

Politechnika Poznańska
Wydział Informatyki
Instytut Automatyki i Robotyki
Zakład Układów Elektronicznych i Przetwarzania Sygnałów

Streszczenie rozprawy doktorskiej

**Automatic vision quality enhancement
of CCTV data**

**(Automatyczna poprawa jakości wizyjnej
danych CCTV)**

Adam Konieczka

Promotor: prof. dr hab. inż. Adam Dąbrowski

Poznań 2017

1. Wprowadzenie

W ostatnich latach obserwuje się gwałtowny wzrost zastosowań urządzeń do rejestracji obrazu. Niezależnie od rynków kamer konsumenckich i profesjonalnych, wykorzystywanych przez stacje telewizyjne, dynamicznie rozwijają się systemy monitoringu wizyjnego, nazywane też telewizją przemysłową – CCTV (ang. closed-circuit television) [1]. W takie systemy wyposażane są obiekty i tereny użyteczności publicznej, jak np. lotniska, szkoły czy dworce. Oprócz tego znajdują one zastosowanie w przemyśle podczas kontroli jakości produkcji. W kamery wyposażonych jest wiele przenośnych urządzeń, takich jak smartfony czy rejestratory jazdy dla kierowców. Systemy wizyjne coraz częściej stanowią też wyposażenie pojazdów, pełniąc rolę wstecznych lusterek lub kamer cofania.

Najprawdopodobniej pierwszy system wizyjny stworzyła firma Siemens AG i był on używany przez nazistowskie Niemcy do nadzorowania eksperymentów z raketami V2 podczas II wojny światowej [2]. Przyjmuje się, że kamery po raz pierwszy wykorzystano w monitoringu miejskim w 1960 r. w Londynie, gdzie brytyjska Policja tymczasowo zamontowała dwie kamery do monitorowania tłumu. Jednakże już w roku 1969 co najmniej 67 kamer było używanych przez różne służby. Wtedy też zaczął się gwałtowny wzrost wykorzystania takich urządzeń [3], by w 2010 r. ich liczba w całej Wielkiej Brytanii przekroczyła 4 miliony, z czego w samym Londynie było używanych ponad pół miliona kamer.

Także w polskich miastach liczba kamer gwałtownie rośnie. W 2012 r. 89% przebadanych miast dysponowało systemami monitoringu miejskiego [4]. W samym Poznaniu było wtedy 356 kamer (w 2016 r. było ich już 525), a w Warszawie – 414. Jednakże zdecydowana większość kamer nie jest zintegrowana. Tylko w Warszawie takich urządzeń było ponad 5000, a np. dalszych 2000 znajdowało się w szkołach. Obecnie szacuje się, że jest ich łącznie ok. 13,5 tys. [5].

Obraz z kamer może być przechowywany z wykorzystaniem rejestratorów i analizowany przez ludzi (zwykle są to operatorzy monitoringu pracujący w centrach monitoringu) bądź przez systemy automatyczne. Taka analiza służy poprawie bezpieczeństwa publicznego, szczególnie w środowiskach silnie zurbanizowanych. Szybka detekcja zagrożeń na podstawie transmitowanego na żywo obrazu przyspiesza interwencję odpowiednich służb. Zarejestrowany materiał filmowy coraz częściej stanowi dowód w postępowaniach sądowych.

Coraz powszechniejsze są także półautomatyczne systemy, które mogą wspierać operatorów monitoringu w wykrywaniu potencjalnie niebezpiecznych sytuacji w oparciu o automatyczną analizę. Systemy te wykorzystują algorytmy rozpoznawania twarzy, chodu bądź innej aktywności ludzkiej [6, 7] i są często nazywane „inteligentnym monitoringiem”. Obecnie umożliwia on wykrywanie, zliczanie i śledzenie pojedynczych osób w czasie rzeczywistym [8, 9]. Niektóre systemy, jak np. [10] wykorzystują połączenie wielu technik przetwarzania obrazu w celu zapewnienia bezpieczeństwa w transporcie zbiorowym. Podobne rozwiązania stosuje się też do kontroli i zarządzania ruchem ulicznym. Wielu producentów oferuje już gotowe rozwiązania do tego celu.

W przemyśle kamery zaczęto wykorzystywać pod koniec lat 60. XX w. [11], jednak dopiero w latach 80. zaczął się gwałtowny wzrost ich użycia. Obecnie tzw. widzenie maszynowe jest stosowane do rozpoznawania wzorów, kolorów, kształtów, wielkości obiektów itp. Dlatego stosuje się je m.in. w kontroli produkcji oraz podczas inspekcji trudno dostępnych miejsc. Kamery są powszechnie stosowane w robotyce, np. do lokalnej nawigacji robotów mobilnych [12]. Ważną gałęzią widzenia maszynowego są także systemy OCR (ang. optical character recognition) służące do rozpoznawania znaków w obrazie.

W systemach CCTV stosuje się wiele rodzajów kamer oferujących obraz o różnych rozdzielczościach. Najczęściej w monitoringu stosuje się kamery kolorowe, natomiast w przemyśle – monochromatyczne. Wiele kamer pracuje także w bliskiej podczerwieni i dzięki zastosowaniu specjalnych oświetlaczy ich użycie jest niewidoczne dla człowieka. Ponadto na rynku są dostępne urządzenia pracujące w dalekiej podczerwieni, które w całkowitej ciemności pozwalają dostrzec człowieka z odległości nawet kilku kilometrów [13]. Do specjalnych zastosowań są oferowane m.in. kamery stereowizyjne (wyposażone w dwa obiektywy), linijkowe i szybkobieżne. Ponadto można wyróżnić tzw. kamery inteligentne (ang. smart cameras), które są wyposażone w procesory, dzięki czemu przetwarzanie rejestrowanego obrazu może się odbywać bez użycia dodatkowych urządzeń. Wiele kamer posiada dodatkowe funkcje, które poszerzają zakres dynamiczny rejestrowanego obrazu (WDR – ang. wide dynamic range), zwiększają czułość obrazu przy niedostatecznym oświetleniu, kompensują wpływ źródeł światła znajdujących się za filmowanym obiektem (BLC – ang. backlight compensation), redukują szum, włączają rejestrację obrazu po wykryciu ruchu [14].

Niestety, jakość rejestrowanego materiału wideo wciąż nie jest wystarczająca i ogranicza możliwość wykorzystania wielu algorytmów automatycznego przetwarzania obrazów [6]. Wiele publikacji, jak np. [15], wskazuje, że jakość strumienia wideo jest bardzo ważna w systemach monitorowania miejsc publicznych i bardzo często zbyt niska. Także raport polskiej Najwyższej Izby Kontroli [16] wskazuje, że niedostateczna jakość obrazu często utrudnia

rozpoznanie poszczególnych osób. Mimo tego, nawet oficjalne rekomendacje dla systemów CCTV nie precyzują szczegółowo parametrów tworzonych systemów, a jedynie zwracają uwagę na konieczność dalszej poprawy jakości rejestrowanego obrazu. Jest to niezbędne w celu dalszego zwiększenia użyteczności takich rozwiązań.

2. Cele, teza naukowa i organizacja pracy

Celem niniejszej pracy była analiza i rozwój metod cyfrowego przetwarzania obrazów, które poprawiają jakość i zwiększą użyteczność obrazów uzyskiwanych w systemach CCTV, szczególnie w trudnych, zróżnicowanych warunkach oświetleniowych. Obszar badawczy ograniczono do dwóch głównych zagadnień:

1. usuwania szumu, który występuje w nagraniach wykonanych przy niedostatecznym oświetleniu,
2. lokalnego poszerzania zakresu dynamicznego obrazu, co jest szczególnie istotne, gdy część kadru jest wyraźnie niedoświetlona, a część – prześwietlona.

Teza naukowa rozprawy doktorskiej została sformułowana następująco: opracowane i eksperymentalnie zweryfikowane metody redukcji szumu oraz powiększenia lokalnego zakresu dynamicznego obrazu w systemach wizyjnych zwiększają skuteczność ekstrakcji zawartości informacyjnej ze strumienia wideo zarówno w systemach nadzorowanych przez człowieka oraz w systemach automatycznych.

W celu rozwiązania problemu badawczego autor opracował dwie metody poprawiające jakość obrazu. Pierwsza z nich jest przeznaczona do redukcji addytywnego szumu o charakterze impulsowym oraz, szczególnie, szumu będącego mieszaniną szumu gaussowskiego i szumu impulsowego. Taki szum jest charakterystyczny dla przetworników stosowanych w kamerach. Druga zaproponowana metoda, będąca lokalnym operatorem mapowania tonalnego, umożliwia zachowanie istotnych szczegółów filmowanego kadru w miejscach prześwietlonych (które byłyby widoczne na finalnym obrazie jako całkowicie białe) i niedoświetlonych (które byłyby czarne). Szczegóły te są zwykle tracone podczas redukcji zakresu dynamicznego obrazu zarejestrowanego przez przetworniki obrazowe w trudnych warunkach oświetleniowych do zakresu możliwego do wyświetlenia przy użyciu standardowych ekranów LCD (ang. liquid-crystal display).

Skuteczność powyższych metod została potwierdzona eksperymentalnie, a porównanie uzyskanych wyników z dostępnymi w literaturze wskazało na wysoką skuteczność zaproponowanych rozwiązań.

Praca składa się z 7 rozdziałów i liczy 200 stron, łącznie z 35-stronicowymi załącznikami. Została zorganizowana w następujący sposób. W rozdziale pierwszym nakreślono obszar badawczy, wskazano na szerokie zastosowanie systemów CCTV oraz problemy, które ograniczają ich rozwój. Następnie przedstawiono cel i tezę pracy oraz wymieniono główne osiągnięcia naukowe.

Rozdział drugi prezentuje stan wiedzy w zakresie badawczym, przedstawiając przeznaczenie, elementy składowe i budowę systemów CCTV, ze szczególnym uwzględnieniem kamer. Wymieniono ograniczenia jakości obrazu w takich systemach koncentrując się na szumach i sposobach ich redukcji, metodach poszerzania zakresu dynamicznego i ocenie jakości przetworzonych obrazów.

Rozdział trzeci przedstawia proponowaną metodę redukcji szumu, dobór jej parametrów pracy oraz sposób przetwarzania różnych rodzajów obrazów. W kolejnym rozdziale przedstawiono eksperymenty weryfikujące skuteczność ww. metody przy zastosowaniu obrazów testowych zawierających różne rodzaje szumu. Porównano także uzyskane wyniki z dostępnymi w literaturze.

Rozdział piąty poświęcono proponowanej metodzie lokalnego poszerzania zakresu dynamicznego obrazów, wykorzystującej informację o jasności i kontraście rejestrowanych obrazów. Opisano sposób uzyskiwania obrazów testowych i algorytm ich przetwarzania z wykorzystaniem autorskiego sposobu generacji map tonalnych oraz opracowanego detektora krawędzi. Następny rozdział prezentuje eksperymenty weryfikujące skuteczność powyższej metody z wykorzystaniem przygotowanych obrazów testowych. Obrazy te przetworzono przy użyciu opracowanej metody oraz konkurencyjnych metod, opisanych w literaturze. Przydatność metod oceniono zarówno dla systemów nadzorowanych przez ludzi jak i automatycznych.

Ostatni rozdział zawiera krótkie podsumowanie pracy oraz wnioski. W załącznikach znajdują się szczegółowe wyniki odszumiania z zastosowaniem proponowanej metody.

3. Stan wiedzy

Standardowy schemat rejestracji obrazu przez kamerę, jak np. [17], zakłada, że obraz padający przez obiektyw za pośrednictwem mechanizmu migawki trafia na przetwornik obrazowy, który go próbkuje i kwantuje. Następnie obraz jest przetwarzany w procesie postprocessingu obejmującym linearyzację, korekcję balansu bieli, usunięcie efektu mozaiki (spowodowanego stosowaniem filtra Bayera), korekcję koloru, jasności i kontrastu. Tak uzyskany obraz jest dopiero zapisywany lub transmitowany.

Kamery zwykle wykorzystują dwa rodzaje przetworników: CCD (ang. charge coupled device) oraz CMOS (ang. complementary metal-oxide semiconductor), które wypierają CCD. Rozdzielczości przetworników wykorzystywanych we współczesnych kamerach CCTV wynoszą zwykle kilka megapikseli, zaś wartość każdego punktu obrazu jest zapisywana przy użyciu od 8 do nawet 20 bitów [14, 18].

W [19] stwierdzono, że o jakości obrazu CCTV decydują cztery parametry: wyraźność (wystarczająca ostrość bez zniekształceń), szczegółowość, naturalność koloru i występowanie artefaktów. Niestety, obraz z nawet dobrej jakości kamer może być niewystarczający do dalszej analizy, gdy kamera jest oślepiana przez silne źródło światła znajdujące się za filmowanym obiektem. W rozprawie zwrócono uwagę, że jest to poważny problem dla operatorów systemów monitoringu miejskiego, gdy jedna część kadru jest oświetlona silnym światłem słonecznym, a druga znajduje się w głębokim cieniu i szczegółowość obrazu w miejscach prześwietlonych oraz niedoświetlonych jest zbyt niska. Taka sama trudność występuje też, m.in., dla motorniczych tramwajów, w których lusterka wsteczne zastąpiono kamerami [20]. Kamera oślepiona przez słońce bądź przez światła przejeżdżających obok samochodów nie jest w stanie pokazać na monitorze o ograniczonym zakresie jasności, czy np. drzwi ruszającego pojazdu nie przytrzasnęły wysiadającego pasażera.

Dotąd nie opracowano powszechnie używanych standardów oceny jakości obrazu w systemach CCTV, pomimo że istnieje wiele rozwiązań do oceny jakości obrazu, jak np. [21, 22]. Do testowania parametrów kamer stosuje się także plansze testowe [1], które nie pozwalają jednak ocenić wszystkich istotnych parametrów.

Jednym z czynników, który ma decydujący wpływ na jakość obrazu jest szum, szczególnie dostrzegalny na obrazach zarejestrowanych podczas niedostatecznych warunków oświetleniowych. Zmniejsza on wyrazistość obrazu oraz poprawność odwzorowania koloru. Szum jest zjawiskiem fizycznym, całkowicie niemożliwym do wyeliminowania, pojawia się podczas rejestracji i przetwarzania obrazu. Zwykle jest traktowany jako dodatkowy sygnał uszkadzający obraz. Między innymi w [23] wymieniono następujące rodzaje szumów występujące dla przetworników CCD: śrutowy (ang. shot noise), resetowania (reset noise), wzmacniania (output amplifier noise/read noise), zegarowy (clocking noise), prądu ciemnego (dark current noise), związany z różną czułością poszczególnych komórek światłoczułych (photo response non-uniformity) i kwantyzacji (quantization noise).

Z kolei w przetwarzaniu obrazów do modelowania powyższych szumów wykorzystuje się dwa modele: addytywny, w którym przyjmuje się, że szum jest wprost dodawany do sygnału, oraz multiplikatywny, który zakłada, że wartości próbek sygnału są mnożone przez wartości próbek szumu [24, 25]. W przypadku obrazów rejestrowanych przez kamery najczę-

ściej rozważa się model addytywny, w którym uwzględnia się szумы: impulsowe, gaussowskie, mieszane – impulsowo-gaussowskie, białe, różowe ($1/f$) i inne.

Do oceny skuteczności odszumiania wykorzystuje się miary PSNR (ang. peak signal-to-noise ratio) [27] i MSE (ang. mean squared error) określające średnie różnice pomiędzy wartościami pikseli obrazu pierwotnego przed zaszumieniem i odpowiednimi wartościami pikseli wyliczonymi w obrazie odszumionym.

Zależnie od rodzaju szumu, zawartości obrazu i jego przeznaczenia, stosuje się różne metody odszumiania, które wciąż są intensywnie rozwijane. Odszumianie przeprowadza się zazwyczaj przed dalszą analizą obrazu. W przypadku kolejnych ujęć nieruchomych kadrów szумы usuwa się przez uśrednienie takich obrazów [24], jednak w przypadku systemów CCTV użyteczność tej metody jest ograniczona – w niniejszej rozprawie posłużyła do wyznaczenia obrazów referencyjnych używanych podczas tworzenia modeli szumów występujących w obrazach z matryc CMOS. Znacznie częściej szумы usuwa się poprzez filtrację pojedynczych obrazów. W [26] zaprezentowano podział metod odszumiania na operujące w dziedzinie przestrzennej i w dziedzinie transformat, które są bardziej złożone i nie są w tej pracy analizowane. Do dziedziny przestrzennej zalicza się metody uśredniające, bazujące na obliczaniu mediany i na filtracji Wienera. Wśród nich do najmniej złożonych (najszybszych) należą metody uśredniające, które dają znacznie gorsze rezultaty od metod wykorzystujących obliczanie mediany z wartości pikseli znajdujących się w pobliżu każdego piksela odszumianego obrazu. Metody medianowe oferują wyższą skuteczność, jednak wymagają sortowania pikseli, co jest bardziej złożone obliczeniowo. Dlatego, mając na uwadze konieczność uzyskania niskiej obliczeniowości, wymaganej w systemach przetwarzających obraz w czasie rzeczywistym, w niniejszej pracy rozwijane są metody uśredniające.

Drugim istotnym czynnikiem decydującym o jakości obrazu dla systemów CCTV jest jego zakres dynamiczny, a więc maksymalna różnica pomiędzy najjaśniejszymi i najciemniejszymi miejscami. Do jego wyrażenia stosuje się takie miary, jak współczynnik kontrastu czy stosunek sygnału do szumu. Do określania różnic w jasności rejestrowanej sceny stosuje się jednostkę EV (ang. exposure value), która jest równa 1 dla dwukrotnej wartości różnicy jasności. Współczesne zaawansowane kamery i aparaty fotograficzne są już w stanie zarejestrować cały zakres dynamiczny sceny widzianej przez człowieka (ok. 14 EV), jednak nie jest możliwe jego wyświetlenie na typowych ekranach, których zakres dynamiczny wynosi jedynie 8–10 EV. Ograniczeniem jest także zakres dynamiczny popularnych formatów plików komputerowych używanych do cyfrowego zapisu obrazu.

Aby umożliwić wyświetlenie obrazów o większej dynamice na typowych ekranach stosuje się techniki kompresji zakresu dynamicznego, popularnie nazywane HDRI (ang. high dy-

namie range imaging). W tym celu wykorzystuje się kilka kolejno naświetlonych obrazów z różnymi wartościami ekspozycji bądź jeden obraz o szerszym zakresie dynamiki, z którego uzyskuje się także obrazy zbyt jasne i zbyt ciemne (jak w tej rozprawie) i łączy je w jeden za pomocą technik mapowania tonalnego [28]. Techniki te wykorzystują operatory mapowania tonalnego do utworzenia mapy tonalnej. Mapa ta wskazuje, które obszary obrazów wejściowych, i w jakich proporcjach, należy użyć w celu otrzymania obrazu finalnego. Początki tych rozwiązań dla obrazów cyfrowych sięgają lat 80. XX w. Jednak dopiero w XXI w. zaproponowano wiele efektywnych operatorów, wśród których najważniejsze są opisane w pracy: Reinharda, Drago, Ashikhmina, Duranda, Fattala, Mantiuka i Pattanaika.

Techniki mapowania tonalnego są coraz powszechniej używane w przetwarzaniu obrazów, np. w systemach do detekcji znaków drogowych [29], do wielokierunkowej wizualizacji sceny [30], a nawet w przetwarzaniu obrazów sonarowych [31]. Do oceny jakości tak uzyskanych obrazów zaproponowano wiele technik psychofizycznych, parametrycznych, oceniających zmiany kontrastu i widoczność szczegółów. Jedną z miar jest JND (ang. just noticeable differences) [32]. Wiele z ww. technik wykorzystuje ocenę subiektywną dokonywaną przez widzów, lecz podjęto także próby stworzenia obiektywnych modeli oceny jakości obrazów mapowanych tonalnie.

W ostatnich latach popularność zyskało oprogramowanie do tworzenia obrazów mapowanych tonalnie (błędnie nazywanych HDR). Do najbardziej znanych programów należą dedykowane fotografom: Aurora HDR, EasyHDR, Photomatix Pro i Luminance HDR. Niektóre kamery CCTV, a nawet smartfony i aparaty fotograficzne mają także wbudowaną możliwość uzyskiwania obrazów o sztucznie powiększonym zakresie dynamicznym. Jednakże użyteczność powyższych rozwiązań w systemach CCTV jest ograniczona, gdyż ich celem jest uzyskanie obrazów o możliwie realistycznym lub efektownym wyglądzie. Tymczasem w systemach monitoringu jest wymagana jak największa szczegółowość obrazu i dlatego w niniejszej rozprawie zaproponowano metodę dedykowaną dla takich systemów.

4. Proponowana metoda redukcji szumu impulsowego i mieszanego szumu impulsowo-gaussowskiego

Autor zaproponował i zaimplementował w środowisku programistycznym Matlab iteracyjną metodę pracującą w zamkniętej pętli sprzężenia zwrotnego. Bazuje ona na filtracji uśredniającej i używa profili szumów, które zawierają informację o charakterystyce szumu występującego w obrazie. Profile te mogą być tworzone dla konkretnej kamery pracującej w określo-

nych warunkach oświetleniowych. Są tworzone automatycznie, a następnie wykorzystywane podczas odszumiania. W rozszerzonej wersji metody zaproponowano dobór profili do odszumiania w oparciu o opracowany sposób automatycznej detekcji stopnia zaszumienia odszumianych obrazów.

Profile szumu są tablicami progów detekcji, które klasyfikują piksele odszumianego obrazu monochromatycznego (w przypadku obrazu kolorowego analogicznie traktuje się poszczególne składowe w systemie RGB) jako niezaszumione albo zaszumione – w celu ich uśrednienia podczas odszumiania. Zarówno tworzenie profili jak i odszumianie jest procesem wymagającym kilku iteracji i . Do utworzenia profili szumu wykorzystuje się przykładowe obrazy zaszumione (oznaczone jako R) oraz niezaszumione (N). Obrazy N można uzyskać np. poprzez wielokrotne uśrednianie wielu statycznych zaszumionych obrazów R.

W celu wyznaczenia progów detekcji iteracyjnie dla każdego piksela $N_{x,y}$ (x,y – pionowe i poziome współrzędne piksela) oblicza się średnią $M_{x,y}$, i odchylenie przeciętne $D_{x,y,i}$ z sąsiednich pikseli znajdujących się w kwadratowej masce, której rozważany piksel jest środkiem. Nie uwzględnia się przy tym pikseli, które w poprzedniej iteracji zostały uznane za zaszumione (w pierwszej iteracji wszystkie piksele traktuje się jako niezaszumione). Następnie, na podstawie wartości $D_{x,y}$, piksele $N_{x,y}$ klasyfikuje się do k zbiorów $P_{k,i}$ na podstawie zależności

$$P_{k,i} = \left\{ N_{x,y} : D_{x,y,i} \in \left\langle \frac{k-1}{v}; \frac{k}{v} \right\rangle \right\}, \quad (1)$$

gdzie v służy do określenia szerokość przedziału wartości $D_{x,y,i}$. Optymalne progi detekcji $T_{k,i}$ tworzą profil szumu i uzyskuje się je na podstawie

$$T_{k,i} = \underset{T=0,0.1,0.2,\dots,100}{\operatorname{argmax}} \operatorname{card} \left\{ \left\{ N_{x,y} : \left\{ \begin{array}{l} N_{x,y} \in P_{k,i} \\ |N_{x,y} - M_{x,y,i}| > T \end{array} \right\} \cup \left\{ N_{x,y} : \left\{ \begin{array}{l} N_{x,y} \in P_{k,i} \\ |N_{x,y} - M_{x,y,i}| \leq T \\ |N_{x,y} - R_{x,y}| \leq U \end{array} \right\} \right\} \right\}, \quad (2)$$

gdzie T jest testowaną wartością progu, zaś U przyjętym parametrem pozwalającym traktować piksele $N_{x,y}$ zaszumione w niewielkim stopniu jako niezaszumione (istotne dla szumu gaussowskiego). Uzyskany próg $T_{k,i}$ pozwala sklasyfikować jako zaszumione jak największą część pikseli, które rzeczywiście były zaszumione oraz jako niezaszumione jak największą część pikseli, które rzeczywiście były niezaszumione (z uwzględnieniem parametru U).

Klasyfikacja pikseli $N_{x,y}$ dla następnej iteracji jako zaszumione ($E_{x,y,i+1}=0$) lub niezaszumione ($E_{x,y,i+1}=1$) odbywa się na podstawie zależności:

$$E_{x,y,i+1} = \begin{cases} 0 & \text{gdy } \begin{cases} |N_{x,y} - M_{x,y,i}| > T_{\lfloor v \cdot D_{x,y} \rfloor i} & \text{gdy } D_{x,y,i} < k_{\max} \cdot v \\ |N_{x,y} - M_{x,y,i}| > T_{1,i} & \text{gdy } D_{x,y,i} \geq k_{\max} \cdot v \end{cases} \\ 1 & \text{w przeciwnym razie} \end{cases} \quad (3)$$

gdzie k_{\max} jest łączną liczbą zbiorów $P_{k,i}$, a $\lfloor \cdot \rfloor$ oznacza zaokrąglenie w dół.

Odszumianie obrazów odbywa się z wykorzystaniem utworzonego profilu szumu oraz obliczanych wartości $M_{x,y,i}$ oraz $D_{x,y,i}$ w oparciu o zależność

$$E_{x,y,i+1} = \begin{cases} 0 & \text{gdy } \begin{cases} |N_{x,y} - M_{x,y,i}| > T_{\lfloor v \cdot D_{x,y} \rfloor i} & \text{gdy } D_{x,y,i} < k_{\max} \cdot v \\ |N_{x,y} - M_{x,y,i}| > T_{1,i} & \text{gdy } D_{x,y,i} \geq k_{\max} \cdot v \end{cases} \\ 1 & \text{w przeciwnym razie} \end{cases} \quad (4)$$

Pikselom $N_{x,y}$ odszumianego obrazu, dla których w ostatniej iteracji odszumiania $E_{x,y,i+1}=0$, zmienia się wartości na $M_{x,y,i}$, dzięki czemu odszumianie cechuje się niskim nakładem obliczeniowym.

Promień maski stosowanej do obliczania $M_{x,y,i}$ oraz $D_{x,y,i}$, początkowo równy 1, jest automatycznie powiększany, jeśli nie znajdują się w nim przynajmniej dwa piksele sklasyfikowane w poprzedniej iteracji jako niezaszumione. Poprawia to skuteczność odszumiania bardzo zaszumionych obrazów, gdy różnice wartości sąsiednich pikseli są znaczne.

W rozszerzonej wersji proponowana metoda dokonuje automatycznej detekcji poziomu zaszumienia obrazów na podstawie liczby pikseli sklasyfikowanych jako zaszumione. Dzięki temu podczas odszumiania jest dobierany profil szumu utworzony dla obrazów zaszumionych w podobnym stopniu, co dodatkowo podnosi skuteczność tego procesu.

Proponowana metoda odszumiania została przetestowana dla wybranych szumów typowo rozważanych w przetwarzaniu obrazów:

- impulsowego typu sól i pieprz (o minimalnych i maksymalnych wartościach mogących wystąpić w sygnale),
- impulsowego o zmiennych wartościach, które posiadają rozkład jednostajny dla całego zakresu wartości sygnału,
- mieszaniny addytywnego szumu gaussowskiego o zerowej wartości średniej i szumu impulsowego typu sól i pieprz,
- mieszaniny addytywnego szumu gaussowskiego o zerowej wartości średniej i szumu impulsowego o zmiennych wartościach,

Szumy te o różnych poziomach dodawano do standardowych obrazów testowych oraz do obrazów testowych przygotowanych przez autora. Obrazy testowe traktowano jako referen-

cyjne (R), natomiast po zaszumieniu – jako zaszumione (N). Na ich podstawie wyznaczano profile szumu. Do oceny skuteczności odszumiania zastosowano miarę PSNR.

W pracy przebadano wpływ doboru promienia maski stosowanej do obliczania $M_{x,y,i}$ oraz $D_{x,y,i}$ na skuteczność odszumiania. Zweryfikowano sposób wyznaczania i zakres wartości progów $T_{k,i}$ oraz liczbę zbiorów $P_{k,i}$ dla kolejnych iteracji. Wykazano, że w kolejnych iteracjach liczba pikseli poprawnie klasyfikowanych jako zaszumione bądź niezaszumione szybko rośnie, co potwierdza skuteczność zaproponowanego sposobu klasyfikacji pikseli.

Następnie dla każdego rodzaju szumu, zarówno w uproszczonej jak i w rozszerzonej wersji, wyznaczono minimalną liczbę iteracji odszumiania oraz dobrano odpowiednią wartość parametru U . Przebadano także, jak użycie profili szumu utworzonych dla jednych obrazów wpływa na skuteczność odszumiania innych obrazów oraz zaprezentowano wyniki odszumiania dla wybranych obrazów testowych.

Dla rozszerzonej wersji metody wyznaczono liczbę iteracji wstępnego odszumiania niezbędną do określenia stopnia zaszumienia obrazu przed wyborem odpowiedniego profilu do odszumiania. Ponadto zweryfikowano skuteczność wykorzystania profili utworzonych dla jednej grupy obrazów do odszumiania innych wyznaczając również optymalne wartości parametrów U oraz zaprezentowano uzyskane wyniki odszumiania na wycinkach obrazów testowych.

Do badań wykorzystano zbiory obrazów testowych zawierające zarówno standardowe obrazy powszechnie stosowane w przetwarzaniu obrazów przy porównywaniu podobnych metod oraz obrazy przygotowane przez autora, pochodzące z przetworników stosowanych w kamerach systemów monitoringu wizyjnego.

W przypadku wersji metody dedykowanej do odszumiania rzeczywistych obrazów z systemów CCTV także określono optymalne wartości parametrów U oraz liczbę iteracji korzystając z odrębnego zbioru obrazów testowych o 14-bitowej głębi koloru, zarejestrowanych z różnym poziomem zaszumienia (różna czułość ISO). Zaszumione obrazy N uśredniano (przebadano 6 sposobów uśredniania) by otrzymać referencyjne obrazy R, które podczas tworzenia profili szumu normalizowano z powodu występującego w obrazach N szumu niskoczęstotliwościowego. Profile szumu wyznaczano osobno dla każdej składowej kolorystycznej RGB, ponieważ stwierdzono zróżnicowanie poziomu ich zaszumienia.

Ze względu na zróżnicowany zakres dynamiczny obrazów testowych, do oceny skuteczności odszumiania autor wykorzystał MSE (ang. mean squared error) zamiast PSNR. Za jego pomocą sprawdzono efektywność wykorzystania profili szumu utworzonych dla jednych obrazów do odszumiania innych. Rezultaty odszumiania także zaprezentowano na wycinkach przetworzonych obrazów.

Wyniki eksperymentów dla proponowanej metody odszumiania dowodzą, że sposób wyznaczania progów detekcji $T_{k,i}$ jest właściwy. Uzyskiwane już w kilku iteracjach progi pozwalają z niewielkim błędem klasyfikować piksele jako zaszumione bądź niezaszumione. Prezentowane w pracy wykresy dowodzą, że podczas odszumiania wykryte zaszumione piksele skutecznie zastępuje się uśrednionymi wartościami. Co ważne, do wyliczenia tych średnich nie używa się pikseli sklasyfikowanych jako zaszumione. Dzięki temu, pomimo braku wykorzystania zaawansowanych metod estymacji obrazu pierwotnego (co znacznie wydłużałoby czas przetwarzania), uzyskany obraz cechuje się dużą szczegółowością uzyskiwaną dla metod bazujących np. na filtracji medianowej.

Autor udowodnił, że profile szumu uzyskane dla jednych obrazów mogą posłużyć do odszumiania innych przy zachowaniu porównywalnej skuteczności wyrażonej miarą PSNR, szczególnie w przypadku szumów mieszanych. W rozszerzonej wersji proponowaną metodę cechuje jeszcze wyższa skuteczność, gdy wykrywany jest poziom zaszumienia i są dobierane odpowiednie profile szumu. Wykazano, że rezultaty odszumiania zależą od rodzaju i stopnia zaszumienia obrazów. Obrazy zaszumione w niewielkim stopniu szumami impulsowymi (do około 30% pikseli o zmienionych wartościach) wymagają zastosowania około 2–3 iteracji, bardziej zaszumione wymagają około 7 iteracji. W przypadku szumów gaussowskich wystarczają maksymalnie 2–3 iteracje. Co ważne, skuteczność odszumiania w większości przypadków nie zmniejsza się po wykonaniu większej liczby iteracji niż wymagana. W niektórych przypadkach obserwuje się nieznaczny spadek jakości.

Uzyskane wartości PSNR dla standardowych obrazów testowych porównano z wynikami opisanymi w literaturze dla konkurencyjnych metod. W przypadku szumów impulsowych rezultaty uzyskane dla zaproponowanej metody w około połowie przypadków są najlepsze, a w pozostałych przypadkach są do nich zbliżone. Dla niektórych obrazów testowych uzyskano PSNR wyższy o 1,5 dB lub więcej. Przy usuwaniu szumów mieszanych proponowaną metodę wyprzedzają jedynie bardziej złożone obliczeniowo rozwiązania, wymagające często nawet 20 iteracji.

W przypadku odszumiania rzeczywistych obrazów z systemów CCTV okazało się, że wystarczy jedna iteracja odszumiania dla osiągnięcia najlepszych rezultatów. Uzyskane wyniki odszumiania porównano z wynikami uzyskanymi dla filtrów medianowych i uśredniających. Uzyskane wartości MSE dowodzą, że z wyjątkiem najbardziej zaszumionych obrazów, proponowana metoda daje zdecydowanie lepsze rezultaty. Potwierdzają to także zaprezentowane wycinki przetworzonych obrazów.

5. Proponowana metoda lokalnego poszerzania zakresu dynamicznego

Proponowana metoda (także zaimplementowana w środowisku programistycznym Matlab) automatycznie poprawia widoczność detali, zwłaszcza w miejscach zbyt jasnych (prześwielonych) i zbyt ciemnych (niedoświetlonych) poprzez redukcję całkowitego zakresu dynamicznego zarejestrowanego obrazu do zakresu możliwego do wyświetlenia na standardowych ekranach. Metoda ta bazuje na informacji o jasności poszczególnych fragmentów obrazu, różnicach pomiędzy wartościami sąsiednich pikseli oraz na detekcji krawędzi zaproponowanym sposobem.

Przedstawiona metoda przetwarza obrazy o większym zakresie dynamicznym niż jest możliwy do wyświetlenia na standardowym ekranie. Typowo są to obrazy o głębi bitowej większej niż 8 bitów na każdą składową kolorystyczną rejestrowaną przez przetwornik kamery. Z takiego pojedynczego obrazu uzyskuje się dwa obrazy: niedoświetlony (DK), który zawiera znacznie więcej szczegółów w miejscach prześwielonych niż obraz nieprzetworzony, oraz prześwielony (BR), w którym z kolei są dobrze widoczne szczegóły w miejscach niedoświetlonych. W pięcioetapowym procesie takie obrazy łączy się w jeden obraz wynikowy (F), który zawiera możliwie dużo szczegółów na całej powierzchni.

W pierwszym etapie oba obrazy wejściowe wyostrza się, aby zwiększyć ich szczegółowość – uzyskuje się obrazy DKS i BRS. Te obrazy następnie się uśrednia, aby uzyskać obraz o optymalnym naświetleniu – IMS, z którego będą wykorzystane fragmenty poprawnie naświetlone.

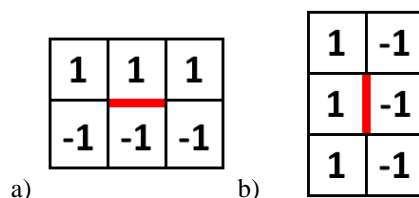
Drugi etap obejmuje analizę jasności obrazów DK i BR. Konwertuje się je do skali szarości uzyskując, odpowiednio, obrazy DKG i BRG. W wyniku progowania tych obrazów powstają mapy obszarów obrazu DKS, które mają dużą jasność i w finalnym obrazie F będą wykorzystane. Analogicznie obraz F będzie zawierał ciemne obszary obrazu BRS, które miały małe wartości w obrazie BRG. W celu wykrycia obszarów, które będą pochodzić z obrazu IMS, oblicza się obraz różnicowy obrazów DKG oraz BRG i również go proguje. W obrazie F wykorzystuje się te jego obszary, które mają największe jasności, a więc zostały prawidłowo naświetlone.

W trzecim etapie generowana jest wstępna mapa tonalna o wymiarach (w pikselach) równych wymiarom przetwarzanych obrazów. Mapa ta zawiera informację o obszarach, które w finalnym obrazie pochodzą z obrazów DKS, IMS i BRS. Ponieważ te obszary nie obejmują całej powierzchni mapy, więc dla pozostałych jej części analizuje się wielkości bezwzględnych różnic pomiędzy sąsiednimi pikselami obrazów DKG i BRG (sumowane w każdym z tych obrazów w pionie i w poziomie). Większe wartości tych różnic w obrazie DKG wska-

zują, że dana część obrazu jest bardziej kontrastowa, a więc zawiera więcej szczegółów i w obrazie F powinna pochodzić z obrazu DKS. Analogicznie większe wartości różnic wyznaczone dla obrazu BRG oznaczają, że dana część obrazu będzie pochodzić z obrazu BRS. Równe wartości wskazują na wykorzystanie obrazu IMS.

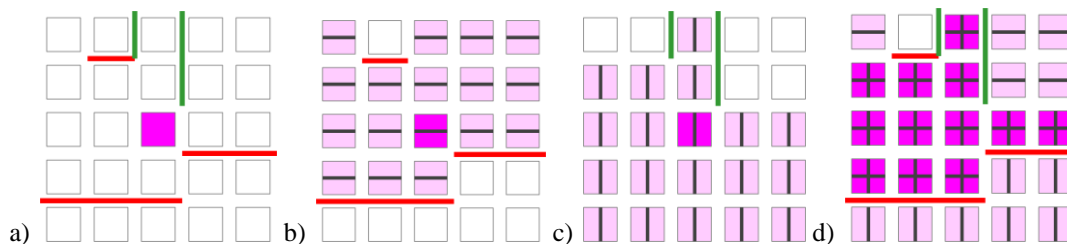
Czwarty etap obejmuje przetwarzanie uzyskanej mapy tonalnej, która musi zostać uśredniona, aby w finalnym obrazie nie powstały pseudo-krawędzie. Pseudo-krawędzie byłyby widoczne na granicach obszarów, które będą pochodzić z obrazów DKS, IMS i BRS, jeżeli w tych miejscach nie występują krawędzie obiektów. Z kolei uśrednienie mapy tonalnej wokół krawędzi obiektów pochodzących z różnych obrazów spowodowałoby zauważalne prze-regulowania, które są charakterystyczne dla niektórych metod mapowania tonalnego. W związku z tym zastosowano uśrednianie mapy tonalnej połączone z detekcją krawędzi.

Uśrednia się tylko wartości (piksele) mapy wewnątrz przyjętej maski, jeśli znajdują się pomiędzy wykrytymi krawędziami. Standardowe detektory krawędzi, takie jak np. Prewitt, Sobela czy Canny'ego przypisują krawędzie wykryte pomiędzy pikselami do sąsiednich pikseli, więc nie wszystkie wartości mapy byłyby prawidłowo wybrane do uśredniania. W związku z tym autor zaproponował zmodyfikowany detektor Prewitt, który zachowuje informację o wykrytych krawędziach pomiędzy pikselami w dwóch dodatkowych tablicach, odpowiednio dla krawędzi poziomych i pionowych. Opracowany detektor wykorzystuje maski przedstawione na rys. 1. Pogrubioną czerwoną linią oznaczono wykrywane krawędzie pomiędzy pikselami.



Rys. 1. Proponowany zmodyfikowany operator Prewitt: a) poziomy, b) pionowy

Ponadto autor opracował efektywną metodę uśredniania mapy tonalnej wykorzystującą tylko jej wartości znajdujące się wewnątrz przyjętej maski pomiędzy wykrytymi krawędziami. Ideę tego rozwiązania przedstawiono na rys. 2.



Rys. 2. Wybór pikseli mapy tonalnej do uśredniania: a) wykryte krawędzie pomiędzy pikselami mapy tonalnej, b) piksele pomiędzy krawędziami poziomymi, c) piksele pomiędzy krawędziami pionowymi, d) piksele wybrane do uśrednienia (oznaczone krzyżami)

Dla systemów, w których ilość obliczeń jest szczególnie istotna, autor opracował także uproszczoną wersję uśredniania mapy tonalnej, która wykorzystuje standardowy detektor Prewitt oraz trzykrotnie uśrednianą (w różnym stopniu) wstępną mapę tonalną.

Ostatni, piąty etap, obejmuje generację finalnego obrazu F. W tym celu jest stosowana średnia arytmetyczna ważona, której wagami są wartości uzyskanej mapy tonalnej. Wagi te wskazują, w jakich proporcjach wartości składowych RGB w finalnym obrazie wykorzystują wartości tych składowych w obrazach DKS, IMS i BRS.

Efektywność proponowanej metody lokalnego poszerzania zakresu dynamicznego zweryfikowano przy użyciu zbioru siedmiu testowych obrazów prezentujących sytuacje typowo występujące w systemach monitoringu wizyjnego nadzorowanego przez człowieka (ze szczególnym naciskiem na monitoring miejski) oraz w systemach, gdzie obraz jest analizowany automatycznie. Obrazy te przetworzono przy użyciu proponowanej metody (zarówno w pełnej jak i w uproszczonej wersji) oraz przy użyciu wybranych popularnych operatorów mapowania tonalnego (Ashikhmina, Drago, Mantiuka i Pattanaika) zaimplementowanych w oprogramowaniu Luminance HDR. Do analizy wykorzystano też obrazy nieprzetworzone oraz o obniżonym kontraście.

W pierwszej grupie eksperymentów uczestniczyło 36 respondentów, którzy oceniali istotne dla operatorów monitoringu parametry obrazu, takie jak: ogólna jakość, widoczność szczegółów, realizm i naturalność. Oprócz tego ich zadaniem było m.in. liczenie ludzi w kadrze (gdy nie wszyscy są wyraźnie widoczni), odczytywanie treści słabo widocznych napisów, wskazywanie i liczenie podejrzanych pakunków, określanie kolorów pojazdów i ubrań wskazanych osób. Testy te były przeprowadzone zarówno przy określonym czasie wyświetlania poszczególnych zdjęć, wynoszącym 3 lub 6 sekund (dla każdego obrazu) oraz bez limitu czasu.

Eksperymenty przeprowadzono w warunkach laboratoryjnych z użyciem skalibrowanego kolorystycznie ekranu [33]. Do oceny parametrów obrazu wykorzystano powszechnie stosowaną pięciostopniową skalę subiektywnej oceny MOS (ang. mean opinion score) [34]. Każdy respondent oceniał 6 obrazów przygotowanych w 8 wersjach i wyświetlanych w przypadkowej kolejności.

Druga grupa eksperymentów dotyczyła automatycznego rozpoznawania danych w obrazach. Zweryfikowano zdolność systemu OCR zaimplementowanego w oprogramowaniu FreeOCR do rozpoznania treści na pięciu wycinkach obrazów testowych. Sprawdzone także wpływ operacji mapowania tonalnego na skuteczność detekcji twarzy w obrazie wykorzystując oprogramowanie Picasa.

Wyniki eksperymentów z udziałem respondentów wskazują, że przy ograniczonym czasie prezentacji obrazów metoda lokalnego poszerzania zakresu dynamicznego pozwala uzyskać tylko nieco gorsze rezultaty niż operator Pattanaika (który jednak działa znacznie wolniej), ale lepsze niż pozostałe metody. W testach bez limitowanego czasu wyświetlania zdjęć okazało się, że proponowana metoda umożliwiła najdokładniejsze określenie kolorów oraz odczytanie napisów w słabo widocznych fragmentach zdjęć. W zakresie poprawy widoczności miejsc niedoświetlonych (w głębokim cieniu) oraz prześwietlonych (jak np. widoczność oznakowania poziomego na drodze) okazała się bezkonkurencyjna w odniesieniu do wszystkich pozostałych metod.

Pomimo że celem opracowanej metody nie jest uzyskanie zdjęć o najlepszej jakości, okazało się, że przetworzone przy jej użyciu obrazy testowe zostały najlepiej ocenione w tej kategorii, uzyskując średnie oceny 3,50 oraz 3,48 w skali MOS (odpowiednio dla pełnej i uproszczonej wersji) w porównaniu do 3,39 dla obrazów o obniżonym kontraście, które zajęły następne miejsce. Jeszcze korzystniej wypadła ocena szczegółowości zdjęć – analogicznie uzyskano wartości 3,80 i 3,69 wobec 3,15 dla operatora Pattanaika, który był następny w tym zestawieniu. Jedynie naturalność uzyskanych obrazów została oceniona gorzej niż obrazów nieprzetworzonych, o obniżonym kontraście oraz przetworzonych operatorami Drago i Mantiuka. Autor jednak zwrócił uwagę, że celem prowadzonych badań było uzyskanie metody pozwalającej uzyskać przede wszystkim obrazy o jak największej szczegółowości nawet za cenę pewnych zniekształceń, które nie utrudniają percepcji tych obrazów. Pomimo tego, nawet w uśrednionym zestawieniu trzech powyższych cech, proponowana metoda uzyskuje najlepsze oceny – odpowiednio 3,41 i 3,37 względem 3,24 dla obrazów o obniżonym kontraście, które uplasowały się na kolejnym miejscu.

Treści znajdujące się na obrazach przetworzonych proponowaną metodą (w obu wersjach) jako jedyne były możliwe w całości do odczytania przez oprogramowanie do OCR – łączna liczba błędów odczytu była dla nich najmniejsza. W przypadku detekcji twarzy tylko na obrazach przetworzonych przy użyciu operatorów Drago, Mantiuka i proponowanego (w obu wersjach) było możliwe jednoczesne rozpoznanie wszystkich twarzy. Ponadto twarze na obrazie przetworzonym proponowanym rozwiązaniem zawierają najwięcej szczegółów, które są istotne np. przy rozpoznawaniu osób.

Pomiary czasu przetwarzania wskazują, że proponowana metoda, w wersji uproszczonej, może być bez przeszkód zaimplementowana w systemach czasu rzeczywistego. Ponadto obie wersje tej metody mogą być efektywnie zastosowane w systemach przetwarzania wieloprocesorowego.

6. Wnioski

Na podstawie eksperymentalnych wyników opisanych w rozprawie można wysnuć poniższe wnioski:

1. Istnieje możliwość poprawy jakości obrazu w systemach CCTV poprzez zastosowanie dodatkowego przetwarzania obrazu dzięki metodom opracowanym przez autora.
2. Odszumianie obrazów może wykorzystywać filtrację uśredniającą zamiast bardziej złożonych metod filtracji bazujących na medianie.
3. Proponowana metoda odszumiania skutecznie redukuje szum o charakterze impulsowym oraz szum będący mieszaniną szumu gaussowskiego i szumu impulsowego dając porównywalne lub lepsze wyniki niż inne, bardziej skomplikowane rozwiązania.
4. Tworzenie profili szumu wydatnie upraszcza odszumianie, a profile szumu utworzone dla zaszumionych obrazów referencyjnych mogą być z powodzeniem wykorzystywane do odszumiania innych obrazów.
5. Mapowanie tonalne może wykorzystywać jednoczesną analizę jasności obrazu, różnic wartości pomiędzy pikselami i wykrywanie krawędzi w celu uzyskania finalnych obrazów o dużej szczegółowości.
6. Zaproponowany detektor krawędzi wykrywający krawędzie pomiędzy sąsiednimi pikselami zapobiega w finalnym obrazie przeregulowaniom, które są zauważalne dla powszechnie stosowanych detektorów.
7. Obrazy przetworzone zaproponowaną metodą lokalnego powiększania zakresu dynamicznego są lepszej jakości i o większej szczegółowości, dzięki czemu mogą być bardziej przydatne w systemach CCTV niż obrazy nieprzetworzone bądź przetworzone innymi porównanymi metodami.

W związku z powyższym cel naukowy, którym był rozwój metod cyfrowego przetwarzania obrazów w zakresie poprawy ich jakości i szczegółowości pod kątem zastosowań w systemach CCTV został zrealizowany, a teza naukowa została udowodniona.

Bibliografia

- [1] V. Damjanovski, „CCTV, Networking and Digital Technology”, Butterworth-Heinemann, 2014 and 2005 edition.

- [2] C. Reuter, „The V2 and the German, Russian and American Rocket Program”, S.R. Research and Publishing, New York 2000.
- [3] C. Norris, M. McCahill and D. Wood, „Editorial. The Growth of CCTV: a global perspective on the international diffusion of video surveillance in publicly accessible space”, *Surveillance & Society, CCTV Special*, 2(2/3), pp. 110–135, 2004.
- [4] Fundacja Panoptykon, „Monitoring w polskich miastach i w oczach społeczeństwa”, [Online]. Zamieszczone: https://panoptykon.org/sites/default/files/panoptykon_cctv_seminarium_10-10-2012.pdf [Dostęp 28.11.2016].
- [5] D. Bartoszewicz, „Wielki Brat w Warszawie. Działa już 13,5 tysiąca kamer”, [Online]. Zamieszczone: <http://warszawa.wyborcza.pl/warszawa/1,34862,21000172,wielki-brat-w-warszawie-dziala-juz-13-5-tysiaca-kamer-komentarz.html> [Dostęp 11.11.2016].
- [6] R. Chellappa, A.K. Roy-Chowdhury and K. Zhou, „Recognition of Humans and Their Activities Using Video”, Morgan & Claypool, USA, 2005.
- [7] M. Ramanathan, W. Y. Yau and E. K. Teoh, „Human Action Recognition with Video Data: Research and Evaluation Challenges”, *IEEE Trans. on Human-Machine Systems*, vol. 44, no. 5, pp. 650–663, 2014.
- [8] A.C. Davies and S.A. Velastin, “A Progress Review of Intelligent CCTV Surveillance Systems”, *2005 IEEE Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications*, IDAACS, Sofia, 2005, pp. 417–423.
- [9] T. Shan, S. Chen, C. Sanderson and B.C. Lovell, „Towards robust face recognition for Intelligent-CCTV based surveillance using one gallery image”, *2007 IEEE Conference on Advanced Video and Signal Based Surveillance*, AVSS, London, 2007, pp. 470–475.
- [10] A. Bigdeli, B.C. Lovell, C. Sanderson, T. Shan and S. Chen, „Vision Processing in Intelligent CCTV for Mass Transport Security”, *2007 IEEE Workshop on Signal Processing Applications for Public Security and Forensics*, Washington, DC, USA, 2007, pp. 1–4.
- [11] D.M. Considine and G.D. Considine, „Standard handbook of industrial automation”, Chapman and Hall Ltd, London, 1986.
- [12] D. Ito, „Robot vision: strategies, algorithms and motion planning”, Nova Science, New York, 2009.
- [13] FLIR, „Noc staje się dniem”, [Online]. Zamieszczone: http://www.flir-termowizja.pl/files/Broszura_FLIR_2015_LincPolska_12052015.pdf [Dostęp 28.11.2016].
- [14] Bosch, „Camera Reference Guide”, [Online]. Zamieszczone: http://resource.boschsecurity.com/documents/Selection_Guide_enUS_1999774347.pdf [Dostęp 28.11.2016].

- [15] N.G. La Vigne, S.S. Lowry, A.M. Dwyer and J.A. Markman, „Using Public Surveillance Systems for Crime Control and Prevention: A Practical Guide for Law Enforcement and Their Municipal Partners”, Community Oriented Policing Services U.S. Department of Justice, [Online]. Zamieszczone: <http://www.urban.org/sites/default/files/alfresco/publication-pdfs/412402-Using-Public-Surveillance-Systems-for-Crime-Control-and-Prevention-A-Practical-Guide-for-Law-Enforcement-and-Their-Municipal-Partners.PDF> [Dostęp 28.11.2016].
- [16] Najwyższa Izba Kontroli, „Funkcjonowanie miejskiego monitoringu wizyjnego”, 2013, [Online]. Zamieszczone: <https://www.nik.gov.pl/plik/id,6400,vp,8169.pdf> [Dostęp 28.11.2016].
- [17] P. Sen and C. Aguerrebere, „Practical High Dynamic Range Imaging of Everyday Scenes: Photographing the world as we see it with our own eyes”, *IEEE Signal Processing Magazine*, vol. 33, no. 5, pp. 36–44, Sept. 2016.
- [18] Basler, „Basler IP Cameras. Security Cameras Overview”, [Online]. Zamieszczone: <https://www.sourcesecurity.com/datasheets/basler-bip2-1000c-ip-camera/co-3637-ga/basler-bip2-1000c-ip-camera.pdf> [Dostęp 28.11.2016].
- [19] N. Cohen, J. Gattuso and K. MacLennan-Brown, „CCTV Operational Requirements Manual. 2009”, [Online]. Zamieszczone: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/378443/28_09_CCTV_OR_Manual2835.pdf [Dostęp 28.11.2016].
- [20] Solaris Bus & Coach S.A., „Instrukcja obsługi dla motorniczych tramwaju Solaris Tramino S105p”, 2012.
- [21] D.W. Hertel and E. Chang, „Image Quality Standards in Automotive Vision Applications”, *2007 IEEE Intelligent Vehicles Symposium*, Istanbul, 2007, pp. 404–409.
- [22] M. Saad, A.C. Bovik and C. Charrier, „Blind image quality assessment: A natural scene statistics approach in the DCT domain”, *IEEE Trans. Image Process.*, vol. 21, no. 8, pp. 3339–3352, Aug. 2012.
- [23] „CCD Image Sensor Noise Sources”, On Semiconductor, AND9189/D, Jan. 2016 – Rev. 3.
- [24] A. Bovik, „The Essential Guide to Image Processing”, Elsevier Academic Press, 2009.
- [25] A. Nath, „Image Denoising Algorithms: A Comparative Study of Different Filtration Approaches Used in Image Restoration”, *2013 International Conference on Communication Systems and Network Technologies*, Gwalior, 2013, pp. 157–163.

- [26] M.C. Motwani, M.C. Gadiya, R.C. Motwani and F.C. Harris, „Survey of image denoising techniques”, *Proceedings of Global Signal Processing Expo and Conference (GSPx '04)*, Santa Clara, Calif, USA, Sept. 2004.
- [27] Z. Wang and A.C. Bovik, „Mean squared error: Love it or leave it? A new look at Signal Fidelity Measures”, *IEEE Signal Processing Magazine*, vol. 26, no. 1, pp. 98–117, Jan. 2009.
- [28] E. Reinhard, S. Pattanaik, G. Ward, P. Debevec, „High dynamic range imaging. Acquisition, display, and image-based lighting”, The Morgan Kaufmann Publisher, 2005.
- [29] A. Vavilin, K. Deb, T.H. Kim and K.H. Jo, „Road sign detection method based on fast HDR image generation technique”, *2010 25th International Conference of Image and Vision Computing*, New Zealand, Queenstown, 2010, pp. 1–8.
- [30] K. Hirai, N. Osawa, T. Horiuchi and S. Tominaga, „An HDR Spectral Imaging System for Time-Varying Omnidirectional Scene”, *2014 22nd International Conference on Pattern Recognition*, Stockholm, 2014, pp. 2059–2064.
- [31] J. Duda, „Zastosowanie techniki HDR do obrazów pozyskiwanych z sektorowego sonaru skanującego”, *Biuletyn WAT*, vol. LX, no. 3, 2011.
- [32] K. Myszkowski, R. Mantiuk and G. Krawczyk, „High Dynamic Range Video”, Morgan & Claypool, 2008.
- [33] x-rite Pantone, „Poor Results With ColorMunki Display On PC”, [Online]. Zamieszczone: http://www.xrite.com/service-support/Poor_Results_With_ColorMunki_Display_On_PC [Dostęp 28.11.2016].
- [34] International Telecommunication Union, „ITU-T Recommendation P.911. Series P: Telephone transmission quality, telephone installations, local line networks. Audiovisual quality in multimedia services. Subjective audiovisual quality assessment methods for multimedia applications”, [Online]. Zamieszczone: <https://www.itu.int/rec/T-REC-P.911-199812-I/en> [Dostęp 28.11.2016].

W wykazie literatury wymieniono tylko te pozycje, które zostały zacytowane w tym streszczeniu.