



PROJEKT Č. VI20172020068

NÁSTROJE A METODY ZPRACOVÁNÍ VIDEO A OBRAZU
PRO ZVÝŠENÍ EFEKTIVITY OPERACÍ BEZPEČNOSTNÍCH A
ZÁCHRANNÝCH SLOŽEK (VRASSEO)

ZPRACOVÁNÍ HDR VIDEO DYNAMICKÝCH SCÉN

TECHNICKÁ ZPRÁVA

Svetozár Nosko, Pavel Zemčík

Vysoké Učení Technické v Brně
Fakulta Informačních Technologií
Božetěchova 1
612 66 Brno, Česká republika

Prosinec 2019

Obsah

1	Úvod	1
2	Analýza HDR pro UAV	2
2.1	Analýza detekce v HDR	2
2.2	Zaznám kompletního HDR na dronu	4
3	Závěr	6

Abstrakt

Technická zpráva popisuje vývoj technologie pro záznam a zpracování HDR (vysoký dynamický rozsah) videa pro drony, jejichž počítačové systémy nedisponují vysokým výpočetním výkonem, musí být lehké a mít malou spotřebu energie.

1 Úvod

HDR je technika používaná pro reprodukování většího dynamického rozsahu než je typicky možné standardními metodami pro záznam obrazu. Cílem je zobrazit rozsah jasů podobně jako je vnímán lidským okem. HDR bývá použito pro analýzu dat (detekce objektů a pod.) v těžkých světelných podmínkách [1]. Fotografie s vysokým dynamickým rozsahem se typicky vytvářejí snímáním scény pomocí více snímků s různou expozicí (tato technika se nazývá expoziční bracketing) a následně jejich složením vzniká HDR. Tenhle postup je ovšem vhodný jen pro statické kamery a statické scény. Problém skládání snímků ze statické kamery pro pohybující se objekty byl řešen v předcházejícím reportu (HDR ve vestavěných zařízeních), kde bylo prezentováno řešení skládání HDR snímků pro pohybující se automobily. Dalším krokem je analýza možností implementace a integrace kamery schopné pořádit dostatečně kvalitní snímky pro skládání HDR v omezených výpočetních a fyzických podmínkách (váha rozměry) na dronu. Dnes se běžně využívají malé drony (patří mezi UAV - Unmanned Aerial Vehicle - Bezpilotní letadlo), které jsou dálkově ovládané a nacházejí využití ve všech oblastech lidské činnosti. Typicky se používají pro monitorování infrastruktury, ropovodů, dálnic nebo elektrického vedení. Ve filmovém a hudebním průmyslu se pomocí dronů natáčejí záběry ze vzduchu. Integrace kvalitního algoritmu pro tvorbu HDR videa přímo na dron umožní aplikaci pokročilých metod počítačového vidění v těžkých světelných podmínkách.

2 Analýza HDR pro UAV

Přesnost detekce je silně závislá na kvalitě zachycených snímků. Standardní kamery nemohou zachytit dynamický rozsah ve scéně tak, jak to umožňuje lidské vidění. Řešením je snímání jasu ve scéně v režimu HDR. Existují dva hlavní přístupy jak pořídit HDR obraz. První přístup předpokládá existenci a využití specializovaných HDR senzorů (např. Zhao et al. [2]), tyto senzory jsou však obecně drahé a technologicky náročné. Druhý přístup je založen na standardních senzorech a na pořízení obrazové sekvence s různou dobou expozice (Debevec et al. [3], Mertens et al. [4]). Detekce objektů ve videu je intenzivní oblast současného vývoje, avšak detekce objektu v HDR nebo snímcích po HDR tone-mappingu je téma stále otevřené.

2.1 Analýza detekce v HDR

Detekce pohybujících se objektů ve video sekvencích zachycených pohybující se kamerou je spojená s mnoha problémy. Yazdi a Bouwmans [5] shrnuli ve své práci hlavní problémy. V této podkapitole budou tyto problémy stručně charakterizovány.

Variance osvětlení sledovaného objektu ve videu se mění v čase. Důvodem je typicky pohyb objektů v scéně nebo pohyb zdrojů světla, odraz světla, zakrytí zdroje světla atd. Důsledkem bývá selhání metod, které nejsou invariantní vůči změně osvětlení.

Dalším problémem je změna tvaru a vzhledu pohybujícího se objektu, protože video zachycuje pouze 2D projekci 3D scény, takže jakákoli rotace může změnit vzhled objektu.

Rychlé změny rychlosti a změny směru pohybu objektu nebo kamery mohou způsobit ztrátu objektu trackerem nebo velkou chybu v odhadu polohy objektu. Dalším problémem při pořizování HDR je vznik ghosting efektu (vysvětleno v reportu z roku 2017).



Obrázek 1: Ukázka rychlého pohybu objektu (Převzato z [5])

Objekt může být také překryt částečně či úplně jinými objekty ve scéně (viz Obrázek 1).



Obrázek 2: Ukázka zakrytí objektu (Převzato z [5])

Problémem pro detektory bývá pozadí ve scénách s komplexní strukturou, kde se vyskytuje pohyb. Typickým příkladem je hladina vody, stromy atd. Problém je kritický pro detekci z kamery na UAV, kde další výzvou je detekovat objekty ve velké vzdálenosti, kdy může být objekt ve videu reprezentován jen několika desítkami pixelů.



Obrázek 3: Ukázka komplexního pozadí (Převzato z [5])

Mnoho algoritmů nedokáže odlišit objekt od jeho stínu a detekuje jej jako součást objektu, bounding box tak obsahuje jak objekt, tak jeho stín. Tento problém se vyskytuje zejména u videosekvencí zachycených z výšky.

Samotné kamery využívají různé senzory, optiky, rozlišení, rychlost snímání a algoritmy pro zpracování obrazu. Nekvalitní video může znemožnit detekci objektů. Dalšími problémy jsou artefakty vznikající pohybem kamery (např při letu UAV), blokové artefakty vznikající při kompresi videa a v neposlední řadě šum senzoru.

Detekce objektů ve videu zachyceném pohybující se kamerou je složitější než ve videu statickém. Často je potřeba rozpoznat pohyb kamery od pohybu objektů a kompenzovat tím pohyb kamery, toto zarovnání je kritické i pro HDR video. Kamera se potenciálně může pohybovat ve všech třech osách. Standardní PTZ (Pann – Tilt – Zoom) bezpečnostní kamery umožňují jednoduchý pohyb, ale složitější řešení je u kamer připevněných na létající prostředky, kde se kamera může pohybovat libovolně ve všech směrech.

Výsledky z publikace Příbyl et al. [1] potvrzují předpoklad, že nezpracované lineární HDR snímky nejsou vhodné pro současné detektory význačných bodů (z anglického Feature Points detectors nebo FP detektory), protože význačné body se detekují převážně ve světlých částech obrazu. Je to způsobeno návrhem současných FP detektorů, které jako vstup vyžadují obraz zobrazitelný na displeji (8 bitová hloubka) a typicky počítají derivační



Obrázek 4: Kamera Basler daA2500-14uc

práh. Tato vlastnost detektoru je neefektivní v případě lineárního obrazu HDR, protože zaznamenává jas v scéně bez ohledu na to, zda je zobrazitelný (32bitové reálné číslo). Toto zjištění otevírá cestu pro budoucí výzkum algoritmů, které jsou schopny detekovat přímo na snímcích HDR, např. pomocí adaptivních prahů.

Jednou z možností je převést snímky z HDR pomocí tónového mapování (tone mapping operator - TMO) na standardní snímky. Použití globálních TMO se podle článku nedoporučuje, protože dochází k rozsáhlé kompresi lokálního kontrastu a počet detekovaných zájmových bodů je poměrně nízký. Silná globální komprese jasu je problematická pro většinu detektorů. Naproti tomu lokální TMO upravuje lokální kontrast a dochází tak k lokální kompresi. Existují však velké rozdíly mezi jednotlivými lokálními TMO v závislosti na tom, jak dochází k lokální kompresi (a případně také šum).

2.2 Záznam kompletního HDR na dronu

Prozkoumání možností implementace technik HDR pro UAV ukázalo neexistenci datasetu, který by umožnil vývoj a otestování algoritmů zpracování HDR videa. Pro vývoj je potřebné vytvořit dataset, který by měl obsahovat nekomprimované snímky se střídající se známou expozicí. Provedené testy ukazují, že záznam nekomprimovaných snímků s vysokou snímkovací frekvencí je obtížně proveditelný, protože USB kamery nemají dostatečnou frekvenci snímání. Byl navrhnout experiment, který otestuje možnost instalace zvolené kamery na dron pro záznam datasetu. Na dron bude instalována kamera a platforma pro záznam nekomprimovaných snímků.

Byly analyzovány možnosti použití běžně dostupných kamer (např. od

společnosti Basler, Obrázek). Výhodou je nízká cena a jednoduchá integrace přes rozhraní USB, které je dobře dostupné také na platformě NVidia Tegra TX2. Pro jednotlivé snímky nebo video je kvalita výsledného obrazu dostatečná, problém nastává při skládání HDR, kde jednotlivé snímky nejsou vhodné pro skládání kvůli postprocessingu obrazu a nelineární odezvě senzoru, což způsobuje velký šum ve výsledných HDR snímcích. Největší problém je v rychlosti snímání, kdy jsou kamery připojené přes rozhraní USB 3.0 schopny poskytovat typicky 10-12 snímků za sekundu (při řízení expozice pro každý snímek samostatně). Vytvoření HDR videa s nízkou snímkovou frekvencí vede na video s nízkou kvalitou, které je velmi náchylné k vzniku artefaktů a ghostingu (vzniku duchů).

Kamery od společnosti Camea poskytují obraz ve formátu RAW, streamovaný rozhraním Ethernet. Výhodou je možnost řízení délky expozice pro každý snímek zvlášť bez omezení rychlosti framerate. Další výhodou je použití kvalitních CMOS senzorů Sony Exmor, které poskytují velmi kvalitní obraz. Nevýhodou těchto kamer je rozhraní Ethernet, přes které jsou vysílány snímky ve formátu RAW, což vyžaduje zařízení schopné takový datový tok přijmout - což může být u některých síťových karet a méně výkonnějších počítačů problém.

Frekvence [FPS]	1	10	20	30	40
Datový tok [$MB.s^{-1}$]	1.977	19.77	39.55	59,31	79,1

Tabulka 1: Rychlost odesílání dat pro různé snímkovací frekvence při FullHD rozlišení.

V případě FullHD rozlišení má velikost jednoho snímku cca. 2MB (1 Byte na pixel), při 30 FPS je třeba přenést až $60MB.s^{-1}$ (více v tabulce 1). Vzhledem k rychlosti odesílání dat z kamery se ukázalo, že Tegra TX2 není schopna přijmout datový tok beze ztráty paketů a uložit tento datový tok na úložiště SSD. Pro experiment byla navržena platforma, která je složena z existujících komponent a zařízení. Kameru poskytla společnost Camea a zařízení pro záznam z Ethernetu bylo spoluvyvinuto na fakultě VUT FIT, kde FIT VUT vyvinul firmware pro speciální záznamové zařízení v rámci jiného projektu. Zařízení bylo otestováno pro příjem dat přes Ethernet a je schopno zaznamenat datový tok z kamery až do rychlostí $1Gbps$, což je pro potřeby experimentu postačující. Pro experiment bude použit dron, který je dostupný na VUT FIT.

3 Závěr

Report obsahuje úvod do problematiky HDR a detekce objektů pro UAV, kde je omezení na výpočetní zdroje ale i hmotnost a velikost zařízení. V reportu byly analyzovány možnosti použití komerčně dostupných i specializovaných kamer pro drony. Byl navrhnut experiment pro pořízení datasetu, který umožní otestovat možnosti implementace pořizování HDR videa přímo dronem. Bude vykonán experiment, při kterém bude pořízen dataset multi-expozičních snímků, který bude využit pro další experimenty a analýzu.

Literatura

- [1] B. Příbyl, A. Chalmers, P. Zemčik, L. Hooberman, and M. Čadík, “Evaluation of feature point detection in high dynamic range imagery,” *Journal of Visual Communication and Image Representation*, vol. 38, pp. 141 – 160, 2016.
- [2] H. Zhao, B. Shi, C. Fernandez-Cull, S.-K. Yeung, and R. Raskar, “Unbounded high dynamic range photography using a modulo camera,” in *ICCP*, 2015.
- [3] P. E. Debevec and J. Malik, “Recovering high dynamic range radiance maps from photographs,” SIGGRAPH '97, (New York, NY, USA), pp. 369–378, ACM Press/Addison-Wesley Publishing Co., 1997.
- [4] T. Mertens, J. Kautz, and F. V. Reeth, “Exposure fusion,” in *Computer Graphics and Applications, 2007. PG '07. 15th Pacific Conference on*, pp. 382–390, Oct 2007.
- [5] M. Yazdi and T. Bouwmans, “New Trends on Moving Object Detection in Video Images Captured by a moving Camera: A Survey,” *Computer Science Review*, 2018.