

Kompresa obrázku za pomoci singulárního rozkladu

Adolf Středa

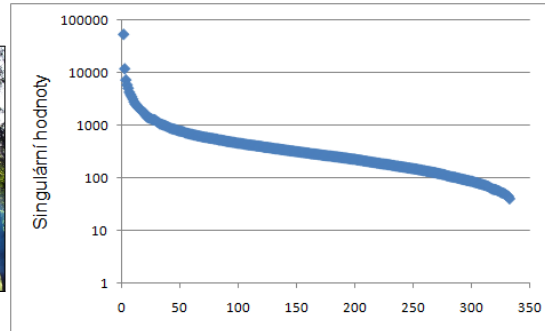
Tím jak se množství obrázků v našich zařízeních zvětšuje, stává se komprese obrázků stále častějším problémem. S nástupem internetových služeb se sdílení fotografií, jako je například Facebook nebo Google+, se problém stal ještě palčivějším. Velmi často se problém oddaluje zmenšením rozměrů nebo rozlišení obrázku, ovšem samotné zmenšování mnohdy neoddlá problém dostatečně. Singulární rozklad matice je často skloňovaným termínem v praktických aplikacích - od metody nejmenších čtverců ve statistice po rekonstrukci obrazu a jedním z dalších možných využití singulárního rozkladu matice je právě komprese obrázku. Jde o ztrátovou kompresní techniku, kdy lze dosáhnout významné úspory dat s minimální ztrátou informace u většiny obrázků. Mnohdy je tato ztráta informace mimo rozlišovací schopnosti našeho oka a, narozdíl od textů, je ve výsledku snesitelná.

Každý barevný kanál obrázku (např. RGB kanály - červená, zelená, modrá) lze reprezentovat maticí s rozměry odpovídajícími rozměrům obrázku v pixelech. Standardně se pracuje s barevnou hloubkou 32 bitů (resp. 24 bitů bez alfa kanálu - průhlednost) označovanou taktéž jako truecolor. Tedy pro každý barevný kanál máme přesně jeden byte (8 bitů), což znamená, že tyto matice mohou obsahovat hodnoty mezi 0 a $2^8 - 1$ ($= 255$). Cílem komprese je snížit hodnotu takové matice za cenu co nejmenšího snížení kvality obrázku. Rozměry této matice jsou identické k rozměrům obrázku, ovšem hodnota této matice je daná až obsahem. Efektivita algoritmu je ovlivněna především rozměry obrázku a právě obsahem; barevná hloubka a počet barevných kanálů nemá na efektivitu zásadní vliv. Snížení hodnoty matice může sloužit jako předzpracování/zakončení pro jiné algoritmy (např. příprava pro skeletní rozklad nebo další stupeň komprese po JPEG kompresi) nebo může být singulární rozklad použit i pro vytvoření vlastního komprimovaného obrázkového formátu.

Singulární rozklad matice je rozklad matice na součin 2 ortogonálních (obecně unitárních) matic U , V^T a diagonální Σ ve vztahu $A = U\Sigma V^T$. Nazveme-li matici reprezentující jeden barevný kanál obrázku maticí A , potom Σ je diagonální maticí s druhými odmocninami nenulových vlastních čísel, tzv. singulárními čísly σ_i , matice $A^T A$ seřazenými sestupně dle hodnoty na hlavní diagonále. Počet vlastních čísel odpovídá hodnotě matice A . Sloupcové vektory matice V tvoří normalizované vlastní vektory $A^T A$, v pořadí odpovídajícímu



(a) Testovací obrázek č. 1



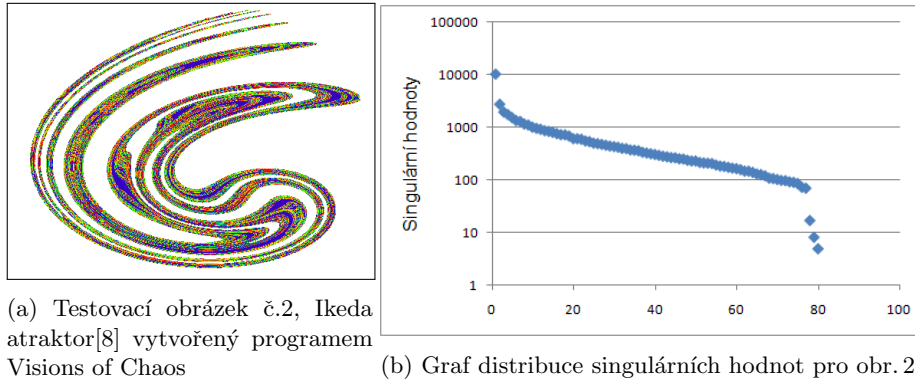
(b) Graf distribuce singulárních hodnot pro obr. 1a

Obrázek 1: Příklad dobře komprimovatelného obrázku

vlastním číslům z nichž byla singulární čísla odvozena. Sloupcové vektory matice U dostaneme ze sloupcových vektorů V a singulárních čísel za pomoci vztahu $u_i = \sigma_i^{-1}Av_i$. Alternativně by bylo možné vynechat singulární hodnotu a výsledný vektor dále normalizovat. Rozklad lze zapsat dvěma ekvivalentními způsoby - jeden, kdy rozměry matice A odpovídají rozměrům matice Σ a je nutné doplnit sloupcové vektory matice U na ortonormální bázi. Ovšem, jelikož jsou prvky pod hlavní diagonálou nulové, pak je možné redukovat Σ na matici čtvercovou a zároveň si tím ušetříme problém s posledními sloupcovými vektory matice U (aby bylo možné násobit U a Σ musí být počet doplněných sloupcových vektorů roven počtu nulových řádků Σ) - takto dostaneme druhou možnost zápisu rozkladu. Proces komprese využívá dyadické expanze matice A , kdy matici A rozložíme na sumu $\sum_{i=1}^n u_i\sigma_i v_i^T$, tedy součet n (počet singulárních čísel) matic s hodnotami 1. Příklad singulárního rozkladu na matici 2×3 vč. dyadické expanze:

$$\begin{aligned} \begin{pmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{pmatrix} &= \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} 5 & 0 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{18}} & -\frac{1}{\sqrt{18}} & \frac{4}{\sqrt{18}} \end{pmatrix} \\ &= \begin{pmatrix} \frac{5}{2} & \frac{5}{2} & 0 \\ \frac{5}{2} & \frac{5}{2} & 0 \end{pmatrix} + \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} & 2 \\ -\frac{1}{2} & \frac{1}{2} & -2 \end{pmatrix} \end{aligned}$$

Z toho na první pohled vyplývá, že čím menší singulární číslo (relativně k ostatním), tím menší má možnost výsledný obrázek ovlivnit. Koneckonců z této vlastnosti vychází definice spektrální normy matice, která je definovaná jako největší singulární číslo dané matice, jejíž přímý důsledek je, že $\|A\| \geq \|Ax\|$ pro libovolný vektor x odpovídající dimenze. Komprese se provede právě odebráním matic s nejmenšími singulárními čísly z dyadického rozkladu. Proto například grafy, kde se velká část sloupců matice liší v drobnostech (např. popisky), jsou komprimovatelné až na 10% velikosti s minimem kompresních artefaktů. Naopak u jiných obrázků, které kladou důraz na detail anebo jsou velmi schematické

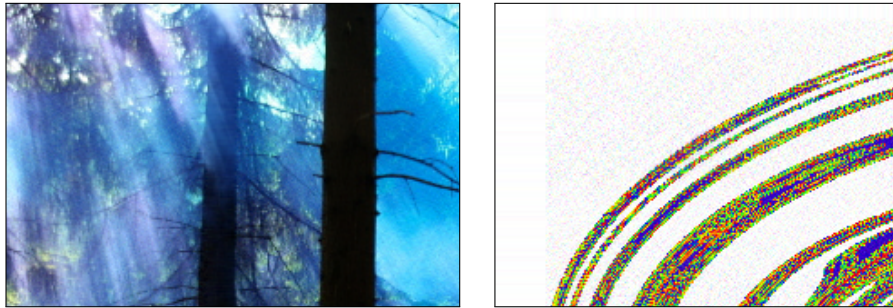


Obrázek 2: Příklad špatně komprimovatelného obrázku

(např. skica), se projeví některé artefakty dříve - třeba už při 50%.

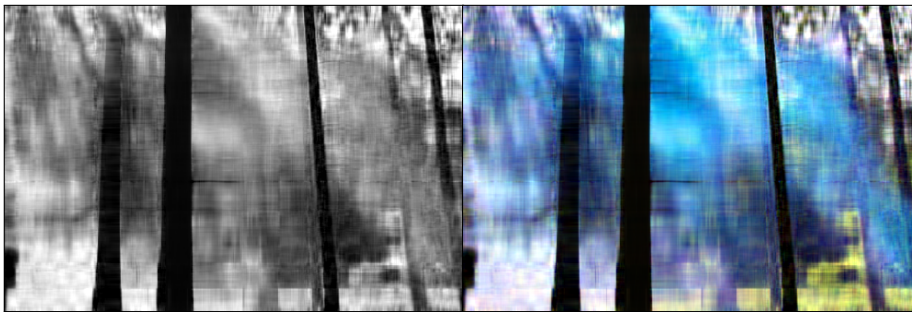
Komprimovatelnost obrázku záleží na rozložení singulárních čísel jednotlivých barevných kanálů - jsou-li singulární čísla řádově přibližně stejná, pak jsou možnosti komprese omezené a redukcí malého počtu singulárních čísel odstraníme pravděpodobně příliš dat, což se na obrázku projeví znatelnými kompresními artefakty. Naopak, je-li např. růst singulárních čísel exponenciální, pak je možné s minimálními artefakty použít vysoký poměr komprese. Mezi nejčastější kompresní artefakty patří změna světlosti obrázku, zrnitost a rozmazání (především vertikální a horizontální). Rozdíl v distribuci singulárních čísel obr. 1a a 2a je patrný na obr. 1b a 2b. V obou grafech je pro názornost použito logaritmické měřítko. V prvním případě je vidět, že pokles velikosti singulárních čísel přibližně od 80. singulárního čísla začíná být stabilní až přibližně k 260. singulárnímu číslu, kdy opět pokles nabírá na rychlosti. V případě druhém je počáteční pokles výrazně slabší a přibližně od 15. singulárního čísla je téměř stabilní, až na poslední tři singulární hodnoty (bohužel necelá 4% úspora dat za tato singulární čísla je takřka bezvýznamná). Přímé porovnání komprimovatelnosti je vidět na výřezech z levého horního rohu již zmiňovaných obrázků na obr. 3. Po odebrání poloviny matic z dyadické expanze nejsou kompresní artefakty na obr. 3a na první pohled viditelné. Ovšem na obr. 3b je již při takovéto kompresi vidět šum v pomyslném obdélníku obsahujícím náš atraktor. Mimo pomyslný obdélník najdeme vertikální (resp. horizontální) pruhy za jeho horizontálními (resp. vertikálními) hranami, nicméně tento kompresní artefakt je už na hraně viditelnosti.

Rozdíl v kvalitě mezi po kompresi barevného a černobílého obrázku, při použití koeficientů 0.3, 0.59, 0.11 pro převod RGB kanálů do kanálu jediného a algoritmu bez detekce hran, není ve většině případů znatelný (viz obr. 4 a 5, 6). Důvodem pro tyto poměry je odlišná intenzita vnímání jednotlivých barev, kdy zelená je vnímána nejjasněji a naopak modrá s nejmenší intenzitou.[5] Při použití



(a) čtvrtinový výřez obr. 1a s kompresí 50% (b) čtvrtinový výřez obr. 2a s kompresí 50%

Obrázek 3: Porovnání komprimovatelnosti obr. 1a a obr. 2a



Obrázek 4: Srovnání kvality po kompresi barevného a černobílého obrázku oproti zdrojovému obrázku (obr 1a). Při kompresi bylo zachováno 5% (16) singulárních hodnot.

poměrů, které minimálně zkrusují světlost objektů na obrázku, se ve většině případů černobílého komprimovaného obrázku projevuje dokonce i případná variabilita v distribuci singulárních čísel u jednotlivých barevných kanálů barevné předlohy. Výjimkou jsou hraniční případy daného převodního algoritmu, kdy dané barvy splynou a může dojít ke „korekci“ případných zdrojů artefaktů při převodu anebo ke změně charakteru těchto artefaktů. Počet barevných kanálů ovlivňuje především nutný počet singulárních rozkladů (pro každý kanál jeden). Pouze je-li rozdílná distribuce vlastních čísel jednotlivých kanálů, pak ovlivní i výsledný koeficient velikosti souboru, který bude průměrem koeficientů jednotlivých barevných kanálů. Barevná hloubka ovlivňuje pouze paměťové požadavky na jeden prvek matice, na koeficient velikosti souboru nemá vliv žádný.

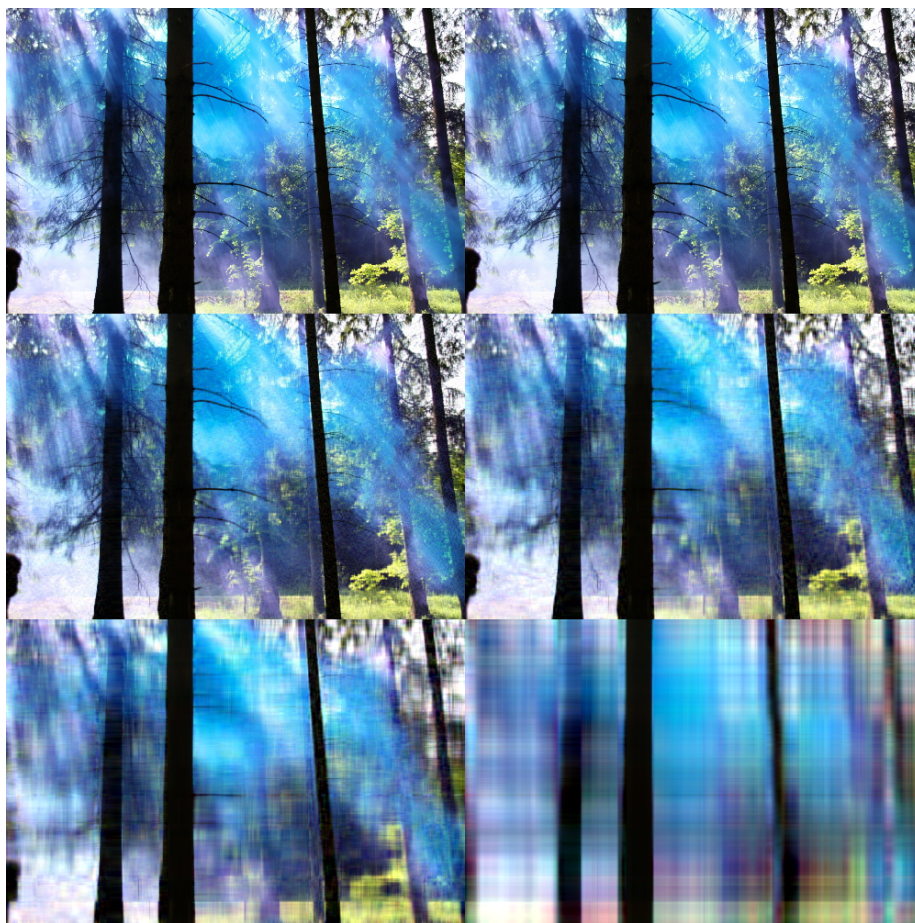
Nechť máme černobílý obrázek o rozměru $m \times n$ s jediným 8bit barevným kanálem. Obrázek můžeme do paměti uložit jako matici $m \times n$ zabírající $m \cdot n$ bytů. Je-li r_1 hodnota této matice, pak můžeme uložit matici za pomoci skeletního rozkladu do $r_1 \cdot (m + n)$ bytů. Potom po kompresi, při které ponecháme r_2 členů dyadické expanze, můžeme uložit matici do $r_2 \cdot (m + n + 1)$ bytů, nebo do

konce díky skeletnímu rozkladu do $r_2 \cdot (m+n)$ bytů. Potom ξ koeficient velikosti souboru oproti originálu je:

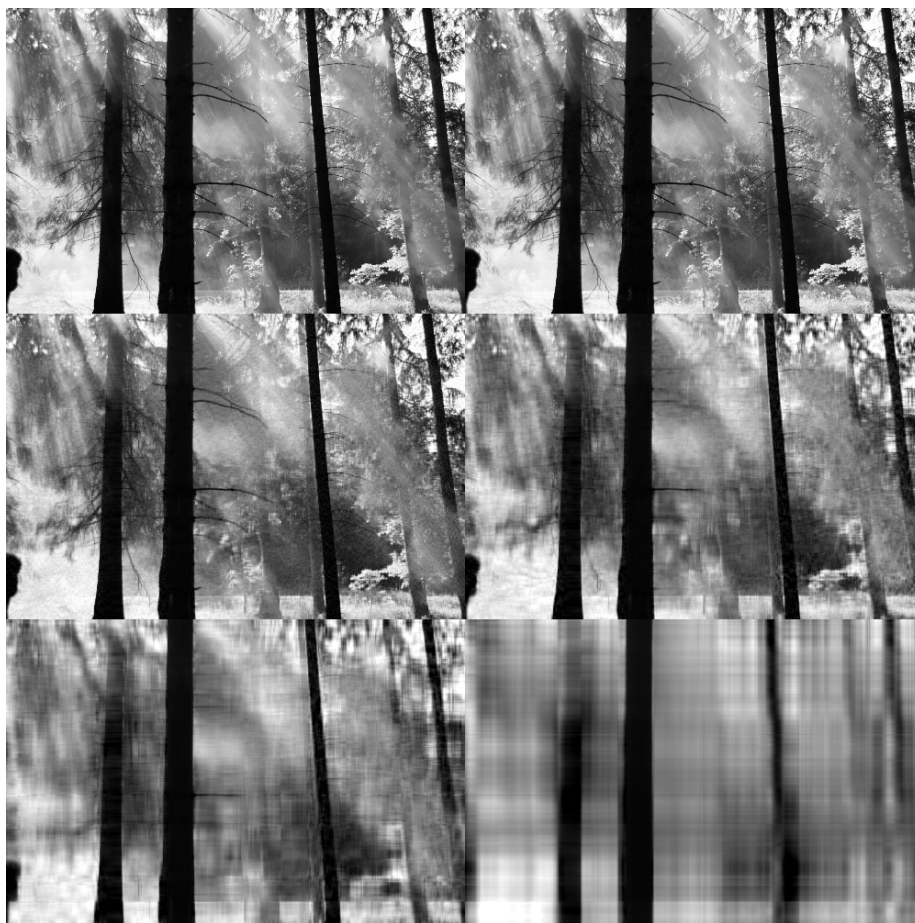
$$\xi = \frac{(m+n) \cdot r_2}{(m+n) \cdot r_1} = \frac{r_2}{r_1}$$

Velikost obrázku v paměti ve formě skeletního rozkladu bude $\xi \cdot r_1 \cdot (m+n)$ bytů, a tedy změna velikosti souboru je přímo úměrná změně počtu singulárních čísel (hodnoty matice). Samozřejmě tyto hodnoty platí pouze pokud data ukládáme ve formě skeletního rozkladu. Vezmeme-li v potaz tradiční datové formáty obrázku, pak poměr velikosti komprimovaného se může lišit oběma směry kvůli kompresi zahrnuté ve formátu (např. JPEG, PNG) anebo velikost souboru může být neměnná v případě formátu bez komprese (např. BMP). [3] Pokud bychom chtěli komprimovat pouze za pomoci singulárního rozkladu, poté je nutné aby $r \cdot (m+n+1) \leq m \cdot n$ (r počet singulárních čísel), ovšem i skeletní rozklad má o r lepší paměťovou náročnost. Tedy, až na výjimky, by bylo samostatné využití singulárního rozkladu pro ukládání dat diskutabilní.

Technicky čelí komprese přes singulární rozklad jednomu zásadnímu problému. Singulární rozklad je velice všestranný nástroj a je možné z něj mnoho vlastností dané matice odvodit. Ovšem tyto, pro kompresi redundantní, informace jsou naneštěstí za cenu velké výpočetní náročnosti. Navíc postup běžný při výpočtu na papíru přes charakteristický polynom je v takto velkých maticích nemyslitelný. Nejrychlejší algoritmy pro hledání vlastních čísel jsou algoritmy iterativní. Mezi ně patří např. QR transformace - jedna z nejpoužívanějších skupin algoritmů. Velmi častou implementací je převod matice pomocí rotací do bidiagonálního tvaru a následná diagonalizace za pomoci QR algoritmu. Iterativní metody nedají přesný výsledek, pouze k němu s každou iterací konvergují a výsledek jsou schopny dodat s deklarovanou přesností (v závislosti na nastavení použité knihovny nebo vlastní implementace). Bohužel problém s náročností při větších obrázcích je stále u většiny knihoven velmi znatelný. Částečně tento problém řeší implementace využívající paralelizace (rozděl a panuj), kdy je SVD rozklad rozdělen na několik menších úkolů, což je obzvláště výhodné u velkých matic. [2] Navíc to umožňuje efektivně využít i mnohem specializovanější zdroje. Například grafické karty se velmi často využívají při výpočtech umožňující masivní paralelizaci a rychlost se může u velkých matic lišit 6-7x [6], v závislosti na použitém algoritmu, jeho implementaci a použitých knihovnách pro výpočty na grafické kartě.



Obrázek 5: Různé poměry komprese barevného obrázku. Procenta zachovaných singulárních hodnot jsou po řádcích: 100% (původní), 75%, 25%, 10%, 5% a 1%.



Obrázek 6: Různé poměry komprese černobílého obrázku. Procenta zachovaných singulárních hodnot jsou po řádcích: 100% (původní), 75%, 25%, 10%, 5% a 1%.

Reference

- [1] *Math.net numerics documentation*. <http://numerics.mathdotnet.com/docs/>. Navštíveno: 22. 8. 2014.
- [2] *Singular value decomposition*. <http://www.alglib.net/matrixops/general/svd.php>. Navštíveno: 22. 8. 2014.
- [3] P. AGUILERA, *Comparison of different image compression formats*. [http : //homepages.cae.wisc.edu/ ece533/project/f06/aguilera_rpt.pdf](http://homepages.cae.wisc.edu/ece533/project/f06/aguilera_rpt.pdf). Navštíveno: 22. 8. 2014.
- [4] L. CAO, *Singular value decomposition applied to digital image processing*. <http://www.lokminglui.com/CaoSVDintro.pdf>. Navštíveno: 22. 8. 2014.
- [5] D. FINLEY, *Hsp color model — alternative to hsv (hsb) and hsl*. <http://alienryderflex.com/hsp.html>. Navštíveno: 22. 8. 2014.
- [6] B. FOSTER, S. MAHADEVAN, AND R. WANG, *A gpu-based approximate svd algorithm*, in *Parallel Processing and Applied Mathematics*, Springer, 2012, pp. 569–578.
- [7] A. KOZLÍK, *Ztrátová komprese obrázku pomocí singulárního rozkladu*. <http://www.karlin.mff.cuni.cz/kozlik/linalg2/svd.php>. Navštíveno: 22. 8. 2014.
- [8] J. RAMPE, *Visions of chaos, version 52.7 [software]*. Dostupné na: <http://softology.com.au/voc.htm>.
- [9] T. VUTHY, *Using the singular value decomposition (svd) for image compression*. <http://volunteerlecturerprogram.com/wp-content/uploads/2013/01/vuthythesis.pdf>. Navštíveno: 22. 8. 2014.