

Efektywna metoda uzyskiwania obrazów HDR dla systemów CCTV

Streszczenie. W artykule zaproponowano metodę przetwarzania obrazów przeznaczoną dla systemów CCTV (ang. closed circuit television) nadzorowanych przez człowieka (operatora monitoringu). Proponowane rozwiązanie poprawia szczegółowość obrazu uzyskiwanego w trudnych warunkach oświetleniowych, szczególnie dla obszarów nierównomiernie oświetlonych po zmroku lub przy silnym nasłonecznieniu. Do przetwarzania wykorzystano pary obrazów niedoświetlonych i prześwietlonych pochodzące z kamer o typowych przetwornikach 8-bitowych oraz z jednokrotnym naświetleniem dla przetworników o większej rozdzielczości niż 8 bitów.

Abstract. In this paper an image processing method designed for CCTV (closed circuit television) systems supervised by the man (typically the monitoring operator) is presented. The proposed method enhances details of image obtained in difficult lighting conditions, especially for areas unevenly illuminated at night or in strong sunlight. Images acquired using cameras with converters of 8-bit resolution (overexposed and underexposed frames) or more than 8-bit are processed. (An effective method of HDR images generation for CCTV systems).

Słowa kluczowe: HDR, mapowanie tonalne, CCTV, przetwarzanie obrazów, poprawa jakości obrazów.

Keywords: HDR, tone mapping, CCTV, image processing, image quality improvement.

Wprowadzenie

Systemy monitoringu wizyjnego CCTV (closed circuit television) są rozwijane w celu zapewnienia bezpieczeństwa ludzi, głównie w obszarach zurbanizowanych. Wykorzystywane w nich kamery charakteryzują się coraz lepszymi parametrami technicznymi, coraz mniejszymi szumami, dużą czułością i rozdzielczością, wciąż jednak występują problemy z jakością obrazu za dnia w przypadku dużych kontrastów oraz w nocy, gdy kamera jest oślepiana [1]. Problem ten dotyczy nie tylko monitoringu miejskiego, ale także wielu innych systemów wizyjnych np. wspomaganie kierowców pojazdów, pojazdów autonomicznych, czy awioniki.

Możliwym rozwiązaniem powyższych problemów jest wyznaczenie obrazów o poszerzonym zakresie dynamicznym – HDR (ang. high dynamic range) [2, 3]. Obrazy HDR typowo uzyskuje się z przetworzenia kilku obrazów o różnych parametrach naświetleń lub z przetworników obrazowych o większej rozdzielczości niż 8 bitów za pomocą technik mapowania tonalnego. Techniki te, jak np. opisane w [4, 5] nie pozwalają uzyskać dostatecznej szczegółowości obrazów w miejscach prześwietlonych i niedoświetlonych równocześnie. Wynikowe obrazy HDR zapisywane są w formatach możliwych do wyświetlenia przez typowe monitory LCD, a więc o rozdzielczości 8 bitów na każdą składową koloru (R, G, B).

Metody HDR są coraz częściej oferowane w aparatach fotograficznych i oprogramowaniu komputerowym, jednak stopień skomplikowania wykorzystywanych w nich algorytmów zazwyczaj nie pozwala na ich wykorzystanie w systemach wideo pracujących w czasie rzeczywistym. Spotykane adaptacyjne metody przyspieszania procesu przetwarzania mogą powodować niestabilność algorytmu zestawiania obrazów [6]. Co prawda na rynku są już dostępne kamery CCTV mające zaimplementowane obrazowanie HDR [7, 8], jednak zastosowane w nich algorytmy przetwarzania obrazów nie zostały dokładnie opisane, a uzyskane efekty nie zawsze są satysfakcjonujące, głównie ze względu na niedostateczną liczbę szczegółów, przeregulowania wokół krawędzi obiektów i nienaturalny lokalny kontrast obrazu.

Tak więc istnieje potrzeba opracowania wydajnej metody uzyskiwania wynikowych obrazów HDR. Umożliwiłaby ona zwiększenie wyrazistości obrazu (zwłaszcza w miejscach prześwietlonych i

niedoświetlonych), co może przełożyć się na skuteczność detekcji zagrożeń [9, 10] oraz zwiększenie niezawodności pracy systemów wizyjnych. Ważną cechą technik HDR jest możliwość dostarczenia bardzo szczegółowych obrazów, nawet za cenę realizmu, tak aby można było dokonać detekcji jak największej liczby istotnych szczegółów [11].

Autorzy w swoich wcześniejszych pracach zajmowali się analizą obrazów ze zwiększonym zakresem dynamicznym [12]. Opisali również zagrożenia występujące w środowisku zurbanizowanym, które można wykrywać za pomocą metod przetwarzania obrazów cyfrowych [13]. Zaprezentowana w niniejszym artykule wydajna i efektywna metoda obrazowania HDR wykorzystuje doświadczenia zdobyte w dotychczasowych badaniach.

Akwizycja obrazów

Zaproponowana metoda uzyskiwania wynikowych obrazów HDR opiera się na założeniu, że w obrazie prześwietlonym szczegółowość (rozumiana jako bezwzględne wartości różnic pomiędzy sąsiednimi pikselami) jest znacznie większa w miejscach niedoświetlonych, natomiast w obrazie niedoświetlonym – w miejscach prześwietlonych. Dlatego obraz wynikowy o powiększonej szczegółowości można uzyskać na podstawie:

- dwóch kolejno naświetlonych obrazów o znacząco różnym czasie ekspozycji – jednego prześwietlonego i drugiego niedoświetlonego. Oba obrazy mogą pochodzić z kamery wyposażonej w standardowy przetwornik obrazowy o rozdzielczości 8-bitowej (w przypadku kamer monochromatycznych lub 8-bitów dla każdej ze składowych R, G, B),
- jednego obrazu pochodzącego z kamery wyposażonej w przetwornik więcej niż 8-bitowy.

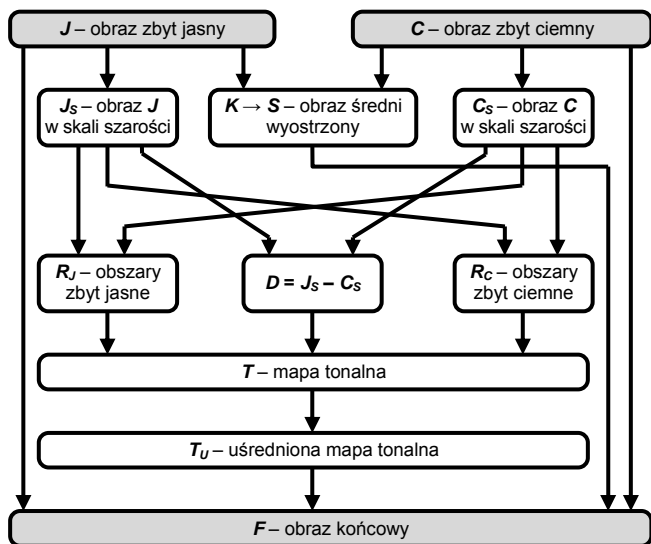
W pierwszym przypadku przyobrazy można uzyskać poprzez cykliczne, naprzemienne zmienianie parametrów naświetlenia matrycy światłoczułej (czasu otwarcia migawki lub czasu akwizycji sygnału w kamerach bez migawki mechanicznej). Z przeprowadzonych badań wynika, że najlepsze efekty uzyskuje się wykorzystując obraz prześwietlony o 2 EV (*Exposure Value*) oraz niedoświetlony o 2 EV. Prześwietlenie i niedoświetlenie jest odniesione do optymalnej ekspozycji wyznaczonej na podstawie średniej jasności wybranych obszarów kadru lub całego obrazu. Różnicę +2 EV uzyskuje się typowo przez czterokrotne wydłużenie czasu naświetlania, a -2 EV przez czterokrotne

skrócenie czasu naświetlania. W przypadku rejestracji obiektów ruchomych, co jest sytuacją typową w monitoringu, możliwe jest powstanie przesunięć obiektów pomiędzy kolejnymi naświetleniami. Okazuje się jednak, że uzyskanie obrazu niedoświetlonego o bardzo krótkim czasie naświetlania zaraz po prześwietlonym sprawia, że przesunięcia obiektów na obrazach są praktycznie niezauważalne.

W drugim przypadku (kamera z przetwornikiem obrazowym o rozdzielczości większej niż 8 bitów) obraz niedoświetlony i prześwietlony pochodzą z jednej zarejestrowanej klatki, przy czym pierwszy z nich wykorzystuje dolny zakres przetwarzania przetwornika, a drugi – zakres górny. Dzięki temu jest wykorzystywany znacznie szerszy zakres dynamiczny przetwornika niż w przypadku pozyskania jednego obrazu o najwyższym kontraście.

Mapowanie tonalne

Pozyskane pary obrazów są przetwarzane przez algorytm mapowania tonalnego w celu uzyskania pojedynczych obrazów wynikowych. Ogólny schemat proponowanej metody przedstawiono na rys. 1.



Rys.1. Schemat proponowanej metody mapowania tonalnego dla systemów CCTV

Załóżmy, że w celu uzyskania obrazu o poszerzonym zakresie dynamicznym wykorzystuje się obraz J (zbyt jasny) i C (zbyt ciemny) zapisane w 8-bitowym systemie RGB (zakres wartości każdej składowej: 0–255).

W pierwszym etapie z obrazów J i C uzyskuje się obraz K o powiększonym kontraście korzystając z formuły:

$$(1) \quad K_{x,y,z} = 2 \cdot \left(\frac{C_{x,y,z} + J_{x,y,z}}{1,5} - 42,5 \right),$$

gdzie: x, y – współrzędne pikseli obrazu, z – składowa R, G lub B obrazu.

Obraz K jest następnie wyostrzony korzystając z zależności:

$$(2) \quad S_{x,y,z} = 2 \cdot K_{x,y,z} - \bar{K}_{x,y,z},$$

gdzie: $S_{x,y,z}$ – wartość składowej R, G lub B pikseli obrazu wyostrzonego, $\bar{K}_{x,y,z}$ – uśredniona wartość składowej R, G lub B pikseli obrazu K dla pikseli z zakresu $x-2:x+2, y-2:y+2, z$.

Niezależnie obrazy J i C konwertuje się do skali szarości korzystając z typowej zależności:

$$(3) \quad J_{x,y} = 0,299R_{x,y} + 0,587G_{x,y} + 0,114B_{x,y}.$$

W ten sposób powstają, odpowiednio, obrazy J_S i C_S .

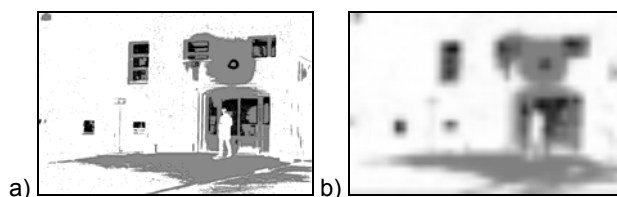
W następnym etapie, na podstawie zależności (4), wyznacza się macierze R_J i R_C zawierające odpowiednio wartości lokalnych zmian kontrastu dla obrazów J_S i C_S :

$$(4) \quad R_{J,x,y} = |J_{S,x,y} - J_{S,x+1,y}| + |J_{S,x,y} - J_{S,x,y+1}| \\ R_{C,x,y} = |C_{S,x,y} - C_{S,x+1,y}| + |C_{S,x,y} - C_{S,x,y+1}|$$

Wartości $R_J > R_C$ dla danego pikseli wskazują, że powinien on w obrazie końcowym pochodzić z obrazu zbyt jasnego, jako zawierającego więcej szczegółów niż obraz zbyt ciemny. Analogicznie wartości $R_J < R_C$ wskazują, że dany fragment obrazu końcowego może być prześwietlony i powinien pochodzić z obrazu zbyt ciemnego. Oprócz tego wyznacza się macierz $D = J_S - C_S$. Wartości tej macierzy większe niż ustalony próg P (jeden parametr regulacyjny metody) wskazują, że odpowiadające im piksele obrazu końcowego powinny pochodzić z obrazu S . Duże wartości D uzyskuje się tylko dla pikseli, które miały bardzo duże wartości w macierzy J_S i bardzo małe w macierzy C_S . Oznacza to, że piksele te reprezentują fragmenty obrazu, które byłyby poprawnie naświetlone na obrazie uzyskanym bez korekcji ekspozycji. W związku z powyższym do wyznaczenia mapy tonalnej T zawierającej informację o tym, które fragmenty obrazu końcowego powinny pochodzić z obrazów J, S i C wykorzystuje się zależność:

$$(5) \quad T_{x,y} = \begin{cases} 0 & \text{gdy } R_{C,x,y} > R_{J,x,y} \quad \wedge \quad R_{x,y} < P \\ 127,5 & \text{gdy } R_{x,y} \geq P \quad \vee \quad R_{C,x,y} = R_{J,x,y} \\ 255 & \text{gdy } R_{J,x,y} > R_{C,x,y} \quad \wedge \quad R_{x,y} < P \end{cases}$$

W celu uniknięcia powstania w obrazie końcowym pseudokrawędzi na granicach obszarów pochodzących z obrazów J, S i C otrzymaną mapę tonalną T uśrednia się otrzymując mapę T_U . Doświadczalnie stwierdzono, że najlepsze do tego celu będzie zastosowanie filtru uśredniającego o masce kwadratowej, której długość boku będzie wynosić około 1/15 krótszego boku obrazu. Przykładowe mapy tonalne T i T_U (uzyskane dla obrazów z górnego wiersza rys. 4) przedstawia rys. 2.



Rys.2. Przykładowa mapa tonalna: a) T – przed uśrednieniem, b) T_U – po uśrednieniu

Ostatnim etapem jest mapowanie tonalne, w którym obraz końcowy powstaje jako złożenie obrazów J, S i C na podstawie uzyskanej mapy tonalnej T_U . Wartości T_U określają zawartość obrazów J, S i C w końcowym obrazie F na podstawie zależności:

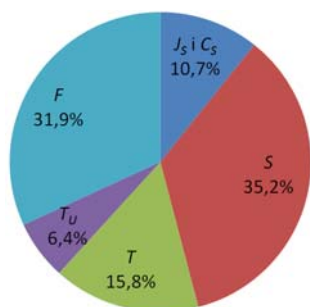
$$(6) \quad F_{x,y,z} = \begin{cases} S_{x,y,z} + \frac{T_{U,x,y,z} - 127,5}{127,5} \cdot (J_{x,y,z} - S_{x,y,z}) & \text{dla } T_{U,x,y,z} > 127,5 \\ C_{x,y,z} + \frac{T_{U,x,y,z}}{127,5} \cdot (S_{x,y,z} - C_{x,y,z}) & \text{dla } T_{U,x,y,z} \leq 127,5 \end{cases}$$

Szczegółowość obrazu F zależy od wartości parametru P (por. (5)). Im większa wartość parametru P , tym udział obrazu średniego wyostrzonego S będzie większy. Doświadczalnie stwierdzono, że wartość parametru P dla obrazów 8-bitowych powinna mieścić się w zakresie od 140 do 200.

Wyniki eksperymentów w systemach monitoringu

Proponowaną metodę zaimplementowano w środowisku programistycznym Matlab. Do eksperymentów wykorzystano kolorowe obrazy testowe uzyskane z aparatu Canon 70D w trudnych warunkach oświetleniowych:

- po zmierzchu, gdy część kadru była wyraźnie niedoświetlona, a pozostała – prześwietlona,
- w słoneczny dzień, gdy jedna część kadru obejmowała obszary pozostające w cieniu, a druga – silnie nasłonecznione.



Rys.3. Procentowy udział czasu potrzebnego na realizację poszczególnych etapów algorytmu – wygenerowanie poszczególnych obrazów (J_S – obraz zbyt jasny w skali szarości, C_S – obraz zbyt ciemny w skali szarości, S – obraz średni wyostrzony, T – mapa tonalna, T_U – uśredniona mapa tonalna, F –

obraz końcowy)

Obrazy te rejestrowano w 14-bitowym formacie RAW z rozdzielczością 5472×3648 pikseli i konwertowano przy użyciu oprogramowania Digital Photo Professional 4 do rozdzielczości 720×480 pikseli tworząc obraz zbyt jasny J z korektą +2 EV oraz zbyt ciemny C z korektą -2 EV. Średni czas uzyskiwania obrazu końcowego na komputerze z procesorem Intel Core i5-6600 oraz 16 GB DDR4 RAM wynosił 0,12 s. Wykres na rys. 3. przedstawia procentowy udział czasu przeznaczanego na wykonanie poszczególnych etapów algorytmu.

W celu przyspieszenia obliczeń proponowana metoda może być zrealizowana wielowątkowo w szybszym niż Matlab środowisku programistycznym lub w systemie wieloprocesorowym. Generacja obrazów J_S , C_S i S , podobnie jak R_J , R_C i R może odbywać się równolegle. Niezależnie od powyższych założeń, poszczególne fragmenty obrazów mogą również być przetwarzane osobno, co pozwoli znacznie skrócić czas obliczeń.

Rysunek 4 przedstawia przykładowe rezultaty uzyskane za pomocą zaproponowanej metody przetwarzania obrazów zarejestrowanych w trudnych warunkach oświetleniowych. Na podstawie eksperymentów przeprowadzonych dla 15 różnych obrazów testowych stwierdzono, że w największą szczegółowość i naturalność obrazów uzyskuje się dla parametru $P = 140$ dla obrazów rejestrowanych w ciągu dnia oraz $P = 200$ dla ujęć nocnych. Należy zwrócić uwagę, że zaproponowana technika mapowania tonalnego umożliwiła równoczesne zachowanie szczegółów w obszarach zbyt ciemnych i zbyt jasnych (por. rys. 4.: wnętrza za drzwiami i w bramie – wiersze górny i dolny, widok zza okna – wiersz środkowy).



Rys.4. Uzyskane rezultaty przetwarzania obrazów proponowaną metodą: a) obrazy zbyt jasne, b) obrazy zbyt ciemne, c) obrazy końcowe F dla parametru $P=140$, d) obrazy końcowe F dla parametru $P=200$

Podsumowanie

Zaproponowano efektywną metodę uzyskiwania obrazów o poszerzonym zakresie dynamicznym, przeznaczoną dla systemów monitoringu nadzorowanych przez człowieka. Do jej najważniejszych zalet należą:

- duża szczegółowość obrazów końcowych w miejscach, które byłyby niedoświetlone lub prześwietlone na obrazach źródłowych,
- brak zauważalnych zniekształceń i artefaktów,

- redukcja efektów oślepienia kamery silnym źródłem światła skierowanym w obiektyw,
- konieczność doboru tylko jednego parametru regulacyjnego.

W dalszym etapie prac autorzy planują implementację proponowanego rozwiązania w rzeczywistym systemie wizyjnym zapewniającym przetwarzanie równoległe, tak aby uzyskać szybkość 25 klatek/s dla obrazów o rozdzielczości D1 (720x576 – PAL lub 720x480 – NTSC).

Autorzy: mgr inż. Adam Konieczka, Politechnika Poznańska, Wydział Informatyki, ul. Piotrowo 3a, 60-965 Poznań, E-mail: adam.konieczka@put.poznan.pl;
mgr inż. Julian Balcerek, Politechnika Poznańska, Wydział Informatyki, ul. Piotrowo 3a, 60-965 Poznań, E-mail: julian.balcerek@put.poznan.pl;
dr inż. Paweł Pawłowski, Politechnika Poznańska, Wydział Informatyki, ul. Piotrowo 3a, 60-965 Poznań, E-mail: pawel.pawlowski@put.poznan.pl;

LITERATURA

- [1] Aguerrebere C., Delon J., Gousseau Y., Muse P., *Best algorithms for HDR image generation. A study of performance bounds*, SIAM Journals, 7 (1), (2013), 1-34.
- [2] Reinhard E., Ward G., Pattanaik S., Debevec P., *High dynamic range imaging. Acquisition, display, and image-based lighting*, The Morgan Kaufmann Series in Computer Graphic, Elsevier, 2006.
- [3] Myszkowski K., Mantiuk R., Krawczyk G., *High Dynamic Range Video*, Morgan & Claypool Publishers, 2008.
- [4] Durand F., Dorsey J., *Fast Bilateral Filtering for the Display of High-Dynamic-Range Images*, ACM Transactions on Graphics, vol. 21, issue 3 (2002), 257-266.
- [5] Reinhard E., Stark M., Shirley P., Ferwerda J., *Photographic Tone Reproduction for Digital Images*, ACM Transactions on Graphics – Proc. of ACM SIGGRAPH, vol. 21 issue 3, (2002), 267-276.
- [6] Li Y., Qiao Y., and Ruichek Y., *Multiframe-Based High Dynamic Range Monocular Vision System for Advanced Driver Assistance Systems*, IEEE Sensors Journal, vol. 15, issue 10, (2015), 5433-5441.
- [7] <http://www.networkwebcams.co.uk/blog/2011/02/02/hdr-high-dynamic-range-video-and-ip-security-cameras/>, dostęp: 24.03.2016 r.
- [8] <http://www.cctvcamerapros.com/Wide-Dynamic-Range-Camera-p/wdr-200dn.htm>, dostęp: 24.03.2016 r.
- [9] Balcerek J., Drgas S., Dąbrowski A., Konieczka A., *Prototype multimedia database system for registration of emergency situations*, Proc. of Signal Processing – Algorithms, Architectures, Arrangements and Applications SPA'2009, IEEE International Conference, 144-148, Poznań, Poland, 24–26 September 2009.
- [10] Balcerek J., Pawłowski P., Konieczka A., Drgas S., Dąbrowski A., Kmiecik M., *Database of emergency telephone calls – system tools for real-time registration and metadata searching*, Proc. of Signal Processing – Algorithms, Architectures, Arrangements and Applications SPA'2010, IEEE International Conference, pp. 89-94, Poznań, Poland, 23–25 September 2010.
- [11] Kuang J., Yamaguchi H., Liu C., Johnson G.M., Fairchild M. D., *Evaluating HDR Rendering Algorithms*, Transactions on Applied Perception, 4 Article 9, 2007.
- [12] Konieczka A., Balcerek J., Dąbrowski A., *Improvement of dynamic range of photography in RAW format representation*, Proc. of Signal Processing – Algorithms, Architectures, Arrangements and Applications SPA'2009, IEEE International Conference, pp. 62-65, Poznań, Poland, 24–26 September 2009.
- [13] Balcerek J., Chmielewska A., Dąbrowski A., Jackowski D., Konieczka A., Marciniak T., Pawłowski P., *Recognition of threats in urban areas by means of the analysis of video sequences*, Proc. of Multimedia Communications, Services and Security, MCSS 2010, IEEE International Conference, pp. 41-48, Kraków, 6–7 May 2010.