

KIV/PRO

Algoritmy pro analýzu a vizualizaci
snímků s vysokým dynamickým
rozsahem na základě lidského vnímání

Zadání

Najděte v anglicky psané odborné literatuře článek v délce alespoň 5 stránek o nějakém algoritmu řešícím libovolný problém. Algoritmus popište do českého referátu tak, aby ho podle vašeho názoru pochopil běžný student 2. ročníku informatiky. Váš text musí svědčit o tom, že algoritmu rozumíte, a musí ho z něj pochopit i nezasvěcený čtenář.

Definice

Dynamický rozsah obrazu označuje poměr mezi nejjasnějšími a nejtmavšími částmi scény. Existují dva typy. Obraz s nízkým dynamickým rozsahem LDR, jenž má rozsah 256 hodnot, což může vést ke ztrátě informací při zobrazování.

Obrazy s vysokým rozsahem HDR mají rozsah hodnot velmi vysoký, díky tomu se pro fotorealistické účely. Většina zařízení pro zobrazení a vytvoření obrazů, tedy kamery a displeje, mají rozsah hodnot scény omezený a nemohou s HDR obrazy přímo pracovat. Je tedy třeba použít tzv. mapování tónů, tedy transformaci stlačující rozsah intenzit scény. V práci [1] je popsáno vylepšení globálně používané verze algoritmu z [2]. Vylepšení vychází z modelu lidského vidění. V této práci je obsažen výtah výše zmíněné práce [1].

Analýza problému

Popis problému

Problém je vyvinout algoritmus mapování tónů, jenž uchovává důležité informace z pohledu lidského vnímání. Autoři [1] chtěli získat transformaci:

$$O = TM(I), \|O - I\|_{human} \rightarrow 0,$$

$I = \{I(p, i) \in [0, M], M < \infty, p \in P\}$ - je HDR obraz

$O = \{O(p, i) \in \{0, 1, \dots, N = 256\}, p \in P\}$ - LDR obraz

p = sada pixelů obrazu, $i = R, G, B$ barevné složky obrazu

O mapování tónů

Algoritmy mapování tónů se rozdělují na globální a místní. Globální algoritmy transformují intenzitu všech pixelů jednotným způsobem společnou kompresní křivkou. Tento postup je jednoduchý a velmi rychlý, ale při širokém rozsahu vstupních intenzit je úroveň detailů nedostačující. Pro zlepšení výsledků se mohou využívat lokální operátory, jenž modifikují intenzitu pixelu pomocí hodnot jeho sousedů, ale tento způsob vede k vytváření artefaktů, které mohou způsobené vylepšení vyrušit.



Obrázek 1 [1] Výsledek vytvořený pomocí různých parametrů základního algoritmu

Pro zachování výhod lokálních operátorů lze použít metodu popsanou v [2]. Tato lokální metoda nevyžaduje rozdělení vstupní obraz, což zabraňuje výskytu artefaktů. Zároveň se mapa záření zpracovává přímo. Globální optimalizace navíc zaručuje lokální konzistenci každého pixelu obrazu.

Jednou z nevýhod těchto algoritmů je, že parametry jsou unikátní pro každý obraz, což má rozhodující efekt, jak je vidět na obrázku 1.

Popis základního algoritmu

Vstupem je HDR obraz a výstupem LDR obraz, oba ve formátu RGB. Kroky algoritmu jsou:

1. Konstrukce mapy záření Q (radiance map) pomocí heuristického algoritmu používající prostor HSV a snižuje celkovou střední intenzitu na pevnou hodnotu. Každý pixel mapy záření obsahuje informaci o jeho intenzitě.
2. Vygenerování naváděcí mapy C dané mapou záření Q . Právě mapa C má největší dopad na kvalitu výsledku algoritmu mapování tónů. Autoři [5] využili vzorec:

$$C(p) = (\mu(p)^{\beta_1} \sigma(p)^{\beta_2} Q(p)^{\beta_3} + \kappa)^{-1},$$

kde p je pixel obrazu, Q je mapa záření (radiance map), C je naváděcí mapa (guidance map), μ je místní průměr (v okně 3×3) a σ je místní standardní odchylka (v okně 3×3) a $\beta_1, \beta_2, \beta_3$ jsou parametry algoritmu.

3. Celý obraz je rozložen na překrývající se 3×3 okna. Takže při definování lineární transformace intenzity pixelu v každém okně, patří všechny pixely do 25 oken. Pro určení parametrů transformace se navrhuje minimalizovat každý bod funkcí:

$$f(q) = \sum_{p \in w(q)} (I(p) - u(p)Q - v)^2 + \varepsilon C^{-2}(u(p) - v(p))^2,$$

kde q je zkoumaný pixel, $w(q)$ je množina oken obsahujících q , p je pixel obsažen v okně, I je výstupem kroku (záření nízkého rozsahu), Q je mapa záření, C je mapa navádění, u a v jsou parametry lineární transformace a ε je regularizační koeficient použitý k vyloučení triviálního řešení.

Funkčnost je minimalizována řešením systému lineární algebraické rovnice. Tato rovnice ukazuje jak moc lze mapu záření Q komprimovat. Cílem je stlačit ostré změny v intenzitě obrazu při zachování detailů, což je zajištěno konstrukcí mapy popsané v kroku 2.

4. Transformace originálního HDR obrazu v RGB formátu Schlickovým vzorcem:

$$Output(p, i) = \left(\frac{Input(p, i)}{Q} \right)^s I$$

kde p je obrazový pixel, i je index barevného kanálu, I je záření nízkého rozsahu, Q je mapa záření, $Input$ jsou vstupní HDR data, $Output$ jsou výstupní LDR data a s je stupeň nasycení (specifikované uživatelem).

Dále je výstupní obraz normalizován pomocí histogramu oříznutím horních a dolních 0,5 % a poté odebrán vzorek do 256 zásobníků intenzity.

Úpravy základního algoritmu

Pro vylepšení algoritmu autoři [1] nahrazují kroky 1, 2 a 4. Kroky 1 a 4 se vzájemně doplňují a jsou určeny vztahem mezi hodnotami barvy a intenzity ve zvoleném barevném modelu. Modifikace kroku 2 navržená v [1] je klíčová, protože umožňuje přesunout zaměření naváděcí mapy z lokální ostroty (zajištěné statistickými veličinami uvnitř malého okna) na množství informací souvisejících s lidským vnímáním.

Barevný prostor

V původním algoritmu z [2] se používá barevný prostor HSV. Ten ovšem neodráží přizpůsobivost vnímání a další důležité jevy lidského zrakového systému. Z tohoto důvodu autoři [1] zvolili model CIECAM02. Jedná se o standardní model přijatý Mezinárodní komisí pro osvětlení (CIE). Tento model není nejpřesnější, ale není přísný na požadavky na vstupní data a přináší solidní výsledky.

Dobrý bodem aplikace tohoto modelu pro mapování tónů je, že barevné složky musí být správně rekonstruované po transformaci intenzity v kroku 4 základního algoritmu. V původním algoritmu byl k tomuto účelu použit Schlickův vzorec, jenž je pro zvolený barevný model neplatný. Pokud jsou všechny kompletní parametry přesunuty do souhrnných koeficientů, pak se v modelu CIECAM02 vypočítá podle vzorce:

$Q = k * (aL + bM + cS)^\alpha$, kde L, M, S jsou výsledky komprese odezvy kužele.

Při předpokladu, že intenzita byla transformovaná na jas Q' . Nové hodnoty pro kužely lze vypočítat podle vzorců:

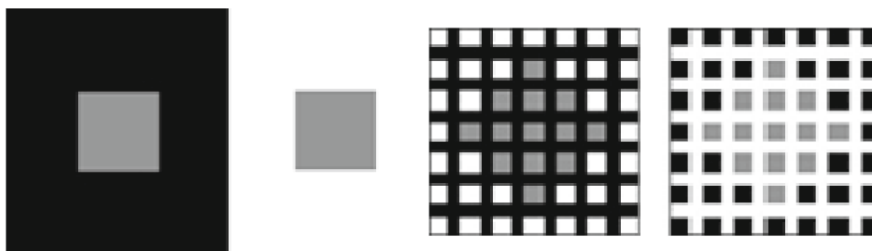
$$L' = L * \frac{I_{\alpha}^{\frac{1}{\alpha}}}{Q}, \quad M' = M * \frac{I_{\alpha}^{\frac{1}{\alpha}}}{Q}, \quad S' = S * \frac{I_{\alpha}^{\frac{1}{\alpha}}}{Q}$$

Zde je koeficient α určen na základě parametrů modelu a uživatel jej nemusí specifikovat.

Pro zabránění chybám způsobeným příliš velkými hodnotami při použití nelineárního koeficientu se vypočítané odezvy kuželů dodatečně odříznutím 0,5 % v horní a dolní úrovni intenzity na obou obrazcích vyrovnají vzhledem k původním reakcím vyrovnáním intenzit na základě histogramu. Dále musí být barevné komponenty LMS transformovány zpět na RGB složky. Poté se provede normalizace histogramem, vzorkováním a gama korekcí.

Výpočet vnímané intenzity

Ke spočítání vnímaného jasu se použije model BiWaM (Brightness Induction Wavelet Model) [14], jenž popisuje efekty kontrastu jasu na základě známých vlastností lidského zrakového systému a lze být použit jako inverzní naváděcí mapa C . Model BiWaM navíc dokáže reprodukovat mnoho iluzí vnímání jasu, jež jsou vidět na obrázku 2. V příkladu vlevo vypadá šedé pole jasněji na tmavším pozadí a méně světlé na světlejším pozadí. V příkladu vpravo je šedý kosočtverec v kontrastu spíše s pozadím než s přilehlou mřížkou.



Obrázek 2 [1] Příklady kontrastních iluzí

Model má dva parametry – základní kontrastní faktor a cílové měřítko odráží podmínky, za kterých je obraz zobrazen. Režim kombinuje jak s kontrastem nastavením vztahů mezi nimi na základě funkce kontrastní citlivosti nezávisle na uživateli. Model používá dekompozici obrazu na sadu frekvenčních složek.

$$C = \max\left(\varepsilon, BIWaM\left(\frac{Q}{\max(Q)}\right)\right)^{-1}, \text{ kde } Q \text{ je mapa záření a } \varepsilon = 0.03 \text{ je malý koeficient.}$$

Algoritmus BiWaM funguje následovně.

1. Vstupní obraz Q Se rozloží na sadu frekvenčních složek $3 * N + 1$, kde $N = \log_2(\max(W, H))$ je maximální dostupná úroveň podrobností (W a H jsou šířka a výška obrázku Q). Všechny úrovně s výjimkou poslední obsahují tři různě orientované komponenty, které odrážejí/reflektují kontrast obrazu. Jediná komponenta na poslední úrovni odráží celkový jas obrazu.

2. Na každé úrovni detailu (kromě poslední) jsou koeficienty pro orientaci každého pixelu opraveny pomocí komponenty a dvou normalizačních křivek:

$w_{out}(s, o, p) = w_{tn}(s, o, p) * (CSF_1(s) * Z_{s,o}(p) + CSF_2(s))$, kde w_{tn}, w_{out} jsou vstupem a výstupem kroku, respektive s a o jsou aktuální měřítko a orientace, p je pixel, $Z_{s,o}(p)$ je kontrast pro aktuální komponentu, $CSF_1(s)$ a $CSF_2(s)$ jsou váhové křivky pro aktuální měřítko.

Pík křivek (který souvisí s pozorovací stupnicí) a dominance křivky 1 nad křivkou 2 lze regulovat. Kontrast pro aktuální komponentu/složku se vypočítá podle vzorce:

$$Z_{s,o}(p) = \frac{\sigma_{cen}^2(p)}{\sigma_{cen}^2(p) + \sigma_{sur}^2(p)}$$

p je pixel, $Z_{s,o}(p)$ je kontrast pro aktuální komponentu, $\sigma_{cen}^2(p)$ je rozptyl v okně 5 x 5 (středová oblast), $\sigma_{sur}^2(p)$ je rozptyl v okně 13 x 13, z něhož je odstraněnou jádro 5 x 5 (okolní oblast)

3. Upravené součásti se shromáždí do výsledného obrázku.

Použití algoritmu BIWam umožňuje zachovat na obrázku detaily, které jsou důležité z hlediska lidského vnímání a odstraňuje nedůležité detaily.

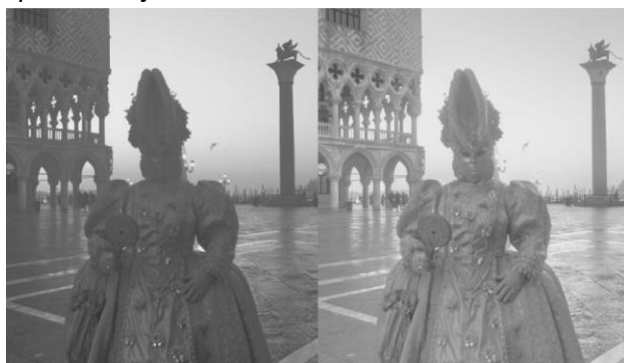
Srovnání s jinými algoritmy



Obrázek 6 [1] Příklad výsledků získaných pomocí různých algoritmů: Od shora dolů – nejlepší klasický, základní, navržený v článku [1], 2011 a 2015

K prokázání výsledků upraveného algoritmu, provedli autoři [1] jeho srovnání hlavně se základním algoritmem [2] a i s algoritmem Ferradase z roku 2011 popsaným v [3] a s algoritmem z roku 2015, jenž je též založen na modelu lidského vnímání. Toto srovnání zobrazuje obrázek 5.

Další obrázky se týkají bližšího srovnání se základním algoritmem z [2]. Na obrázku 3 je vidět, že úroveň detailů se výrazně zlepšila v oblastech masky a šatů. Okno z barevného skla na obrázek 4 si zachovalo jak vysokou úroveň detailů, tak jas vzhledem k okolním objektům. Nakonec obrázek 6 ukazuje výsledek komplikovaného obrázku diskutovaného na začátku článku. Je vidět, že kromě absence zkreslení je barva obrazu přirozenější.



Obrázek 3 [1] Příklad výsledků původních (vlevo) a upravených (vpravo) algoritmů.

Obrázek 4 [1] Příklad výsledků původních (vlevo) a upravených (vpravo) algoritmů.



Obrázek 5 [1] Příklad výsledků původních a upravených algoritmů

Závěr

Autoři [1] navrhli modifikaci algoritmu mapování tónů na základě globálně optimalizovaného algoritmu lineárního mapování tónu v okně. Přesněji se navrhuje:

- použít barevný model CIECAM02 pro výpočet mapy intenzity,
- použít model BIWaM pro výpočet naváděcí mapy C

Použití modelu vnímání jasu umožňuje opustit složitý postup ladění parametrů pro každý obrázek. Upravený algoritmus ukazuje vylepšení základního algoritmu [2] z hlediska počtu a intenzity artefaktů. Kvalita upraveného algoritmu je srovnatelná s nejlepšími nejmodernějšími algoritmy.

Bibliografie:

1. Zipa, K.S., Ignatenko, A.V. Algorithms for the analysis and visualization of high dynamic range images based on human perception. *Program Comput Soft* **42**, 367–374 (2016). <https://doi.org/10.1134/S0361768816060086>
2. Shan, Q., Jia, J., and Brown M.S., Globally optimized linear windowed tone mapping, *IEEE Trans. Visualization Comput. Graphics*, 2010, vol. 16, no. 4.
3. Ferradans, S., Bertalmio, M., Provenzi, E., and Caselles, V., An analysis of visual adaptation and contrast perception for tone mapping, *IEEE Trans. Pattern Anal. Machine Intell.*, 2011, vol. 33, no. 10, pp. 2002–2012.