Podobně jako v případě funkcí můžeme i v případě distribucí pracovat s různými argumenty, tj. např. distribuce

$$\delta(x-2) = \lim_{p \rightarrow \infty} \delta_p(x-2) = \lim_{p \rightarrow \infty} \sqrt{\frac{p}{\pi}} \cdot e^{-p(x-2)^2}$$

má své „nekonečné“ maximum posunuto z nuly do bodu  $x=2$ .

Celou úvahu lze zopakovat pro dvě proměnné. Funkce  $\delta_p(x; y)$  jsou tvaru

$$\delta_p(x; y) = \sqrt{\frac{p}{\pi}} \cdot e^{-px^2} \cdot \sqrt{\frac{p}{\pi}} \cdot e^{-py^2} = \frac{p}{\pi} e^{-p(x^2+y^2)} \quad (5.5.7).$$

Diracova distribuce dvou proměnných má tvar

$$\delta(x; y) = \lim_{p \rightarrow \infty} \delta_p(x; y) = \lim_{p \rightarrow \infty} \frac{p}{\pi} e^{-p(x^2+y^2)}$$

a v „posunuté“ podobě

$$\delta(x-m; y-n) = \lim_{p \rightarrow \infty} \frac{p}{\pi} e^{-p[(x-m)^2+(y-n)^2]} \quad (5.5.8)$$

**Registrace založená na Fourierově transformaci:** Vzájemné posunutí, otočení a změnu měřítka dvou obrazů můžeme identifikovat pomocí jejich Fourierových spekter resp. pomocí tzv. křížové, fázové resp. smíšené korelace. Mějme dva obrazy  $(A, B)$  stejného rozlišení, tj. matice reálných čísel  $A = (a_{ij})$ ;  $B = (b_{ij})$ ;  $i = 0; 1; \dots; W-1$ ;  $j = 0; 1; \dots; H-1$  a jejich Fourierovy transformace, tj. matice komplexních čísel  $F(A) = (F(a_{ij}))$ ;  $F(B) = (F(b_{ij}))$ ;  $i = 0; 1; \dots; W-1$ ;  $j = 0; 1; \dots; H-1$ . Označme  $F^*(b_{ij})$  číslo komplexně sdružené s  $F(b_{ij})$  a  $|F(a_{ij})|$  resp.  $|F^*(b_{ij})|$  absolutní hodnotu čísla  $F(a_{ij})$  resp.  $F^*(b_{ij})$ . Pak matici

$$A \times B = (F(a_{ij}) \cdot F^*(b_{ij})) \quad (5.5.9)$$

nazýváme **křížovou korelací obrazů**  $A; B$ . Matici

$$A \otimes B = \left( \frac{F(a_{ij}) \cdot F^*(b_{ij})}{|F(a_{ij})| \cdot |F^*(b_{ij})|} \right) \quad (5.5.10)$$

nazýváme **fázovou korelací obrazů**  $A; B$ . Matici

$$A \oplus B = \left( \frac{F(a_{ij}) \cdot F^*(b_{ij})}{|F(a_{ij}) + c_1| \cdot |F^*(b_{ij}) + c_2|} \right); \quad c_1, c_2 \in \mathbb{R}^+ \quad (5.5.11)$$

nazýváme **smíšenou korelací obrazů**  $A; B$ .

**Poznámka:** V literatuře se často setkáváme s „jednoduššími“ tvary výrazů (5.5.8); (5.5.9) a (5.5.10), a sice

$$\begin{aligned} A \times B &= F(A) \cdot F^*(B) && \text{(křížová korelace)} \\ A \otimes B &= \frac{F(A) \cdot F^*(B)}{|F(A)| \cdot |F^*(B)|} && \text{(fázová korelace)} \\ A \oplus B &= \frac{F(A) \cdot F^*(B)}{|F(A) + c_1| \cdot |F^*(B) + c_2|} && \text{(smíšená korelace)} \end{aligned}$$

Tyto zápisy jsou však značně zavádějící, neboť svádějí k domněnce, že  $F(A) \cdot F^*(B)$  je součin matic,  $|F(A)|$  je některá z (nekonečně mnoha) norem matice  $F(A)$  apod. Proto na tomto místě zdůrazňujeme, že všechny operace ve výrazech (5.5.9); (5.5.10) a (5.5.11), je třeba provádět prvek po prvku.

Pro detekci **posunutí dvou obrazů** má stěžejní význam následující věta:

**Fázová korelace dvou vzájemně posunutých funkcí (spojitý případ):** Necht'  $a(x; y)$ ;  $b(x; y)$  dvě funkce dvou proměnných takové, že  $b(x; y) = a(x - x_0; y - y_0)$ . Pak pro jejich fázovou korelaci platí:

$$a \otimes b = \delta(x - x_0; y - y_0) \quad (5.5.12)$$

Fázovou korelací je tedy Diracova distribuce posunutá do bodu  $[x_0; y_0]$ . Důsledkem pro diskrétní obrazy je skutečnost, že fázovou korelací vzájemně posunutých a jinak zcela identických obrazů dostaneme obraz, který má ve všech bodech hodnotu nula, kromě bodu  $[x_0; y_0]$ , kde sice není „nekonečný“ Diracův impuls, ale maximální možná hodnota pixelu při daném dynamickém rozsahu, tj.  $D$ . Nejsou-li obrazy zcela identické (což je v praxi nejčastější případ), dostaneme fázovou korelací obraz, který má v naprosté většině pixelů hodnotu nula, až na více či méně rozsáhlý shluk nenulových pixelů s jedním maximálním pixelem. Posunutí takových obrazů optimálně detekují souřadnice pixelu s maximální hodnotou.

Místo fázové korelace se v praxi používá korelace smíšená, kde reálná čísla  $c_1; c_2 \in \mathbb{R}^+$  brání možnému dělení nulou. Velikost a tvar výše popsaného shluku pixelů pak závisí jednak na míře odlišnosti obrazů a jednak na konkrétní volbě čísel  $c_1; c_2 \in \mathbb{R}^+$ .

Pro detekci **vzájemného otočení dvou obrazů** má stěžejní význam následující věta:

**Věta: Amplitudové spektrum dvou funkcí vzájemně otočených kolem obecného středu (spojitý případ):** Necht'  $a(x; y); b(x; y)$  dvě funkce dvou proměnných takové, že  $b(x'; y') = a(x; y)$  právě tehdy, když

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$$

Pak pro jejich amplitudová spektra  $r(\kappa; \lambda) = F[a(x; y)]; s(\kappa; \lambda) = F[b(x; y)]$  platí:  $s(\kappa'; \lambda') = r(\kappa; \lambda)$  právě tehdy, když

$$\begin{pmatrix} \kappa' \\ \lambda' \end{pmatrix} = \begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix} \begin{pmatrix} \kappa \\ \lambda \end{pmatrix} \quad (5.5.13)$$

Tato věta poskytuje návod, jak určit otočení dvou obrazů:

- 1) Sestrojíme amplitudová spektra obou obrazů – dle předchozí věty jsou vzájemně otočena o úhel  $\varphi$  kolem počátku
- 2) Převedeme spektra do polárních souřadnic, tím převedeme otočení na posunutí
- 3) Detekci posunutí provedeme podle (5.5.12)

Detekci změny měřítka umožňuje následující věta:

**Věta: Amplitudové spektrum při změně měřítka (spojitý případ):** Necht'  $a(x; y); b(x; y)$  dvě funkce dvou proměnných takové, že  $a(x; y) = b(cx; cy)$ ,  $c > 0$ . Pak pro jejich amplitudová spektra  $r(\kappa; \lambda) = F[a(x; y)]; s(\kappa; \lambda) = F[b(x; y)]$  platí:

$$r(\kappa; \lambda) = \frac{1}{c^2} \cdot s\left(\frac{\kappa}{c}; \frac{\lambda}{c}\right) \quad (5.5.14)$$

Tato věta modifikuje předchozí návod a umožňuje určit otočení dvou obrazů spolu se změnou měřítka:

4) Sestrojíme amplitudová spektra obou obrazů – dle předchozí věty jsou vzájemně otočena o úhel  $\varphi$  kolem počátku

5) Převedeme amplitudová spektra do log-polárních souřadnic (polární soustava se zlogaritmovanými osami)  $\Rightarrow$  spektra jsou posunutá o vektor  $(-\log c; \log \varphi_0)$ . Transformace složená z otočení a změny měřítka je opět převedena na posunutí.

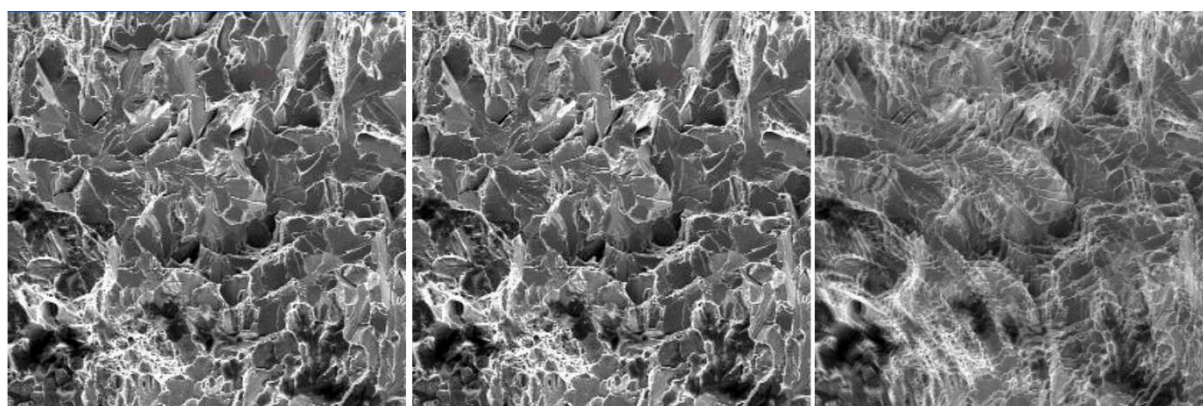
Chceme-li tedy na cílový obraz  $A$  sesadit transformovaný obraz  $B$ , který se od  $A$  liší posunutím, otočením a změnou měřítka, postupujeme následovně:

6) Nejdříve detekujeme otočení a změnu měřítka dle bodů 4) a 5)

7) Podle výsledků v bodě 6) transformovaný obraz  $B$  otočíme a změníme mu měřítko. Tím získáme obraz  $B'$

8) Obrazy  $A$ ;  $B'$  jsou vzájemně posunuté, posun zjistíme dle (5.5.12).

9) Dle výsledků v bodě 8) posuneme obraz  $B'$ .

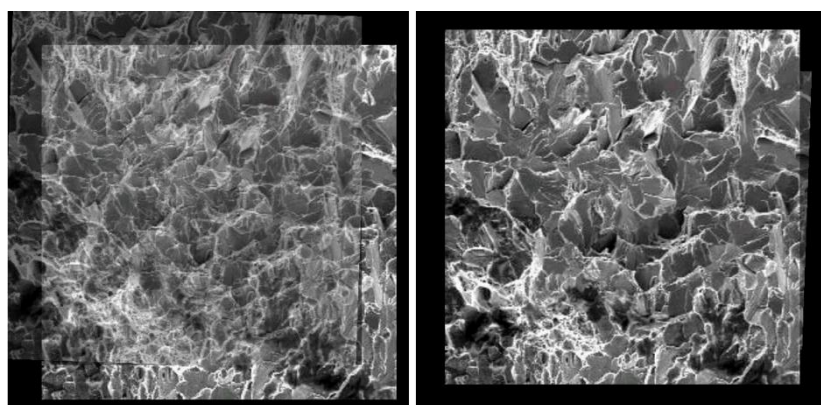


Cílový obraz

Transformovaný obraz

Kompresní součet

**Obr. 5.5.4:** Dva snímky křehkého lomu oceli a jejich kompresní součet. Je zřejmé, že obrazy jsou vzájemně otočené.



**Obr. 5.5.5:** Registrace snímků z předchozího obrázku. Detekce a eliminace rotace – obrazy jsou již jen posunuté (vlevo), detekce a eliminace posunutí (vpravo).

## 6 Obrazové filtry

Obrazové filtry jsou matematické nástroje k úpravě obrazu určené ke zvýraznění či potlačení určitých informací. Lze je dělit podle použitých matematických operací na **filtry lineární a nelineární**, nebo podle schopnosti přizpůsobit svoji činnost lokálními vlastnostem obrazu na **filtry neadaptivní a adaptivní**. Mnohé z filtrů jsou určeny k potlačení šumu v obraze, tuto kapitolu tedy začneme právě základními informacemi o šumu.

### 6.1 Šum

Šumem rozumíme náhodný signál, část obrazu, která znehodnocuje užitečné obrazové informace. K šumu přispívá každé zařízení, které je při pořizování a uchování obrazu používáno. Je většinou charakterizován vysokými frekvencemi, malými amplitudami a malou korelací s přenášeným signálem. Šum se z obrazu snažíme odstranit nebo ho alespoň potlačit. Aby naše snaha byla úspěšná, je třeba znát původ a především charakteristiky šumu. Šum můžeme dělit podle mnoha hledisek:

#### Podle původce:

- **výstřelový šum** - jeho původcem je kvantová povaha proudu v polovodičích, poměrně slabý
- **plápolavý, praskavý** - jeho původcem jsou většinou výbojky, silný na malých frekvencích, se stoupající frekvencí klesá
- **tepelný** - jeho původcem je chaotický pohyb elektronů a jejich kolize s atomy vodiče, závislý pouze na teplotě, ne na proudu nebo napětí (způsobem většinou odporovými snímači).

#### podle statistických vlastností rozložení na frekvencích

- **bílý** – rovnoměrný výkon na všech frekvencích signálu
- **barevný** – nerovnoměrný výkon na různých frekvencích signálu (růžový –  $1/f$ , hnědý, modrý...)

#### podle statistických vlastností rozložení na intenzitě

- **rovnoměrný** – jeho hodnota má na každé intenzitě signálu rovnoměrné rozdělení. Jednoduše se modeluje – generátory náhodných čísel mají většinou právě toto rozdělení
- **gaussovský** – jeho hodnota má na každé intenzitě signálu normální rozdělení. Jeho modelování je (trochu) komplikovanější, je však velmi blízký realitě.

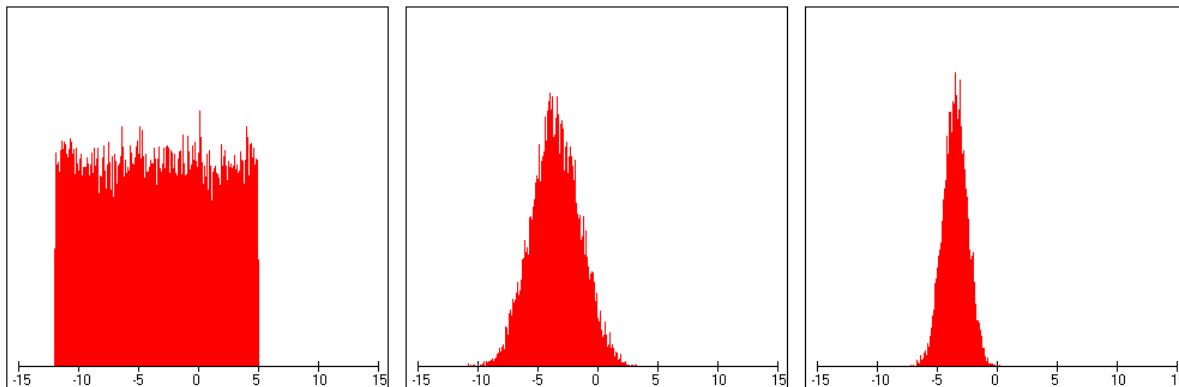
#### podle závislosti na obraze

- **aditivní** – přičítá se k hodnotě pixelu
- **impulsní** - nezávisí na obraze, může být způsoben např. vadnými čipy snímacího zařízení, prachem na objektivu apod.

Nejúčinnější a nejkorektnější metoda potlačování aditivního šumu využívá skutečnosti, že tento šum můžeme považovat za realizaci náhodné veličiny a k jeho potlačení lze tedy v některých případech využít **Linderbergovu-Léviho větu**, která říká: Jestliže  $X_1; X_2; \dots; X_n$  jsou náhodné veličiny s libovolným (ale stejným) rozložením, stejnou střední hodnotou  $\mu$  a stejným (konečným) rozptylem  $\sigma^2$ , pak jejich průměr konverguje k normálnímu rozložení se stejnou střední hodnotou  $\mu$  a rozptylem  $\bar{\sigma}^2 = \sigma^2 / n$ .

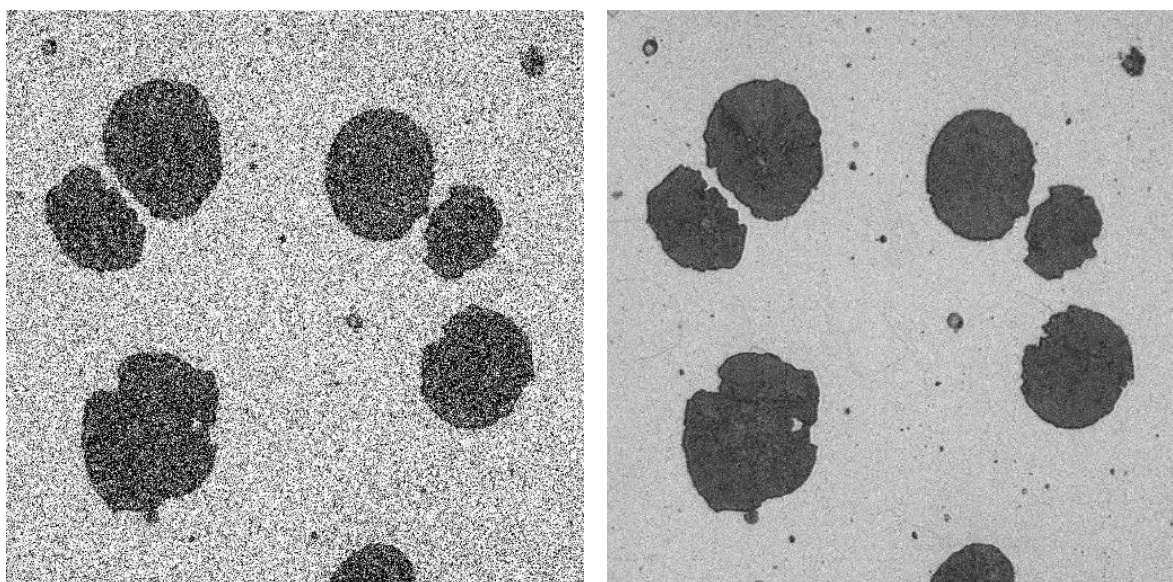


Situace je ilustrována na obr. 6.1.1. Vlevo je generováno rovnoměrné rozložení s rozpětím  $\langle a; b \rangle = \langle -12; 5 \rangle$ . Tato náhodná veličina má tedy střední hodnotu  $\mu = 3.5$  a rozptyl  $\sigma^2 \approx 24$ . Jestliže tyto veličiny budeme sčítat, dostaneme náhodnou veličinu s normálním rozdělením, jehož rozptyl bude s přibývajícím počtem sčítanců klesat. Pro  $n = 6$  je rozptyl normálního rozložení jen  $\bar{\sigma}^2 \approx 4$ ; pro  $n = 24$  pak již jen  $\bar{\sigma}^2 \approx 1$ .



**Obr.6.1.1:** Ilustrace Linderbergovy-Léviho věty: Náhodná veličina uprostřed resp. vpravo vznikla součtem šesti resp. čtyřiceti náhodných veličin vlevo.

Máme-li tedy možnost pořídít více „zcela identických“ snímků, budou se tyto snímky lišit jen aditivním šumem. Ten na každém snímku představuje náhodnou veličinu se stejnou střední hodnotou ( $\mu = 0$ ) a stejným rozptylem. „Užitečná“ informace je na každém snímku stejná. Budeme-li snímky průměrkovat, šum na tomto průměru bude mít normální rozložení se střední hodnotou nula a klesajícím rozptylem. Šum bude tlumen, aniž by došlo ke ztrátě užitečných informací na snímku.



**Obr.6.1.2:** Redukce aditivního šumu průměrováním obrazů. Silný aditivní šum byl podstatně redukován průměrem 25 obrazů.

Ne vždy ovšem máme možnost pořídít sérii takto identických obrazů. Proto jsme odkázáni i na jiné metody odstraňování šumu, a to pomocí obrazových filtrů, o kterých pojednají následující podkapitoly.



## 6. 2 Lineární filtry

Tyto filtry pracují tak, že nahrazují hodnotu zpracovávaného pixelu lineární kombinací hodnot okolních pixelů. Pracují tedy na principu konvoluce vyložené v kpt. 5.3.

**Filtry typu dolní propust:** pracují tak, že „propouštějí“ nízké prostorové frekvence ve smyslu Fourierovy transformace (viz kpt. 5.2.) a potlačují (popř. zcela likvidují) frekvence vysoké. Principiálně je tedy možno postupovat tak, že sestrojíme amplitudové spektrum filtrovaného obrazu, vymačkujeme v něm vysoké frekvence ( viz např. obr. 5.2.3 a 5.2.4). Důsledky takto přímého zásahu do amplitudového spektra se však většinou těžko odhadují, a proto se k účelům filtrace využívá konvoluční teorém - viz vztahy (5.3.2), (5.3.3).

Uvažujme obraz litiny (viz obr. 6.2.1 vlevo) a sestrojme jeho konvoluci s maticí  $C = (c_{ij})$ ;  $i, j = -3; \dots; 3$  kde  $c_{ij} = 1/49$ . Tato konvoluce nahradí ve zpracovávaném obraze hodnotu každého pixelu aritmetickým průměrem hodnot čtyřiceti devíti pixelů ve čtvercovém okolí pixelu zpracovávaného. Výsledkem této konvoluce je obraz  $B = A \otimes C$ , který si můžeme prohlédnout na obr. 6.2.1 vpravo. Zmizely ostré kontrastní přechody mezi kovem a uhlíkovými zrnky, tato konvoluce zřetelně potlačila vysoké prostorové frekvence. Abychom zjistili konkrétně, které frekvence byly potlačeny a jak, využijeme konvoluční teorém.

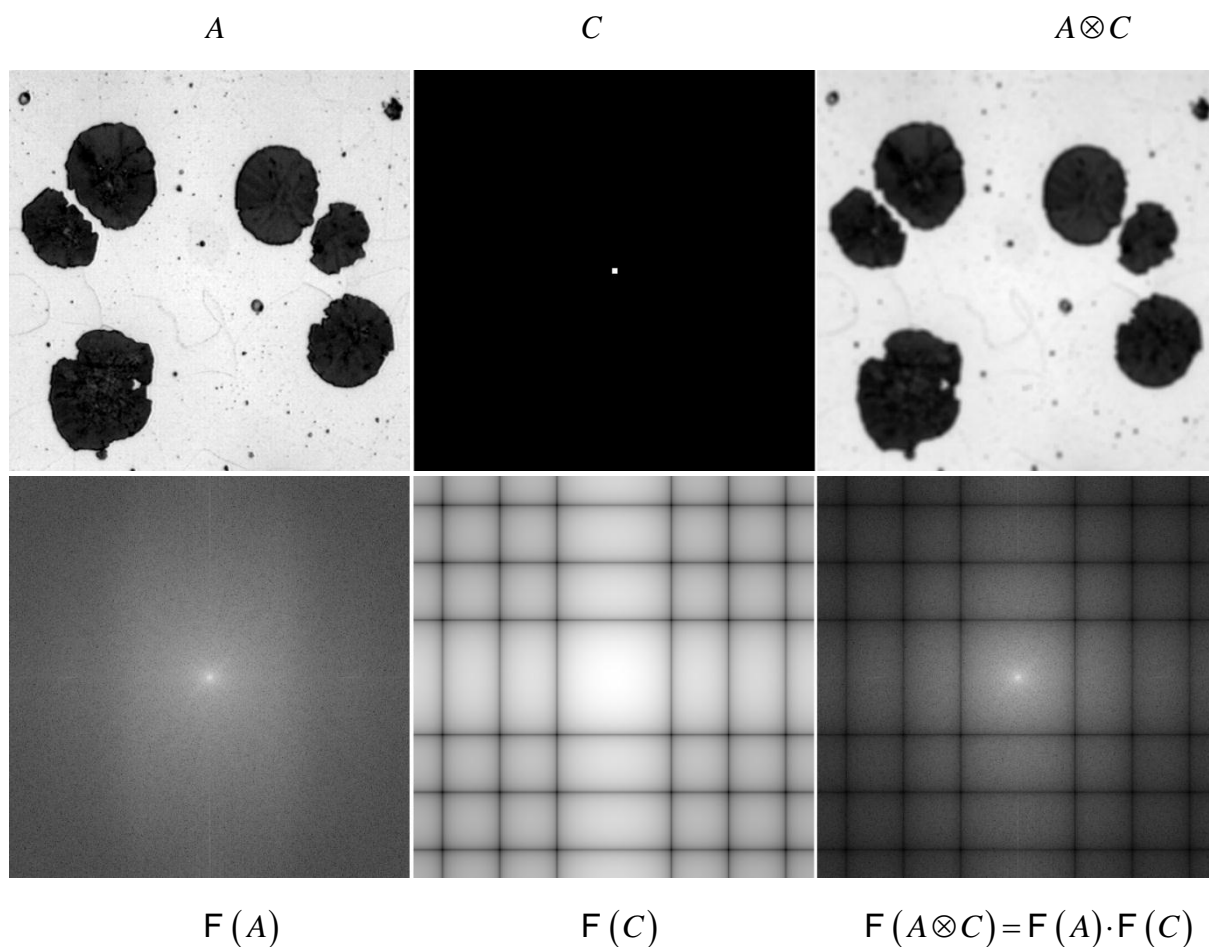
Matici  $C$  z vizualizačních důvodů reprezentujme obrazem - viz obr. 6.2.2 nahoře uprostřed, kde je centrální prvek  $c_{00}$  umístěn ve středu obrazu a pro zvýšení kontrastu je každý prvek  $c_{ij}$  vynásoben číslem  $49 \cdot D = 49 \cdot 255$ . Nyní sestrojme amplitudová spektra obrazů  $A; C$  - viz obr. 6.2.2 dole vlevo a uprostřed. Ve spektru obrazu  $C$  jsou zřetelné tmavé „mřížkovité mezery“, které signalizují potlačené popř. zcela chybějící frekvence. Spektrum  $F(A \otimes C)$  konvoluce  $A \otimes C$  dostaneme podle konvolučního teorému jakou součin spekter  $F(A); F(C)$  obrazů  $A; C$  - viz obr 6.2.2 vpravo dole. Je zřejmé, že matice  $C$  potlačila popř. zcela odstranila ve spektru obrazu  $A$  ty frekvence, které jsou potlačeny popř. zcela odstraněny v jejím spektru.

Ze spektra  $F(C)$  je zřejmé, že filtr  $C$  potlačuje vysoké frekvence jen ve velmi úzkých pásmech a frekvence mezi těmito pásmy (i vysoké) propustí. To má za následek, že výsledek filtrování většinou není dobrý, mnohdy jsou na obraze dokonce patrné slabé několikanásobné kopie původně ostrých hran („duchy“). Amplitudové spektrum filtru by mělo mít jediný extrém, a to v bodě  $F(c_{00})$  (tedy v centrálním prvku) a směrem k okrajům klesat. Jak je vidět z obr 6.2.1, konstantní filtr (tj. filtr složený ze samých jedniček) se k tomuto požadavku ani zdaleka nepřibližuje. Proto se setkáváme s jeho nejružnějšími modifikacemi.

Než je uvedeme, poznamenejme ještě jednu věc. Většinou je potřeba, aby filtrovaný obraz zachoval celkový jas. Toho lze dosáhnout jedině tak, že součet všech prvků konvoluční matice je roven jedné. Konvoluční matice pak mohou být např. tohoto tvaru:

$$C = \begin{pmatrix} \frac{1}{12} & \frac{1}{12} & \frac{1}{12} \\ \frac{1}{12} & \frac{1}{3} & \frac{1}{12} \\ \frac{1}{12} & \frac{1}{12} & \frac{1}{12} \end{pmatrix} = \frac{1}{12} \cdot \begin{pmatrix} 1 & 1 & 1 \\ 1 & 4 & 1 \\ 1 & 1 & 1 \end{pmatrix} \quad (6.2.1)$$

(matice pracuje na tzv. osmiokolí pixelu)



**Obr.6.2.1:** Obráz litiny (vlevo nahoře) a jeho konvoluce (vpravo nahoře) s maticí reprezentovanou obrazem (uprostřed nahoře). Dole příslušná amplitudová spektra.

$$C = \begin{pmatrix} 0 & \frac{1}{5} & 0 \\ \frac{1}{5} & \frac{1}{5} & \frac{1}{5} \\ 0 & \frac{1}{5} & 0 \end{pmatrix} = \frac{1}{5} \cdot \begin{pmatrix} 0 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 0 \end{pmatrix} \quad (6.2.2)$$

(matice pracuje na tzv. čtyřokolí pixelu)

Softwarové prostředky proto často nabízejí možnost zadávat konvoluční matice bez potřebného dělení (tzv. normalizace) a na přání uživatele ji provedou automaticky. Například matice (6.2.1) a (6.2.2) lze tedy zadat takto:

$$C = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 4 & 1 \\ 1 & 1 & 1 \end{pmatrix}; \quad C = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

a po zaškrtnutí možnosti „normalizace“ provede program dělení součtem zadaných prvků automaticky. Následující matice jsou rovněž uvedeny bez normalizace.

Konstantní filtr

Zvýšení váhy  
centrálního prvku

Binomický (Pascalovský) filtr

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix}; \quad \begin{pmatrix} 1 & 1 & 2 & 1 & 1 \\ 1 & 1 & 4 & 1 & 1 \\ 2 & 4 & 8 & 4 & 2 \\ 1 & 1 & 4 & 1 & 1 \\ 1 & 1 & 2 & 1 & 1 \end{pmatrix}; \quad \begin{pmatrix} 1 \\ 4 \\ 6 \\ 4 \\ 1 \end{pmatrix} \cdot (1 \ 4 \ 6 \ 4 \ 1) = \begin{pmatrix} 1 & 4 & 6 & 4 & 1 \\ 4 & 16 & 24 & 16 & 4 \\ 6 & 24 & 36 & 24 & 6 \\ 4 & 16 & 24 & 16 & 4 \\ 1 & 4 & 6 & 4 & 1 \end{pmatrix}$$

Z hlediska frekvenční charakteristiky nemá nejlepší vlastnosti Gaussovský filtr, kde prvky konvolučního jádra jsou hodnoty funkce

$$f(x; y) = \frac{1}{2\pi\sigma_1\sigma_2} \cdot \exp \left[ -\frac{(x-\mu_1)^2}{\sigma_1^2} - \frac{(y-\mu_2)^2}{\sigma_2^2} \right]$$

Např. pro hodnoty  $\sigma_1 = \sigma_2 = 1$ ;  $\mu_1 = \mu_2 = 0$ , má konvoluční matice řádu pět tvar

$$C = \begin{pmatrix} 0 & 0,001 & 0,003 & 0,001 & 0 \\ 0,001 & 0,022 & 0,056 & 0,022 & 0,001 \\ 0,003 & 0,056 & 0,159 & 0,056 & 0,003 \\ 0,001 & 0,022 & 0,056 & 0,022 & 0,001 \\ 0 & 0,001 & 0,003 & 0,001 & 0 \end{pmatrix}$$

Filtry typu dolní propust se používají k vyhlazování obrazů, ke snižování ostroty obrazu a k potlačování šumu. Zde je ovšem třeba poznamenat, že tyto filtry nejsou schopny určit původ vysokofrekvenční informace, a tak spolu s šumem mohou potlačovat i užitečné informace v obraze.

**Filtry typu horní propust:** pracují naopak tak, že propouštějí vysoké prostorové frekvence a potlačují frekvence nízké. To umožňuje zvýšit kontrast drobných detailů v obraze, a naopak snížit kontrast mezi většími světlými a tmavými plochami. Mezi filtry tohoto typu patří např. filtry vytvořené metodou neostré masky známé z fotografie. Tato metoda skládá ostrý pozitiv s neostrým negativem. Výsledkem může být matice normovatelná, anebo nenormovatelná (součet jejich prvků je roven nule). Normovatelné matice lze použít k zaostřování snímků, např.

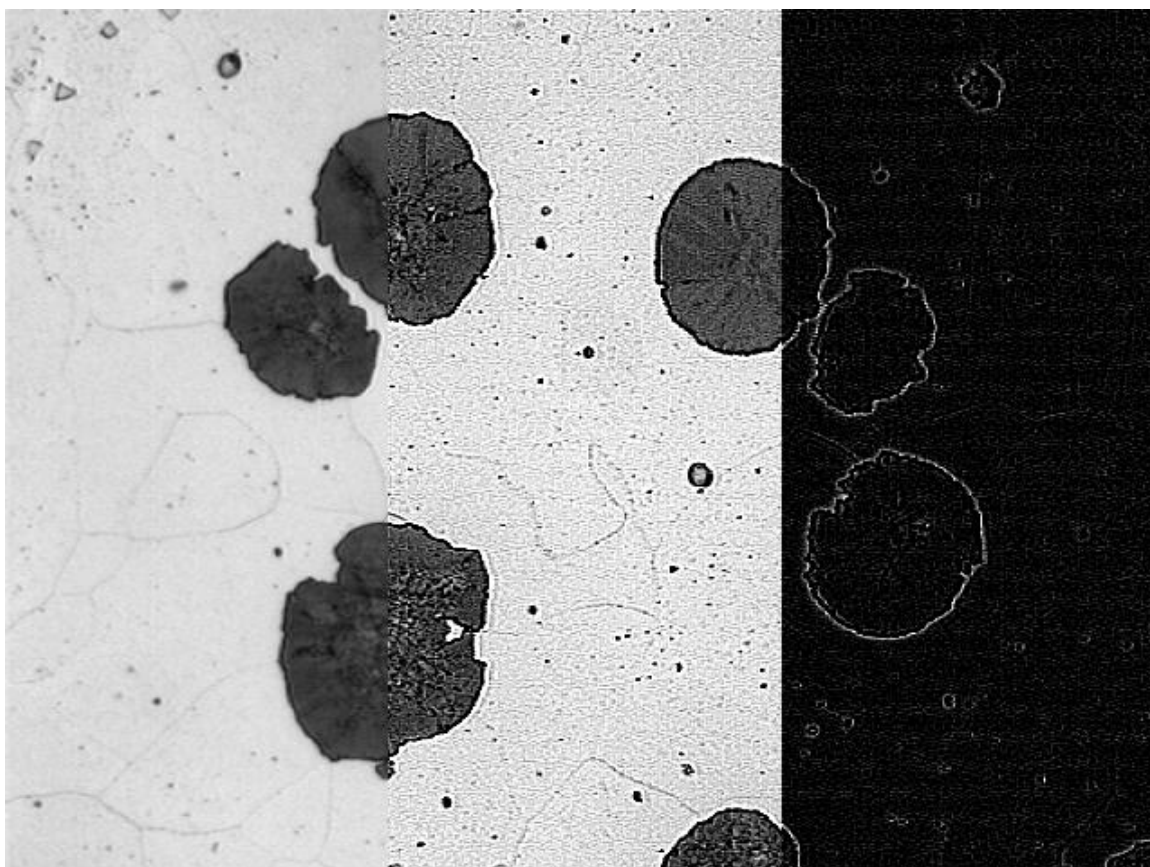
$$\begin{aligned} \mathbf{C} &= a \cdot \mathbf{E} + b \cdot \mathbf{B} = 6 \cdot \underbrace{\begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}}_{\text{původní obraz}} + 1 \cdot \underbrace{\begin{pmatrix} 0 & -1 & 0 \\ -1 & -1 & -1 \\ 0 & -1 & 0 \end{pmatrix}}_{\text{neostrý negativ}} = \underbrace{\begin{pmatrix} 0 & -1 & 0 \\ -1 & 5 & -1 \\ 0 & -1 & 0 \end{pmatrix}}_{\text{čtyřkolí}} \\ \mathbf{C} &= a \cdot \mathbf{E} + b \cdot \mathbf{B} = 10 \cdot \underbrace{\begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}}_{\text{původní obraz}} + 1 \cdot \underbrace{\begin{pmatrix} -1 & -1 & -1 \\ -1 & -1 & -1 \\ -1 & -1 & -1 \end{pmatrix}}_{\text{neostrý negativ}} = \underbrace{\begin{pmatrix} -1 & -1 & -1 \\ -1 & 9 & -1 \\ -1 & -1 & -1 \end{pmatrix}}_{\text{osmiokolí}} \end{aligned}$$

Nenormovatelné matice, např.

$$\mathbf{C} = a \cdot \mathbf{E} + b \cdot \mathbf{B} = 5 \cdot \underbrace{\begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}}_{\text{původní obraz}} + 1 \cdot \underbrace{\begin{pmatrix} 0 & -1 & 0 \\ -1 & -1 & -1 \\ 0 & -1 & 0 \end{pmatrix}}_{\text{neostrý negativ}} = \underbrace{\begin{pmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{pmatrix}}_{\text{čtyřkolí}}$$

$$\mathbf{C} = a \cdot \mathbf{E} + b \cdot \mathbf{B} = 10 \cdot \underbrace{\begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}}_{\text{původní obraz}} + 1 \cdot \underbrace{\begin{pmatrix} -1 & -1 & -1 \\ -1 & -1 & -1 \\ -1 & -1 & -1 \end{pmatrix}}_{\text{neostrý negativ}} = \underbrace{\begin{pmatrix} -1 & -1 & -1 \\ -1 & 9 & -1 \\ -1 & -1 & -1 \end{pmatrix}}_{\text{osmiokolí}}$$

odstraňují nízké frekvence úplně a lze je použít k detekci hran. Ke konvoluci s těmito maticemi se obvykle připočítává konstanta  $\frac{1}{2}$ , a to proto, aby hodnoty nového obrazu byly co nejméně ořezány. Nenormovatelné filtry aproximují tzv. Laplaceův operátor, o kterém se zmíníme v následující kapitole.



**Obr. 6.1.2:** Obraz litiny (originál vlevo) upravený filtrem typu horní propust pracujícím na osmiokolí – normovatelný filtr (uprostřed), nenormovatelný filtr (vpravo).

**Speciální lineární filtry:** Slouží k nejrůznějším speciálním účelům. Jedná se např. o tzv. 1D vyhlazovací filtry

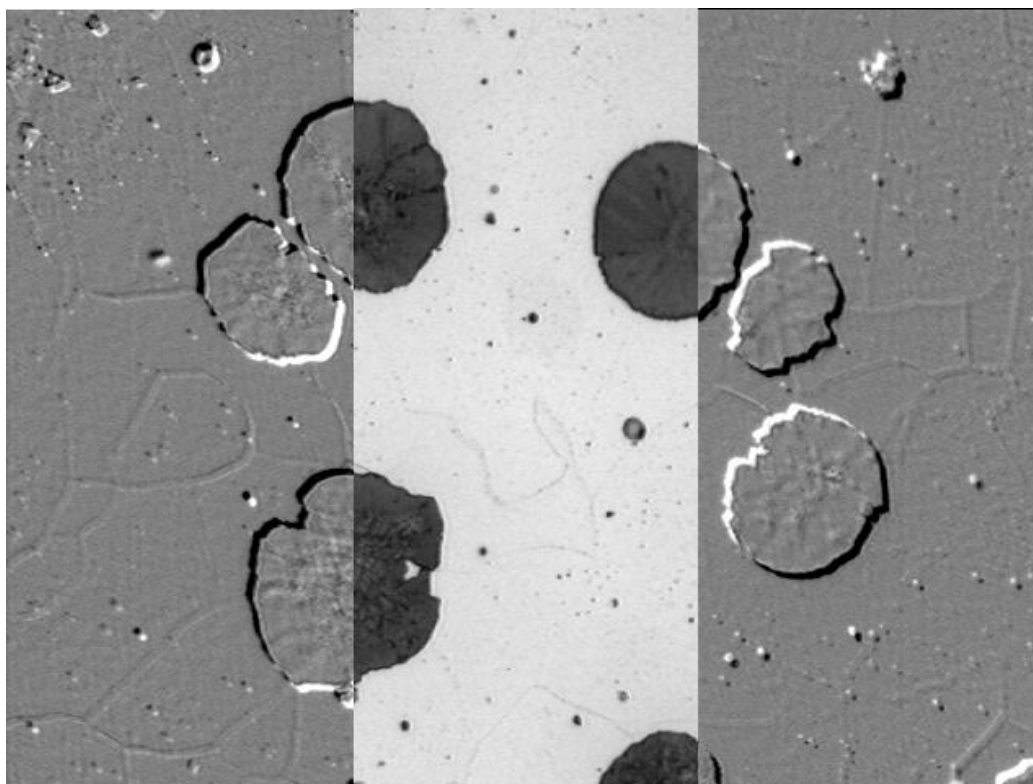
$$C = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix}; \quad C = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

které vyhlazují obraz ve směru řádků resp. sloupců, a používají se k úpravě obrazů, které jsou ve směru řádků či sloupců výrazně „nataženy“ (většinou předchozím použitím procedur typu stretch).

Další skupinou speciálních filtrů jsou filtry typu reliéf (embos), které filtrují obraz do podoby reliéfu osvětleného bočním světlem. Tyto filtry jsou schopny zvýraznit i drobné detaily i v případě, že obraz obsahuje hodně šumu. Tyto filtry skládají pozitiv s posunutým negativem.

Jsou opět nenormovatelné, a proto se opět přičítá hodnota  $\frac{D}{2}$ . Na obr. 6.1.2 je opět snímek litiny, na který byl aplikován filtr

$$C = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \text{ (vlevo)} \quad \text{resp.} \quad C = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix} \text{ (vpravo)}$$



**Obr. 6.1.3:** Obráz litiny (originál uprostřed) upravený filtry typu embos (po stranách).

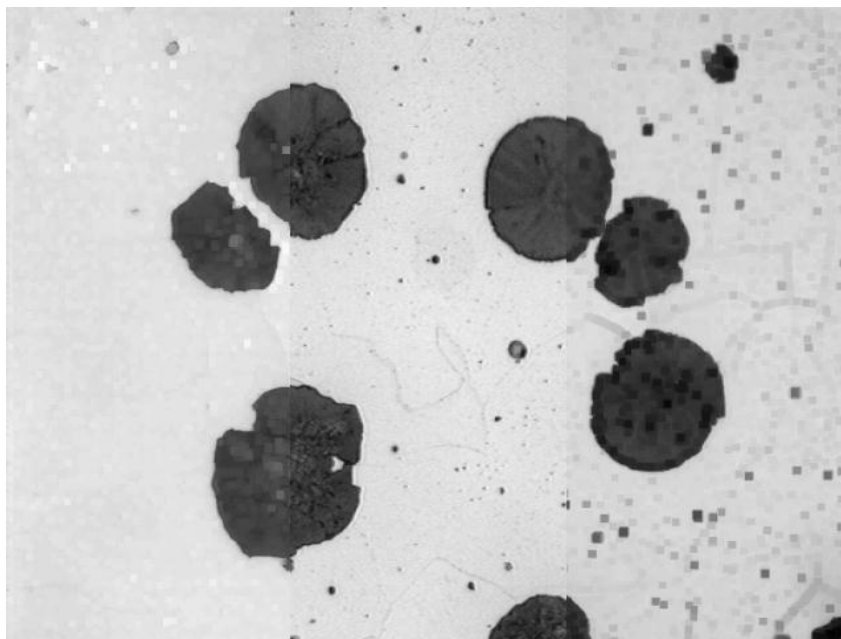
### 6. 3 Nelineární filtry

Tyto filtry pracují rovněž na okolí zpracovávaného pixelu. Nenahrazují ho však lineární kombinací hodnot, ale hodnotou získanou jiným způsobem.

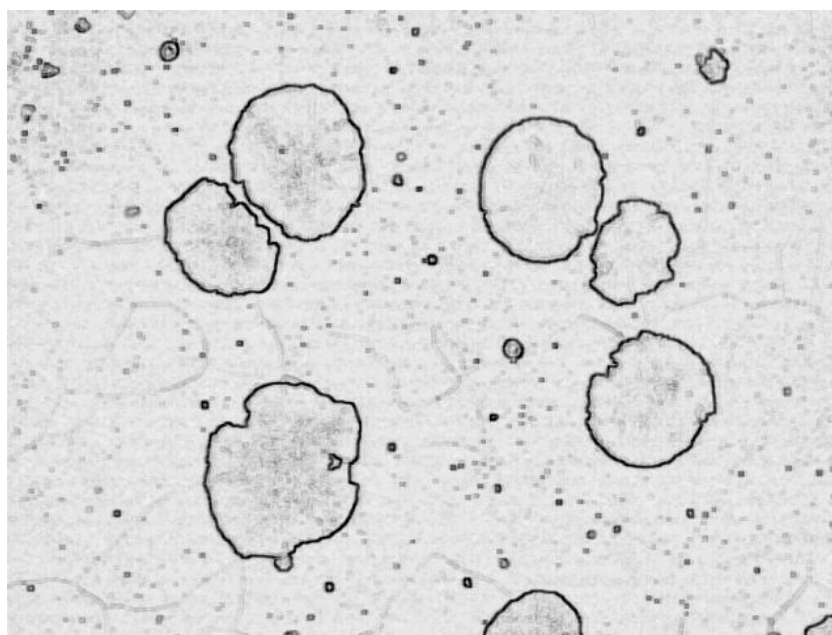
**Filtry typu minimum resp. maximum, dilatace, eroze:** nahrazují hodnotu zpracovávaného pixelu minimální resp. maximální hodnotou pixelů v předem definovaném okolí. Filtry typu minimum zmenšují (erodují) světlé objekty na tmavém pozadí a zvětšují (dilatují) tmavé objekty na světlém pozadí. Filtry typu maximum pracují obráceně. Na obr 6.3.1. vidíme uprostřed opět snímek litiny s tmavými uhlíkovými zrnky, na který byl aplikován filtr typu maximum (vlevo), který zrna erodoval, a minimum (vpravo), který je dilatoval.

Filtr typu minimum může zesilovat tmavé křivky na světlém pozadí a ztenšovat silné světlé křivky na tmavém pozadí (filtr maximum opět obráceně). Dilatace a eroze následovaná odečtením obrazů může sloužit k detekci hranic objektů tak, jak to ilustruje obr. 6.3.2.

Eroze tmavých objektů následovaná jejich dilatací může z obrazu zcela odstranit drobné tmavé objekty (může přitom bohužel dojít ke změně tvaru objektů). Filtry typu minimum mohou „uzavírat“ tmavé objekty a „otevírat“ objekty světlé, filtry typu maximum pracují opět obráceně.



**Obr. 6.3.1:** Účinek filtru typu maximum (vlevo) a minimum (vpravo) na uhlíková zrnka v litině.



**Obr. 6.3.2:** Detekce hranic tmavých objektů odečtením eroze od dilatace

**Mediánový filtr:** nahrazuje hodnotu zpracovávaného pixelu mediánem hodnot pixelů v předem definovaném okolí. Slouží jako velmi účinný prostředek pro potlačování šumu a k vyhlazování obrazů. Na rozdíl od lineárních filtrů typu dolní propust nerozmazává hranice. Podrobněji v následující podkapitole.

**Gradientní filtr:** gradient funkce v bodě je vektor udávající směr největšího růstu funkce několika proměnných. Ve dvojrozměrném případě má tvar :

$$\text{grad } f(x_0; y_0) = \nabla f(x_0; y_0) = \left( \frac{\partial f}{\partial x}(x_0; y_0); \frac{\partial f}{\partial y}(x_0; y_0) \right)$$



Jeho velikost

$$\|\text{grad } f(x_0; y_0)\| = \|\nabla f(x_0; y_0)\| = \sqrt{\left[\frac{\partial f}{\partial x}(x_0; y_0)\right]^2 + \left[\frac{\partial f}{\partial y}(x_0; y_0)\right]^2}$$

udává velikost tohoto růstu. Protože obraz je diskrétní, nahrazují se potřebné parciální derivace diferencemi:

$$\text{grad } f(x_i; y_j) = \nabla f(x_i; y_j) \approx (f(x_{i+k}; y_j) - f(x_{i-k}; y_j); f(x_i; y_{j+k}) - f(x_i; y_{j-k}))$$

$$\|\text{grad } f(x_i; y_j)\| = \|\nabla f(x_i; y_j)\| \approx \sqrt{[f(x_{i+k}; y_j) - f(x_{i-k}; y_j)]^2 + [f(x_i; y_{j+k}) - f(x_i; y_{j-k})]^2}$$

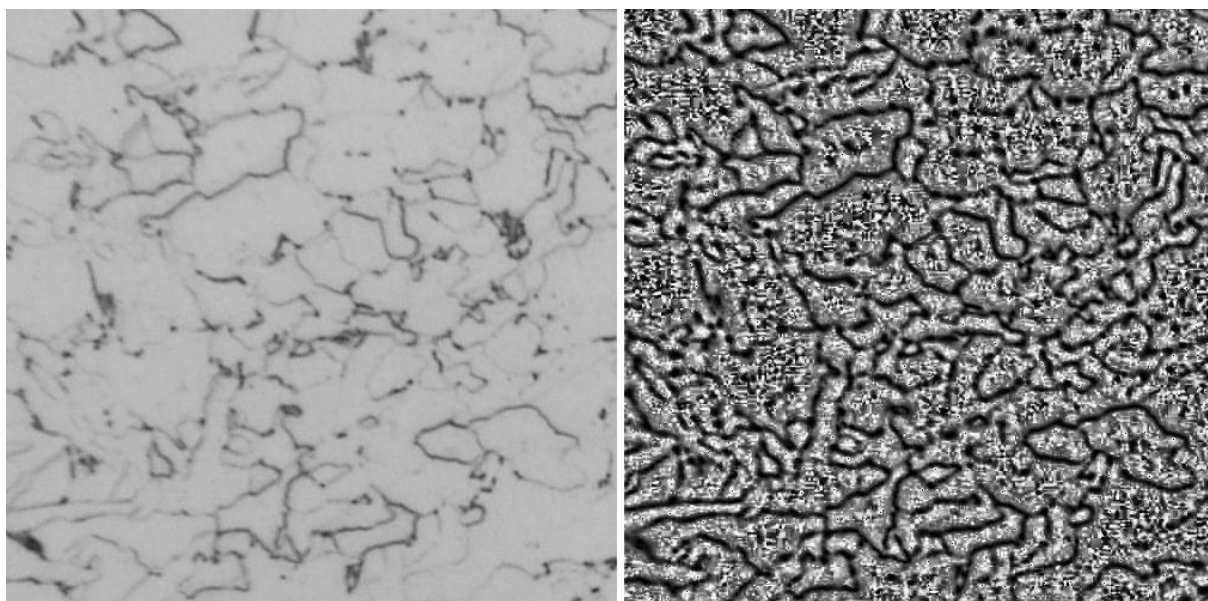
Je-li tedy  $A = (a_{ij})$  původní obraz, pak po aplikaci gradientního filtru obdržíme obraz  $B = (b_{ij})$ , pro jehož pixely platí:

$$b_{ij} = \text{round} \left[ \frac{\sqrt{2}}{2} \cdot \sqrt{[a_{i+k;j} - a_{i-k;j}]^2 + [a_{i;j+k} - a_{i;j-k}]^2} \right]$$

kde se většinou volí  $k = 1$  a konstanta  $\sqrt{2}/2$  funguje jako ochrana před přetečením.

Filtr zvýrazňuje místa, kde dochází k rychlým změnám hodnot pixelů. Slouží tedy i detekci hranic a ke zvýrazňování jemných struktur v obraze.

**Klouzavá ekvalizace:** nahradí hodnotu zpracovávaného pixelu hodnotou získanou ekvalizací hodnot pixelů v předem definovaném okolí. Používá se všude tam, kde obrazy obsahují málo kontrastní detaily, které jsou špatně zřetelné, anebo dokonce nepozorovatelné, přitom obraz má celkově vysoký kontrast. Velikost zvýrazňovaných detailů se dá velmi jemně nastavit volbou velikosti okolí.



**Obr. 6.3.3:** Mikrosnímek oceli upravený klouzavou ekvalizací s okolím 19.

## 6. 4 Adaptivní filtry

Tyto filtry jsou (na rozdíl od předchozích) schopny přizpůsobit svoji činnost lokálním vlastnostem obrazu. Nepracují s celým okolím pixelu, ale jen s tzv.  $k-w$  **clusterem**. Je to podmnožina čtvercového okolí zpracovávaného pixelu se stranou  $w$ , která může být definována v podstatě dvojím způsobem. Buď je to množina pixelů, jejichž hodnoty se od hodnoty zpracovávaného pixelu neliší o více než  $k$ , anebo je to podmnožina, která obsahuje právě  $k$  pixelů. Své „adaptivní varianty“ přitom může mít většina filtrů uvedených v předchozích dvou kapitolách (jedinou výjimkou je gradient). Uvedme především následující dvě:

**Adaptivní průměr** používáme k potlačení aditivního šumu. Má daleko lepší vlastnosti než „neadaptivní varianta“, neboť vhodnou volbou  $k-w$  clusteru lze dosáhnout potlačení šumu bez výraznějšího zhoršení kvality obrazu. Volbu je třeba provést následujícím způsobem:

$k$  je obsah (v pixelech) nejmenšího objektu, který má být v obraze zachován.

$w$  je voleno tak, aby  $\sqrt{k} \leq w \leq 2k$ .

Se vzrůstajícím  $w$  roste adaptivnost filtru, filtr méně rozmazává hranice. Zároveň však klesá (podstatně pomaleji) schopnost filtrovat šum.

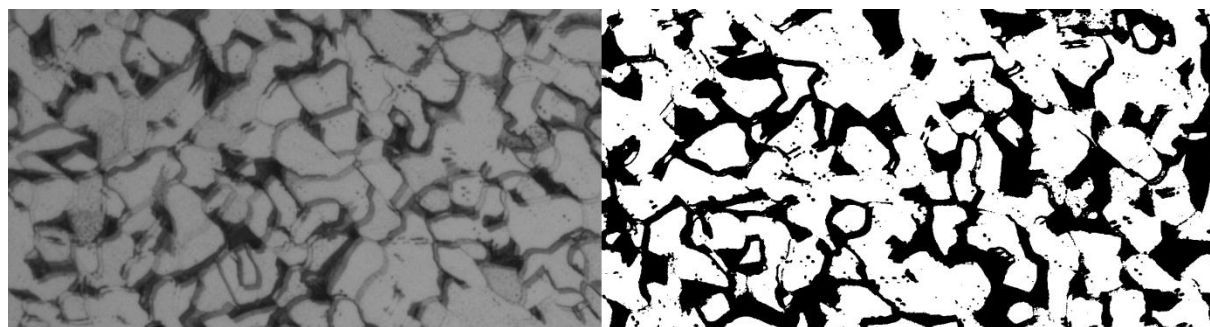
**Adaptivní medián** je používán ke stejným účelům jako adaptivní průměr, navíc je možno odstranit malé objekty z obrazu či detekovat, a to (na rozdíl od „neadaptivní varianty“) beze změny tvaru a velikosti objektů větších.

## 7 Měření objektů

**Objektem** budeme v této kapitole rozumět souvislou oblast, která je zobrazena stejnou barvou. Na snímcích nejrůznějších materiálů se takové objekty většinou nevyskytují, proto je musíme vhodným způsobem definovat. Jedním z možných a často používaných způsobů je prahování

### 7. 1 Prahování

Prahování je metoda, kterou uměle snižujeme počet kvantizačních úrovní (viz kpt. 1.2.) v obraze. Tím docílíme toho, že pixely, jejichž hodnoty se původně lišily, budou zobrazeny stejnou barvou. Nejhrubší prahování je tedy na dvě úrovně (černou a bílou), přičemž hranici jasu která o barvě rozhoduje (práh) je možno volit. Poznamenejme, že prahování je vhodné nejen pro účely měření objektů, ale může zkvalitnit např. dilataci a erozi objektů a identifikaci jejich hranice.



**Obr. 7.1.1:** Prahování snímku oceli

## 7.2 Geometrické momenty objektu

**Geometrické momenty** jsou nástrojem pro analýzu velikosti a tvaru jednotlivých objektů. Obecný geometrický moment rovinného útvaru  $G$  je definován vztahem

$$M(m;n) = \iint_G x^m y^n dx dy$$

Číslo  $m+n$  se nazývá **řád momentu**. Vzhledem k diskrétní povaze obrazu je třeba nahradit tento integrál sumou, tj.

$$M(m;n) = p \cdot \sum_{[r;s] \in G} r^m s^n$$

kde  $p$  je plošný obsah jednoho pixelu. V dalším textu budeme tento obsah považovat za jednotkový. Položíme tedy  $p=1$ , v tom případě dostáváme geometrický moment tvaru

$$M(m;n) = \sum_{[r;s] \in G} r^m s^n \quad (7.2.1)$$

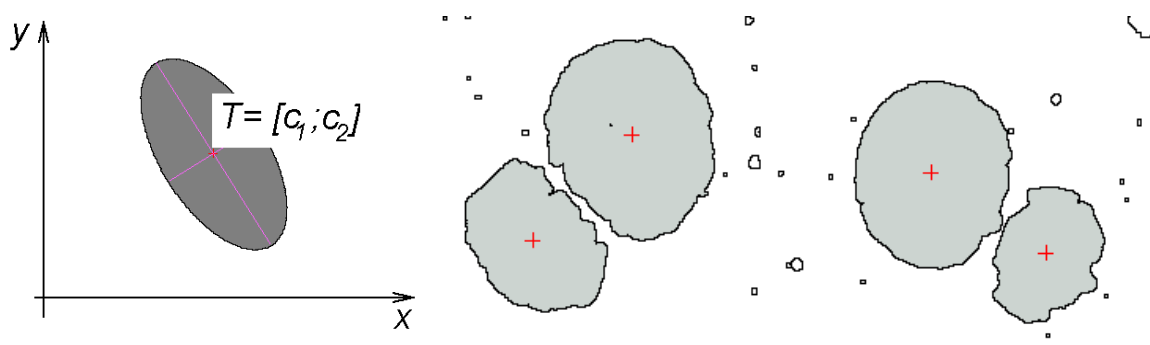
**Geometrický moment  $M(0;0)$** : dosadíme-li do (7.2.1)  $m=n=0$ , dostaneme

$$M(0;0) = \sum_{[r;s] \in G} r^0 s^0 = \sum_{[r;s] \in G} 1$$

Tento geometrický moment je tedy roven obsahu objektu. V našem případě, kdy obsah pixelu považujeme za jednotkový, je zároveň roven počtu pixelů, které objekt obsahuje.

**Těžiště objektu**: je charakteristika vhodná k popisu polohy objektu v obraze.

$$T = [c_1; c_2] = \left[ \frac{M(1;0)}{M(0;0)}; \frac{M(0;1)}{M(0;0)} \right] \quad (7.2.2)$$



**Obr. 7.2.1:** Těžiště elipsy (vlevo) a uhlíkových zrn v litině (vpravo) určená dle (7.2.2.)

Geometrický moment  $M(0;0)$  objektu se nezmění, jestliže objekt zobrazíme nějakým shodným zobrazením. Říkáme, že moment  $M(0;0)$  je invariantní vůči shodnostem. Tento moment však není invariantní vůči dalším geometrickým transformacím. Geometrické momenty vyšších řádů nejsou invariantní ani vůči shodnostem. Při popisu geometrických vlastností objektů nás však zajímají charakteristiky, které naopak invariantní jsou. Například

to, co intuitivně nazýváme tvar objektu, zřejmě nezávisí na jeho poloze, ani na velikosti. Tato charakteristika by tedy měla být invariantní vůči shodnostem i změně měřítka.

**Centrální momenty:** jsou momenty vztažené k souřadnicové soustavě s počátkem v těžišti objektu, jejíž osy jsou rovnoběžné s osami původní soustavy (viz obr. 7.2.2.). Jsou definovány vztahy

$$CM(m;n) = \iint_G (x-c_1)^m (y-c_2)^n dx dy$$

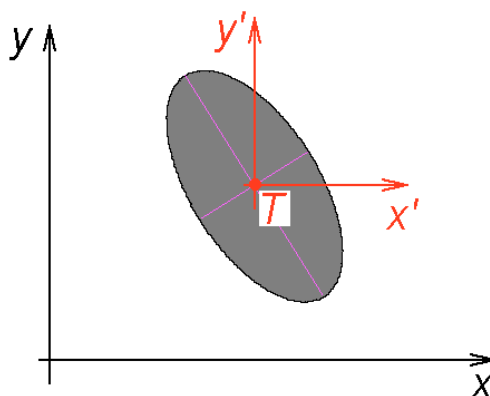
resp.

$$CM(m;n) = \sum_{[r;s] \in G} (r-c_1)^m (s-c_2)^n$$

Zřejmě platí

$$CM(0;0) = M(0;0); CM(1;0) = CM(0;1) = 0$$

Centrální momenty jsou invariantní vůči posunutí, mění se však otočením a změnou měřítka.



**Obr. 7.2.2:** K definici centrálních momentů objektu

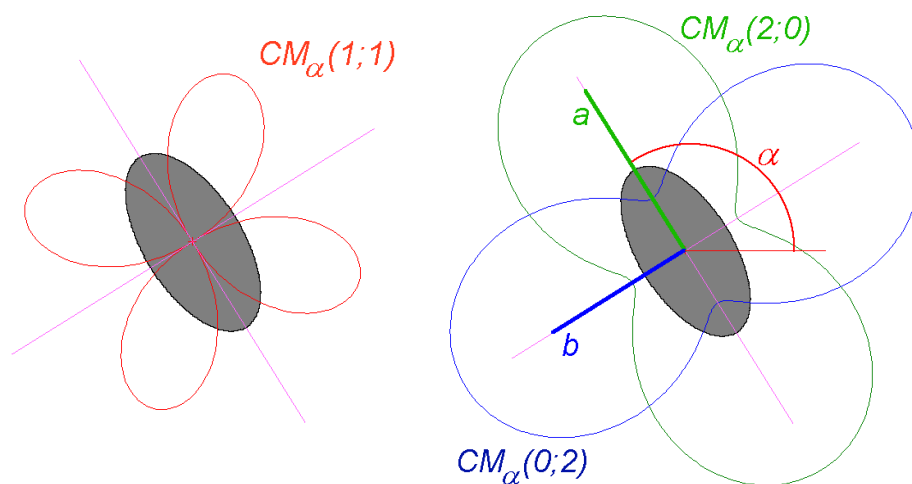
**Normované centrální momenty:** jsou momenty vztažené ke stejné souřadnicové soustavě jako centrální momenty, centrální moment je však normován momentem  $M(0;0)$ . Platí

$$NCM(m;n) = \frac{CM(m;n)}{M(0;0)^{\frac{m+n+2}{2}}}$$

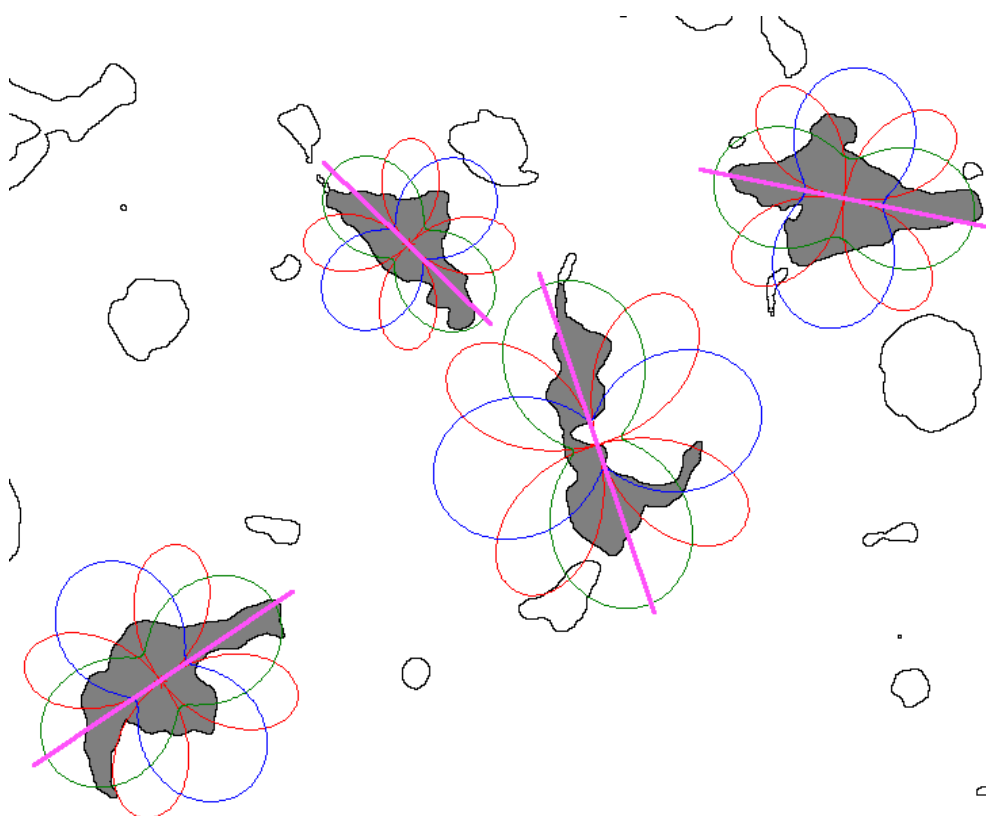
Normované centrální momenty: jsou invariantní vůči posunutí a změně měřítka a nejsou invariantní vůči otáčení. Umožňují však toto otočení kvantifikovat. Na obr. 7.2.3. je znázorněn průběh centrálních momentů  $CM_\alpha(1;1); CM_\alpha(2;0); CM_\alpha(0;2)$  elipsy v závislosti na úhlu  $\alpha$  její hlavní osy vůči kladné poloosy x. Hlavní osa objektu má směr  $\alpha$  určený maximy momentu  $CM_\alpha(2;0)$  resp. minimy momentu  $CM_\alpha(0;2)$ .

Na obr. 7.2.4 je několik tvarově značně rozdílných objektů (zrn uhlíku v oceli) a průběh momentů  $CM_\alpha(1;1); CM_\alpha(2;0); CM_\alpha(0;2)$  s vyznačenou **hlavní osou** objektu, jejíž odchylka od kladné poloosy x určuje orientaci objektu.

**Legendreova (ekvimomentální) elipsa objektu:** je elipsa, která má s daným objektem shodné všechny momenty až do řádu dva včetně. Lze ukázat, že její hlavní poloosa má směr hlavní osy objektu a velikosti poloos jsou



**Obr. 7.2.3:** Závislost centrálních momentů 2. řádu na úhlu otočení.



**Obr. 7.2.3:** Užití centrálních momentů 2. řádu k určení směru hlavních osy objektu

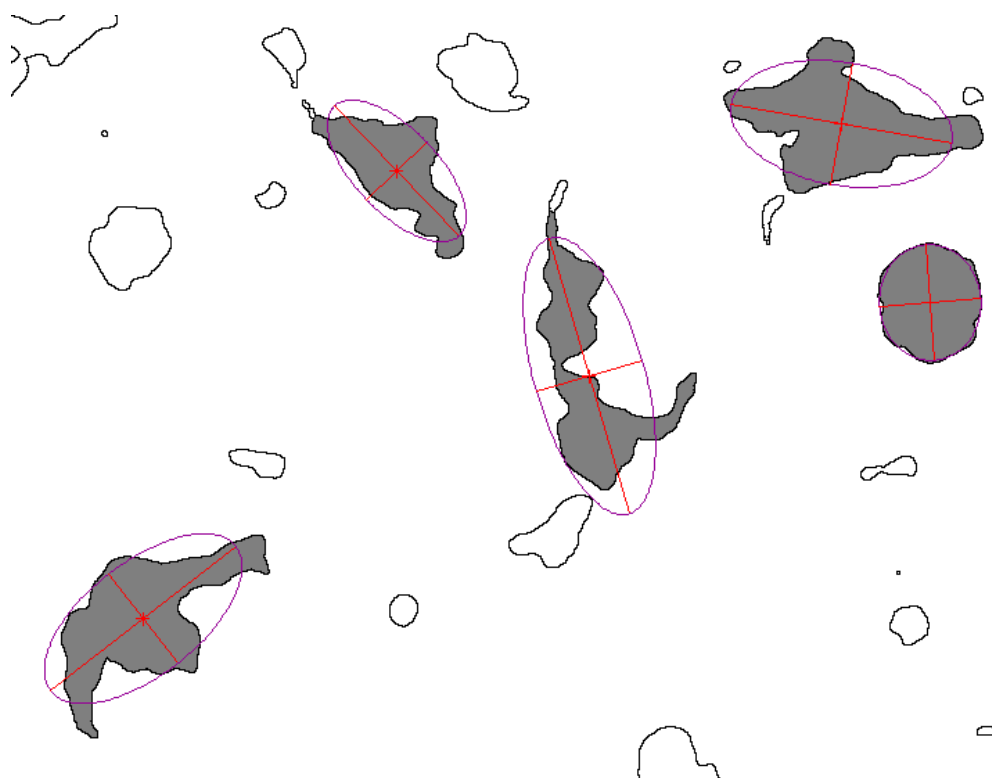
$$a = 2 \cdot \sqrt{PM(2;0)} ; b = 2 \cdot \sqrt{PM(0;2)}$$

kde

$$PM(2;0) = NCM(2;0) \cdot \cos^2 \alpha + 2 \cdot NCM(1;1) \cdot \sin \alpha \cos \alpha + NCM(0;2) \cdot \sin^2 \alpha$$

$$PM(0;2) = NCM(2;0) \cdot \sin^2 \alpha - 2 \cdot NCM(1;1) \cdot \sin \alpha \cos \alpha + NCM(0;2) \cdot \cos^2 \alpha$$

jsou tzv. hlavní momenty (principal moments) objektu, a to hlavní momenty 2. řádu.



**Obr. 7.2.4:** Legendreovy (ekvimomentální) elipsy objektu

Parametry Legendreovy elipsy mohou sloužit jako „tvarové“ charakteristiky objektu:

Elongace  $EL = \log_2 \frac{a}{b}$

charakterizuje „štíhlost“ objektu. Kruh má  $EL = 0$ , elipsa s poměrem os 2:1 má  $EL = 1$  atd.

Disperse  $DP = \log_2 \left[ \frac{\pi ab}{M(0;0)} \right]$

Charakterizuje „odlišnost od elipsy“. Elipsa má  $DP = 0$ .

Extenze  $EX = EL + DP$

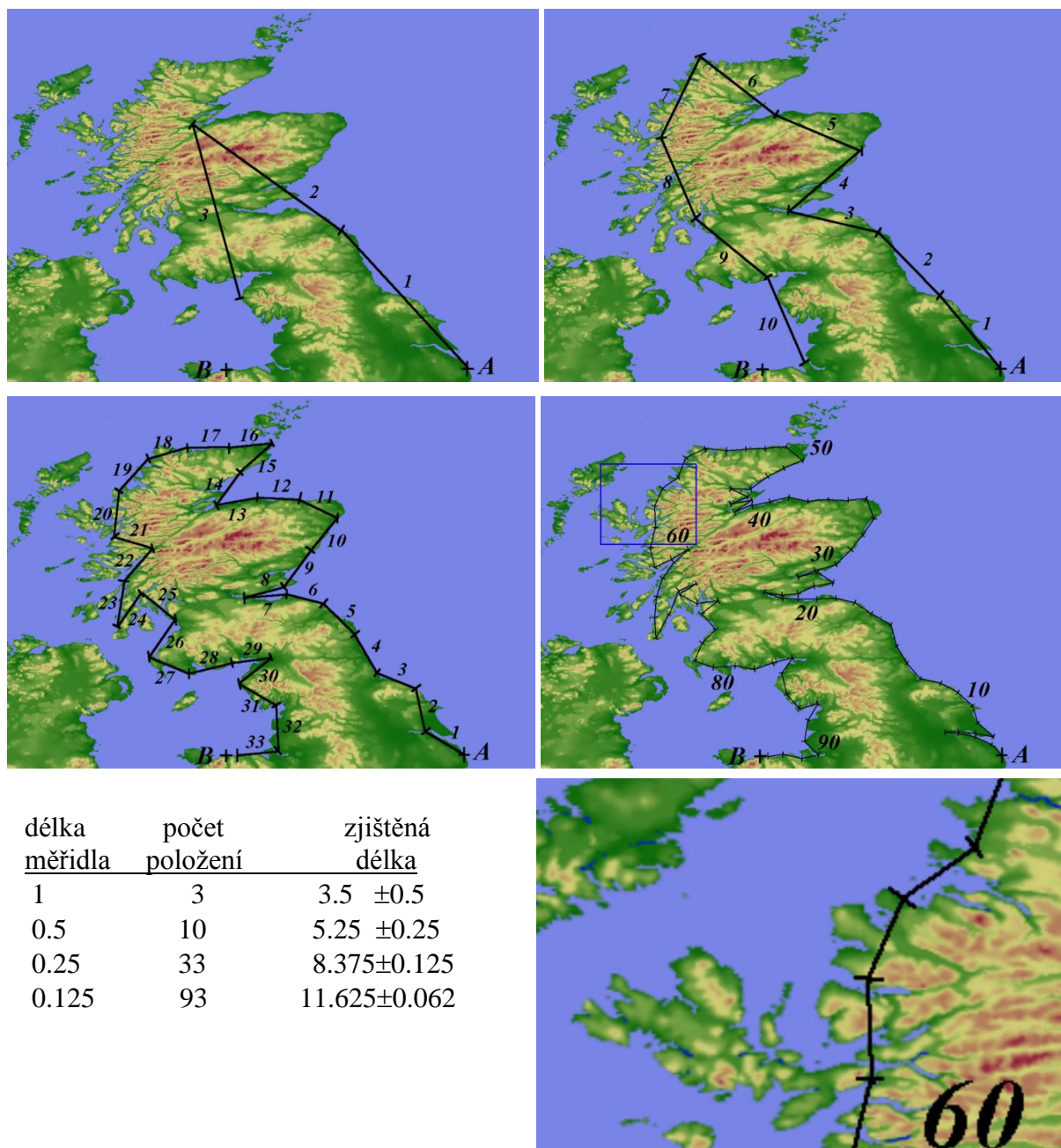
Charakterizuje „odlišnost od kruhu“. Kruh má  $EX = 0$ .

### 7. 3 Fraktální dimenze objektů

Měříme-li délku hrany stolu s přesností na centimetry a pak měření zopakujeme s milimetrovým měřítkem, budou se naše výsledky lišit nanejvýš o půl centimetru. Se zkracujícím se měřidlem bude zjišťovaná délka zpřesňována. Měříme-li délku mořského pobřeží, je to podstatně problematičtější. Zpřesňování měření má za následek mnohonásobné „prodlužování“ pobřeží. Celá situace je ilustrována na následujících obrázcích, kde je modelově měřena délka severní části pobřeží Velké Británie mezi body *A*, *B*. Délka nejdelšího měřidla byla brána za jednotku a na dalších obrázcích postupně zkracována vždy



na polovinu. V připojených tabulkách se můžeme přesvědčit o neustálém „prodlužování“ pobřeží. Na zvětšeném výřezu je patrné, že ani po posledním zkrácení měřidla nejsme délece pobřeží o nic blíže než na začátku.

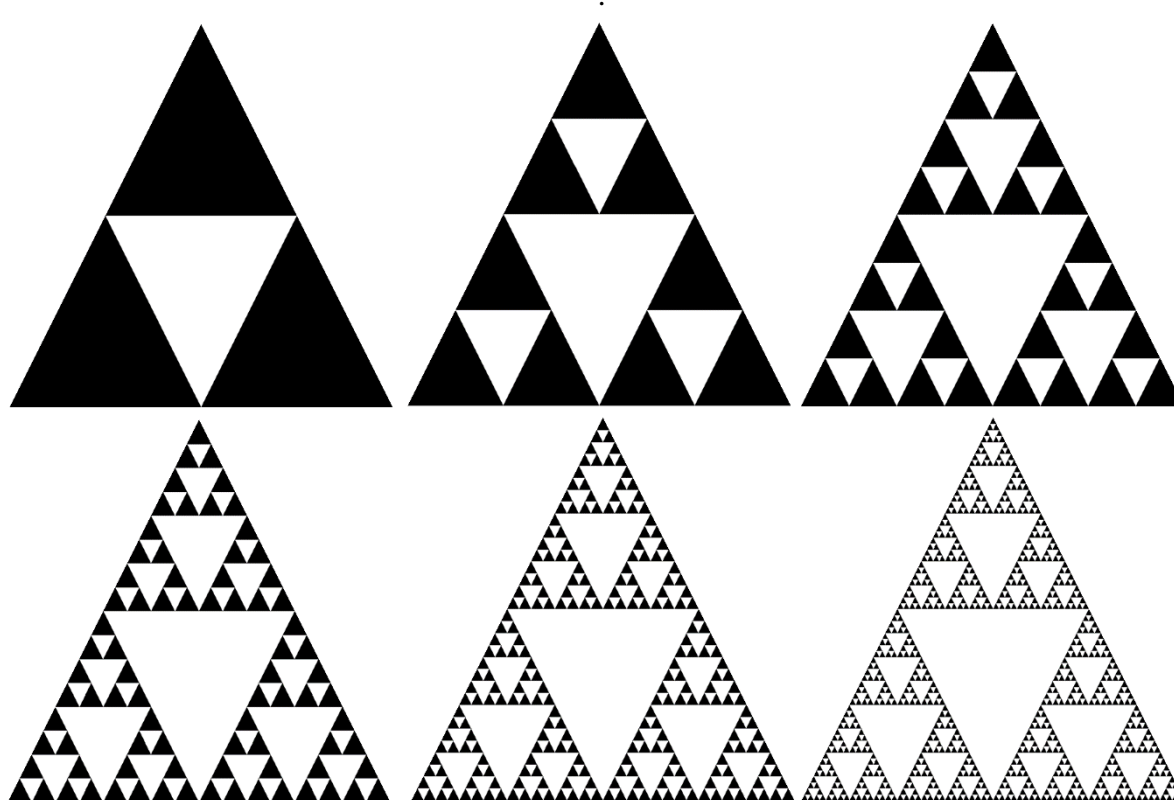


**Obr. 7.3.1:** Měření délky pobřeží Velké Británie

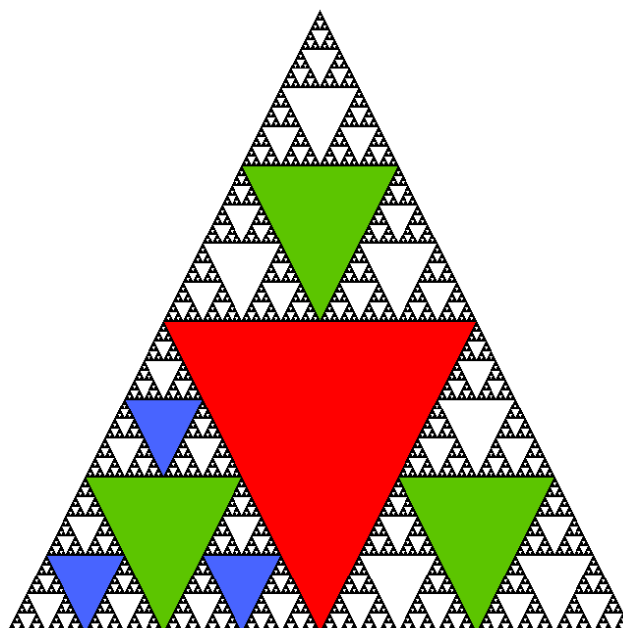
V 50. letech minulého století došel **L. F. Richardson** na základě studia empirických dat k závěru, že kdybychom k měření používali stále kratší tyč, došli bychom k závěru, že všechna pobřeží jsou stejně, a to nekonečně dlouhá. To bylo na první pohled zcela absurdní. Podobné vlastnosti však vykazují nejen mořská pobřeží, ale prakticky všechny reálné útvary.

Tuto zdánlivou absurdnost odhalil **Benoit Mandelbrot** a položil tak základy fraktální geometrie. Pokusme se naznačit význam těchto pojmů.

Začneme tím, že si také vyrobíme podobně „absurdní“ útvar. Vezmeme rovnoramenný trojúhelník, v němž se výška na základnu rovná základně, a vyjmeme z něho vnitřek trojúhelníka sestaveného z jeho středních příček. Na každý ze tří trojúhelníků, které takto vzniknou, aplikujeme tutéž konstrukci a takto pokračujeme do nekonečna. Na obr 7.3.1. je několik prvních kroků této konstrukce, jejímž autorem je polský matematik **Sierpinski**.



**Obr. 7.3.2:** Několik prvních aproximací Sierpinského trojúhelníka



**Obr. 7.3.3:** K obsahu Sierpinského trojúhelníka

Základní charakteristikou geometrického útvaru je jeho dimenze, tj. „počet jeho rozměrů“. Je-li ohraničený útvar jednorozměrný, má konečnou délku a nulový obsah. Je-li dvojrozměrný, má nekonečnou délku a konečný obsah. Otázkou nyní je, kolikadimenzionální je Sierpinského trojúhelník. Předpokládejme nejprve, že je dvojrozměrný a pokusme se spočítat jeho obsah. Označme obsah původního trojúhelníka  $S_0$  a vypočtěme obsah vyjmutých částí:

$$S_1 = \frac{1}{4} S_0$$

$$S_2 = 3 \cdot \frac{1}{4} S_1 = 3 \cdot \frac{1}{4^2} S_0$$

$$S_3 = 3 \cdot \frac{1}{4} S_2 = 3^2 \cdot \frac{1}{4^3} S_0$$

.....

$$S_n = \frac{3^n}{4^{n+1}} S_0$$

.....

$$S = \sum_{n=0}^{\infty} \frac{3^n}{4^{n+1}} \cdot S_0 = \frac{1}{3} S_0 \cdot \sum_{n=0}^{\infty} \frac{3^{n+1}}{4^{n+1}} = \frac{1}{3} S_0 \cdot \sum_{n=1}^{\infty} \left(\frac{3}{4}\right)^n = \frac{1}{3} S_0 \cdot \frac{\frac{3}{4}}{1 - \frac{3}{4}} = S_0$$

Obsah vyjmutých částí je tedy přesně roven obsahu původního trojúhelníka, což znamená, že obsah Sierpiňského trojúhelníka je nulový. Sierpiňského trojúhelník tedy není dvojrozměrný útvar a měl by být jednorozměrný. V tom případě by měl mít konečnou délku. Pokud se ji ale pokusíme spočítat, dospějeme k divergentní geometrické řadě: Označme obvod původního trojúhelníka  $o_0$ , obvody postupně vyjmutých trojúhelníků jsou

$$o_1 = \frac{1}{2} o_0$$

$$o_2 = 3 \cdot \frac{1}{2} o_1 = 3 \cdot \frac{1}{2^2} o_0$$

$$o_3 = 3 \cdot \frac{1}{4} o_2 = 3^2 \cdot \frac{1}{2^3} o_0$$

.....

$$o_n = \frac{3^{n-1}}{2^n} o_0$$

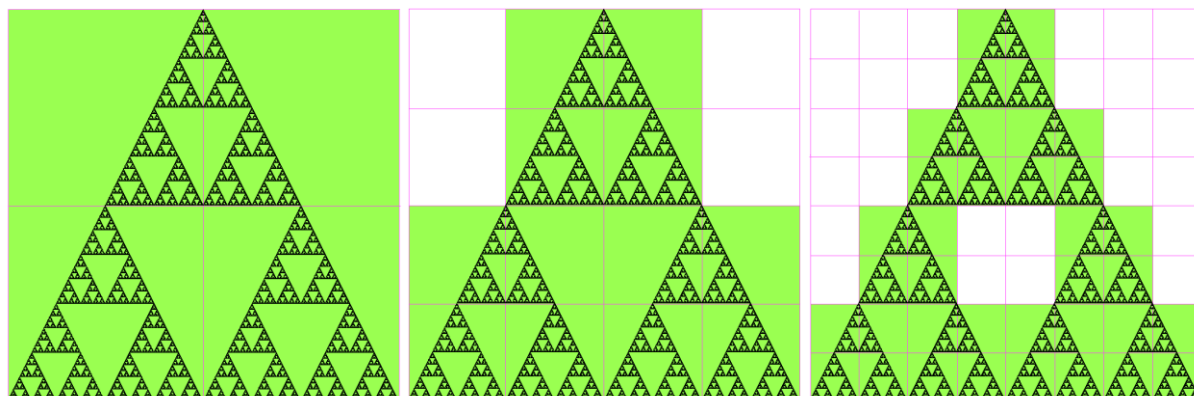
.....

$$o = \sum_{n=1}^{\infty} \frac{3^{n-1}}{2^n} \cdot o_0 = \frac{1}{2} o_0 \cdot \sum_{n=0}^{\infty} \left(\frac{3}{2}\right)^n = \infty$$

„Počet rozměrů“ tohoto útvaru bychom tedy měli hledat někde mezi jedničkou a dvojkou. Jestliže ovšem není celočíselný, je pojem „počet“ poněkud nepatřičný a nahrazujeme ho pojmem **dimenze**. Dimenzi, která připouští neceločíselné hodnoty, nazýváme **dimenzí fraktální**.

Zavedme nyní jednu z možných fraktálních dimenzí. Sierpiňského trojúhelník pokrývejme zjemňující se čtvercovou sítí – viz obr 7.3.4. Zmenšující se velikosti stran čtverce označme

$v_0; v_1; \dots$  a sledujme počty  $p_0; p_1; \dots$  čtverců potřebných k pokrytí útvaru. V případě jednorozměrného útvaru by se součet  $p_n \cdot v_n^1$  měl blížit délce útvaru, v případě útvaru dvojrozměrného by se měl součet  $p_n \cdot v_n^2$  blížit obsahu. V našem případě je ale



$n$	$v_n$	$p_n$	$p_n \cdot v_n^1$	$p_n \cdot v_n^D$	$p_n \cdot v_n^2$
0	1	4	4	<i>konst</i>	4
1	$\frac{1}{2}$	12	6	<i>konst</i>	3
2	$\frac{1}{4}$	36	9	<i>konst</i>	2,25
$k$	$\frac{1}{2^k}$	$4 \cdot 3^k$	$4 \cdot \left(\frac{3}{2}\right)^k$	<i>konst</i>	$4 \cdot \left(\frac{3}{4}\right)^k$
$\rightarrow \infty$	$\rightarrow 0$	$\rightarrow \infty$	$\rightarrow 0$	$\rightarrow \text{konst}$	$\rightarrow \infty$

**Obr. 7.3.4:** K fraktální dimenzi

$$\lim p_n \cdot v_n^1 = \infty; \quad \lim p_n \cdot v_n^2 = 0$$

Je tedy přirozené zobecnit „počet rozměrů“ útvaru (dimenzi) následujícím způsobem:

**Fraktální dimenze útvaru  $U$ :** Nechť  $p_n$  je počet nepřekrývajících se čtverců o straně  $v_n$  nutných k pokrytí útvaru  $U$ ; přičemž  $\lim v_n = 0$ . Dále nechť existuje číslo  $D \in \mathbb{R}$  takové, že

$$0 < \lim p_n \cdot v_n^D < \infty \quad (7.3.1)$$

Pak číslo  $D$  nazýváme **fraktální dimenzí** útvaru  $U$  a číslo

$$M = \lim p_n \cdot v_n^D \quad (7.3.2)$$

nazýváme **mírou** útvaru v dimenzi  $D$ .

Míra útvaru v neceločíselné dimenzi sama o sobě příliš velký význam nemá. U útvarů s neceločíselnou dimenzí je samostatným problémem určit nikoli míru, ale právě dimenzi, jak ukazuje následující příklad.

Určeme fraktální dimenzi Sierpiňského trojúhelníka. Pro jeho míru musí dle (7.3.1) platit

$$0 < \lim p_n \cdot v_n^D < \infty$$

K posloupnostem čísel  $\{p_n\}_{n \in \mathbb{A}}$ ;  $\{v_n\}_{n \in \mathbb{A}}$  sestrojme posloupnost  $\{D_n\}_{n \in \mathbb{A}}$  tak, aby pro každé  $n \in \mathbb{A}$  platilo

$$p_n \cdot v_n^{D_n} = M$$

Tedy 
$$D_n = \frac{\ln M}{\ln v_n} - \frac{\ln p_n}{\ln v_n}$$

Hledaná dimenze  $D$  bude zřejmě limitou

$$D = \lim D_n = \lim \left[ \frac{\ln M}{\ln v_n} - \frac{\ln p_n}{\ln v_n} \right] \quad (7.3.3)$$

Protože  $M \in \mathbb{R}$  a  $v_n \rightarrow 0+$ , je

$$\lim \frac{\ln M}{\ln v_n} = 0$$

takže ze (7.3.3) dostáváme

$$D = -\lim \frac{\ln p_n}{\ln v_n} = -\lim \frac{\ln(4 \cdot 3^n)}{\ln 2^{-n}} = \lim \left[ \frac{\ln 4}{\ln 2^n} + \frac{\ln 3^n}{\ln 2^n} \right] = \lim \frac{n \ln 3}{n \ln 2}$$

Tedy 
$$D = \frac{\ln 3}{\ln 2} \approx 1,584\ 962... \quad (7.3.4)$$

Vztah (7.3.1) poskytuje návod, jak fraktální dimenzi měřit. Vzhledem k diskrétní povaze útvaru nemůžeme realizovat limitu, ale položíme přibližně přímo

$$M \approx p_n \cdot v_n^D$$

Tento vztah zlogaritmujeme

$$\ln M \approx \ln p_n + D \cdot \ln v_n$$

a upravíme na tvar

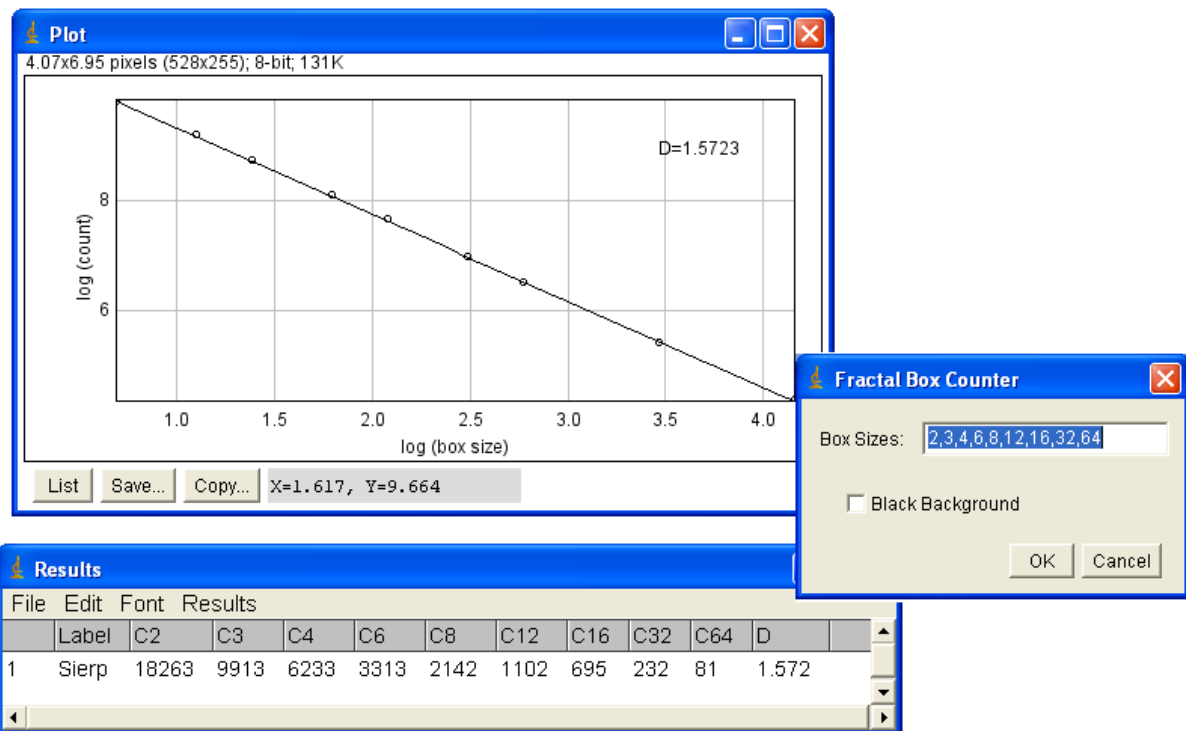
$$\ln p_n \approx -D \cdot \ln v_n + \ln M \quad (7.3.5)$$

V tomto vztahu známe posloupnost  $v_n$  velikostí čtverců sítě, kterou volíme, a počty  $p_n$  pokrývajících čtverců v jednotlivých sítích. Přibližný vztah (7.3.4) tedy určuje rovnici přímky – ke známému tvaru přímky stačí označit

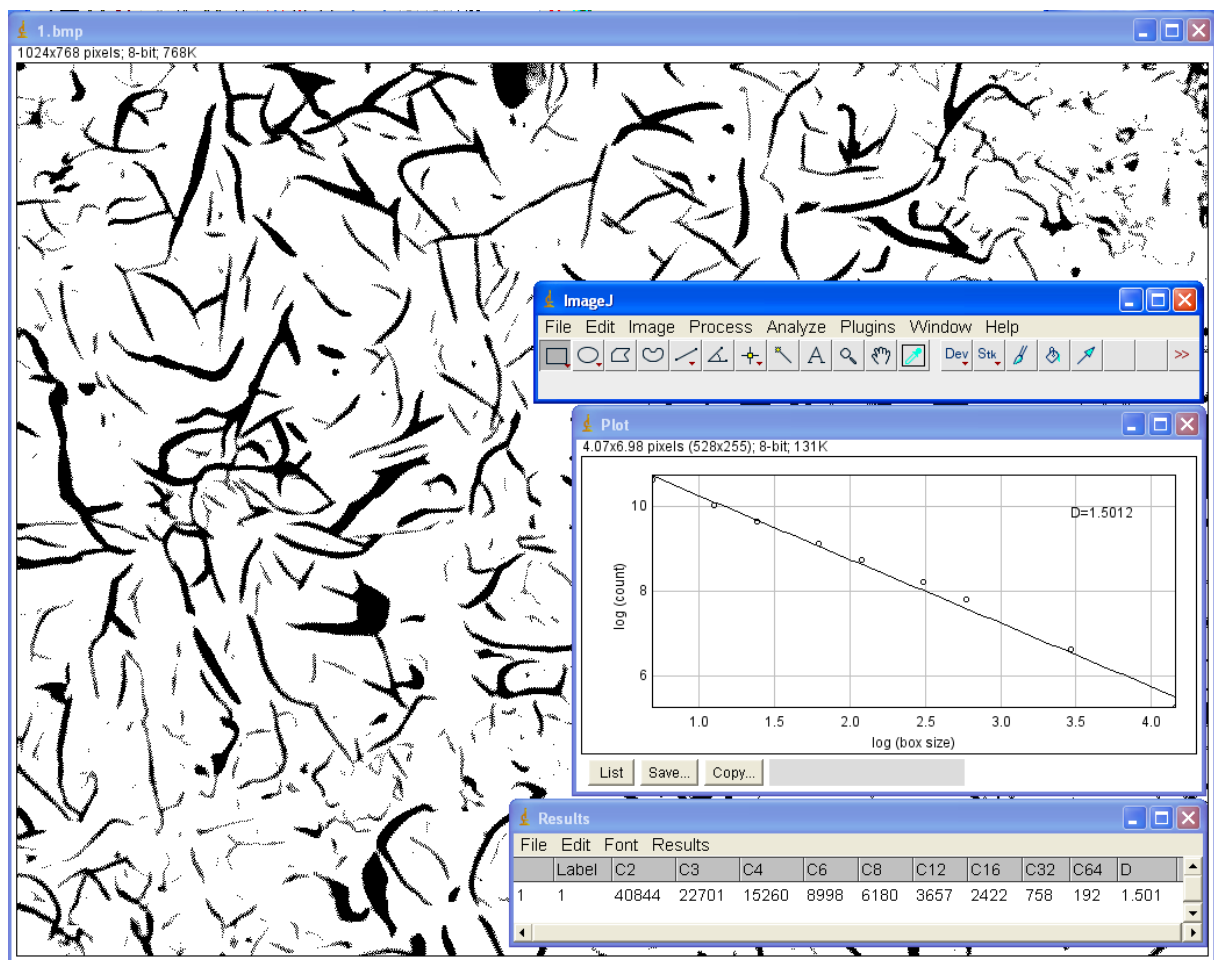
$$\underbrace{\ln p_n}_{y_n} \approx \underbrace{-D}_{k} \cdot \underbrace{\ln v_n}_{x_n} + \underbrace{\ln M}_q$$

Znamé body  $[x_n; y_n]$  vyneseme do grafu a určíme rovnici přímky, která jimi prochází. Body neleží přesně na přímce, přímku je třeba proložit tak, aby se od naměřených hodnot co nejméně lišila. To se provádí metodou nejmenších čtverců, se kterou se seznámíte v matematice. Na obr 7.3.5. máme měření fraktální dimenze Sierpiňského trojúhelníka vygenerovaného na obraze s rozlišením 1024×1024 tak, jak ji provedl program ImageJ - porovnejte výsledek se (7.3.3)

Fraktální dimenze udává „členitost“ měřených objektů, která může být v mnoha aplikacích velmi důležitá. Na obr. 7.3.6. máme výsledky měření fraktální dimenze uhlíkových částic v oceli.



**Obr. 7.3.5:** Výsledky měření Sierpiňského trojúhelníka (program ImageJ)



**Obr. 7.3.6:** Výsledky měření fraktální dimenze uhlíkových struktur v litině (program ImageJ)