Usuwanie tła w wideo nagraniach pochodzących z monitorowania basenu pływackiego

Konrad Reiter, Zdzisław Kowalczuk
Politechnika Gdańska, Wydział Elektroniki, Telekomunikacji i Informatyki, Katedra Systemów Decyzyjnych i Robotyki, Narutowicza 11/12, 80-233 Gdańsk

Streszczenie: Automatyczna obróbka obrazu w czasie rzeczywistym jest kluczowa dla wielu rozwiązań monitoringu wykorzystywanych m.in. w celach bezpieczeństwa. Często jednym z ważniejszych etapów obróbki jest oddzielenie tła od obiektów na pierwszym planie, tak aby wykluczyć wszystkie nieistotne informacje z obrazu. Celem pracy jest podsumowanie doświadczenia zdobytego podczas śledzenia pływaków oraz pokazanie możliwości skutecznego automatycznego nadzoru wideo osób korzystających z basenu. Porównano skuteczność działania dwóch wybranych algorytmów (MOG i KNN) przy użyciu różnych odwzorowań kolorów oraz omówiono zalety i wady analizowanych metod.

Słowa kluczowe: monitoring, baseny pływackie, przetwarzanie obrazów, usuwanie tła, eliminacja szumów, OpenCV

1. Wprowadzenie

Liczba obiektów objętych monitoringiem ciągle się powiększa. Obecnie w większości sklepów, szkół i innych obiektów użyteczności publicznej działają systemy kamer. Najczęstszym powodem ich instalacji jest chęć poprawy bezpieczeństwa użytkowania publicznych przestrzeni. Oczywiście w takich warunkach pojawiają się też alternatywne zastosowania zarejestrowanych nagrań, takie jak np. zliczanie przychodzących klientów czy pomiary potrzebne do wyznaczania statystyk.

Śledzenie obiektów jest w ogólności zadaniem bardzo trudnym, gdyż wskutek zmiany ułożenia lub swoich możliwości funkcjonalnych, obiekty mogą wykonywać gwałtowne ruchy, zmieniać swoją postać lub strukturę albo konfigurację między obiektami, mogą pojawiać się okluzje lub inne zmiany w obserwowanej scenie albo też niepożądane ruchy śledzącej kamery.

W niniejszej pracy analizujemy problem przetwarzania nagrań wideo zarejestrowanych w przestrzeni basenowej, umożliwiających monitoring oraz zbieranie danych statystycznych związanych z pływakami, używającymi publicznego basenu. Dla skupienia uwagi, zagadnienie to sprowadzamy do 'prostego' zadania śledzenia obiektów ruchomych w środowisku wodnym basenu.

W literaturze istnieje wiele sposobów obserwacji obiektów z wykorzystaniem nagrań wideo, opartych na uniwersalnych obserwatorach lub filtracji Kalmama, albo innych specyficznych podejść, takich jak *meanshift*, *camshift*, czy *optical flow* [10]. Znane są też rozmaite procedury śledzenia obiektów ruchomych oparte na metodach sztucznej inteligencji, np. na sieciach neuronowych [3, 11], które wymagają jednak przeprowadzenia treningu na odpowiedniej próbie uczącej.

Jako metodę rozwiązania problemu śledzenia, rozważymy tu procedurę opartą na segmentacji obrazu, polegającą na wyizolowaniu wszystkich znaczących obiektów znajdujących się w obrazie wideo, analizowanym po usunięciu wodnego tła. Implementowana w ramach przetwarzania obrazów [4] segmentacja wykonywana na nietrywialnych danych wideo jest bardzo trudnym zagadnieniem. Jako ważny etap takiej operacji zwykle proponuje się usuwanie, bądź odcinanie tła (ang. *background subtracting* lub *background removal*), które stosuje się już od ponad 30 lat [1, 9]. Najczęściej rozważa się systemy dotyczące obiektów, które są względnie niezmienne (stacjonarne), z prostym tłem. W przypadku ulicznego monitoringu mogą pojawić się poruszające się liście lub gałęzie drzew oraz rozmaite cienie. W przypadku basenu pływackiego, na powierzchni (tafli) wody może pojawić się typowe zafalowanie, wywołane przez pływaków. Im algorytm będzie bardziej odporny na tego typu zaburzenia, tym łatwiejsza będzie analiza obrazu i obliczanie wymaganych statystyk.

Praca jest zorganizowana w następujący sposób. Sekcja 2 wyjaśnia, jaki problem przetwarzania jest analizowany oraz czego oczekuje się od syntezowanego algorytmu. W sekcji 3 omawia się kwestie związane z ustawieniem i doborem kamery. Sekcja 4 poświęcona jest wyjaśnieniu działania różnych wariantów algorytmów odcinania tła. W sekcji 5 prezentuje się wyniki zrealizowanych eksperymentów. Sekcja 6 kończy raport, formułując podsumowanie i wnioski projektowe.

2. Analiza zadania

Przedmiotem pracy jest analiza rozwiązań stosowanych do usuwania tła, w celu dostarczenia materiału odpowiedniego do śledzenia pływaków, znajdujących się w obszarze basenu,

Autor korespondujący:
Zdzisław Kowalczuk, kova@pg.gda.pl

Artykuł recenzowany
nadesłany 08.12.2017 r., przyjęty do druku 28.09.2018 r.

Zezwala się na korzystanie z artykułu na warunkach licencji Creative Commons Uznanie autorstwa 3.0

w postaci maski z zaznaczonymi osobami (bez punktów tła). Warto dodać, że w kolejnych krokach analizy należy wykryć kontury oraz śledzić ich środki, na przykład za pomocą filtru Kalmana [7], co jednak nie jest już przedmiotem niniejszego artykułu.

Obiekt – jakim jest basen pływacki – ma specyfikę, którą warto uwzględnić w proponowanym rozwiązaniu, a która nie występuje w takim stopniu w innych popularnych zastosowaniach (np. przy monitoringu ulic). Po pierwsze należy dołożyć starań, aby śledzić każdego pływaka, nawet jeśli znajduje się on pod powierzchnią wody – tak aby nie zgubić żadnego śladu. Po drugie należy uwzględnić fakt, że tafla wody zachowuje się jak duże lustro. Występują bowiem na niej małe refleksy świetlne pochodzące ze sztucznego oświetlenia oraz duże odbicia światła dziennego, co powoduje, że nawet cała szerokość toru na dłuższym odcinku (nawet do kilku metrów) może być niewidoczna. Na delikatnie falującej wodzie refleksy świetlne się przemieszczają i rozdzielają. Jeśli zaś woda zostanie mocno zaburzona, odbicia znikają kompletnie – co również może być kłopotliwe, szczególnie w przypadku dużych refleksów świetlnych.

Ponadto obiekt basenowy ma też cechy, które można wykorzystać. Jednym z nich jest relatywnie korzystne oświetlenie, gdyż w czasie pracy basenu zwykle włączone jest sztuczne światło (które zmniejsza niebezpieczeństwo gwałtownej zmiany siły oświetlenia). Inną ważną zaletą tego środowiska jest względna stałość barw tła (mimo możliwych ruchów wody i refleksów).

3. Sprawy sprzętowe

Wybór miejsca, z którego prowadzony jest monitoring, wpływa na sposób prowadzenia analizy oraz minimalizacji wpływu refleksów świetlnych na dalszą analizę. Rozważmy wykorzystanie jednej kamery. Zauważmy, że w przypadku korzystania z kilu kamer dochodzą nowe problemy, ale podstawowe wnioski z systemu mono wizyjnego będą stosować się także do systemów z wieloma kamerami.

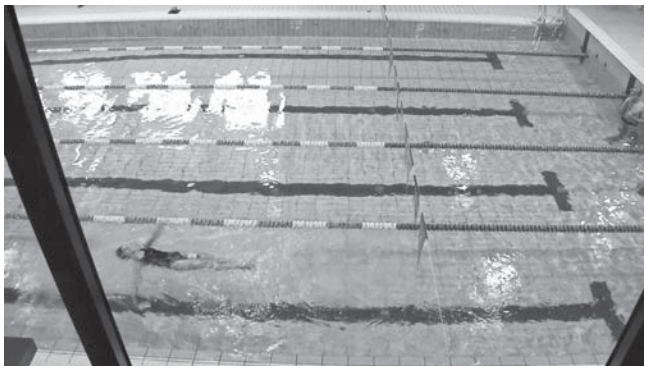
3.1. Penetracja wody

Podstawowym zadaniem monitoringu jest uchwycenie tego, co jest najważniejsze w przestrzeni basenu. Aby kamera mogła rejestrować osoby nurkujące pod powierzchnię wody, należy zapewnić, aby kąt patrzenia na wodę był jak najbliżej pionu [8]. Dzięki temu, nawet głęboko nurkujący pływak zostanie zarejestrowany na nagraniu i algorytm śledzenia go nie 'zgubi'. Oczywiście pożądane jest też, aby kamera zawieszona była na środku dłuższego boku basenu – tak aby można było zmieścić obszar całej pływalni w jednym ujęciu. Jeśli basen jest zbyt niski, lub z innego powodu nie można uchwycić całego obszaru, należy rozważyć użycie obiektywu szerokokątnego albo skłonić się do zastosowania kilku kamer.

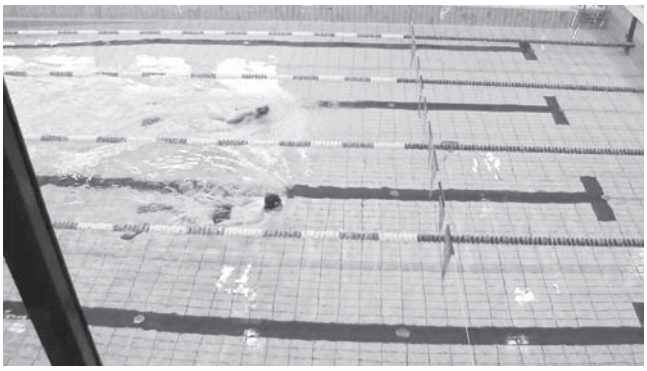
3.2. Refleksy słoneczne

Nieodpowiednie ustawienie kamery może spowodować, że na wodzie pojawią się refleksy świetlne. Na obiektach pływackich często duża część ściany jest przeszklona, przez co do wnętrza pomieszczenia wpada światło słoneczne. Jeśli kamera zostanie ustawiona dokładnie na przeciw okien, istnieje duża szansa, że część wody będzie odbijać światło słoneczne wprost do kamery (przy spokojnej wodzie). Ilustruje to rys. 1 przedstawiający ramkę filmu, w której refleks świetlny pokrywa kilkumetrowy odcinek toru na całej jego szerokości. Jak się wydaje, nie powinno to stanowić dużego problemu, gdyż osoba, która przepływa przez tego typu obszar, zaburzy wodę i rozproszy światło – to zaś spowoduje, że pływaka zobaczymy wyraźnie, jak to pokazuje rys. 2.

Takie rozumowanie jest jednak problematyczne, ponieważ adaptacyjne algorytmy usuwania tła zwykle dopasowują się do ustalonej postaci odbicia światła. Kiedy zaś pływak zaburzy wodę i zniknie refleks świetlny, doprowadzi to do wykrycia dużego pozornego obiektu. Dodatkowo kiedy słonce chowa się za chmurami, to również powoduje zmianę, do której algorytm musi się dopasować. Proces takiej adaptacji może trwać nawet kilkanaście sekund, co stanowi duży problem w bieżącej analizie.



Rys. 1. Refleks świetlny pokrywający dużą część toru
Fig. Sunlight reflex covering a large part of the track



Rys. 2. Refleks świetlny całkowicie rozproszony przez pływaka
Fig. 2. Sunlight reflex completely dispersed by a swimmer

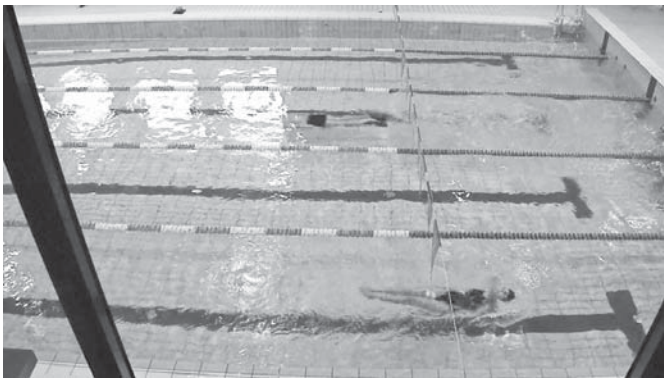
Podobnie, jeśli ktoś popłynie pod wodą i pod refleksem świetlnym. Efekt taki można zaobserwować na trzech ramkach filmu, przedstawionych na rys. 3, które zostały pobrane w odstępie około jednej sekundy. Doskonale widać, jak sylwetka człowieka znika pod refleksami świetlnymi.

Warto zauważyć, że basen, z którego pochodzą zdjęcia, ma jedynie pięć okien standardowej wielkości na przeciwległej ścianie względem kamery. Zatem można przypuszczać, że zjawisko to może się nasilać na obiektach pływackich, w których większość ściany jest przeszklona. Najlepiej jest zapewnić, aby kierunek patrzenia kamery nie był bezpośrednio w kierunku okien (aby jak najmniejsza część tafli wody mogła odbijać światło słoneczne bezpośrednio do obiektywu).

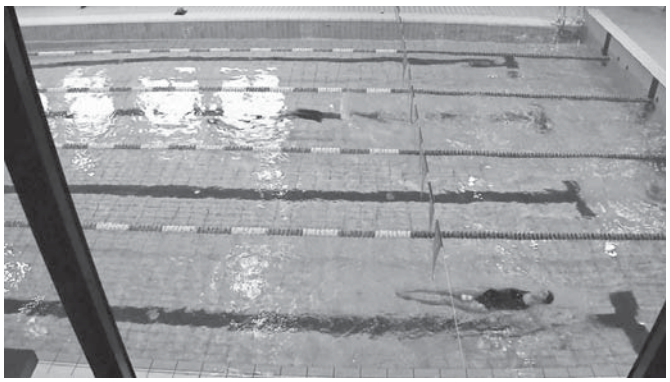
3.3. Refleksy od sztucznego oświetlenia

Oprócz refleksów spowodowanych światłem słonecznym, na wodzie powstają także odbicia światła lamp (można to zaobserwować np. na rys. 1 w prawym dolnym rogu obrazu). Jednak w stosunku do odbić światła słonecznego stanowi to znacznie mniejszy problem, łatwiejszy do wyeliminowania. Istnieje małe prawdopodobieństwo, że z powodu takiego odbicia kamera nie zarejestruje czegoś ważnego. Jednak im bardziej prostopadle kamera jest ustawiona względem wody, tym problem staje się większy – zwłaszcza w świetle analizy przedstawionej w punkcie 3.1. Rozwiązaniem tego problemu może być też zastosowanie filtru polaryzacyjnego, który zmniejsza skutki odbić światła.

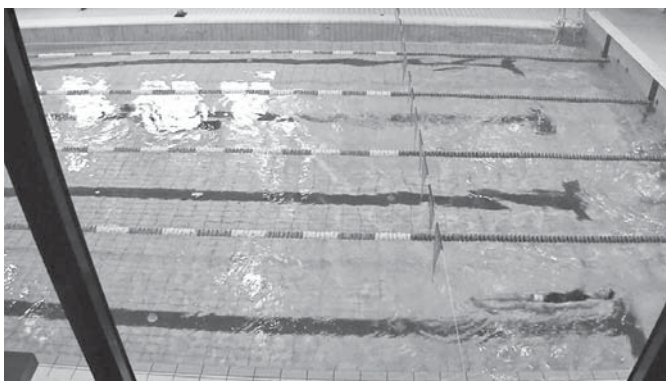
MOST WIEDZY Pobrano z mostwiedzy.pl



(a)



(b)



(c)

Rys. 3. Pływak znikający pod refleksem świetlnym: (a) pływak z prawej strony refleksu, (b) widać tylko głowę i płetwy, (c) widać wynurzającą się głowę

Fig. 3. Swimmer disappearing under the sunlight reflex: (a) swimmer to the right side of the reflex, (b) visible only head and fins, (c) visible emerging head

4. Algorytmy usuwania tła

Zamierzonym etapem projektu jest wybór algorytmu służącego odcinaniu tła tak, aby na obrazie pozostały jedynie istotne obiekty. Działanie takich algorytmów opiera się zwykle na rekursywnym wyznaczaniu modelu tła, a następnie traktowanie wszystkiego, co wykracza poza ten model jako plan pierwszy. Algorytmy przeanalizowano pod kątem łatwości w implementacji, czasu działania, dokładności oraz możliwości przystosowania zarówno do nowego obiektu, jak i do zmieniającego się tła w basenie. Usuwanie tła z obrazu przeprowadzać można na wiele sposobów. Istnieją prace, np. [6], poświęcone porównaniu różnych algorytmów, ich wad i zalet.

Biblioteka OpenCV przeznaczona do komputerowej analizy obrazu, do odcinania tła przeznacza algorytm mieszaniny rozkładów Gaussa MOG (ang. *Mixture of Gaussians*) oraz metodę *k*-najbliższych sąsiadów KNN (ang. *K-Nearest Neigbors*). Poza nimi rozważymy też inne, opisane dalej algorytmy (nie mieszczące się w tej bibliotece).

4.1. Pojedynczy rozkład Gaussa

Przeznaczeniem algorytmu Gaussa [9] jest odkrywanie modelu tła oraz następnie wskazywanie tych miejsc w nowej ramce obrazu, gdzie występują znaczące różnice. Przeszła informacja w kolejnych ramkach zmniejsza swój wpływ na model tła w tempie wykładniczym. Dzięki zaś odpowiedniemu doborowi wartości tzw. stałej zapominania, można regulować efektywną długość takiej 'pamięci'. W prostym przypadku, aktualizacja modelu tła odbywa się według filtru pierwszego rzędu:

$$m_t = \alpha y + (1 - \alpha) m_{t-1} \tag{1}$$

gdzie m_t jest modelem w czasie t, y to nowa ramka filmu, natomiast współczynnik α to stała zapominania z przedziału [0,8; 1). Miejsca na obrazie, które są różne od modelu m_t wykrywane są z zastosowaniem odpowiedniego progu:

$$|y - m_t| > \kappa \sigma_t \tag{2}$$

opartego na założonym odchyleniu standardowym σ_t oraz współczynniku skalującym κ. Przyjmując $\kappa = 2{,}5$ wymagamy, aby wartość piksela mieściła się w 'odległości' nie większej niż 2,5-krotność odchylenia standardowego od wyznaczonej wartości oczekiwanej, aby został uznany za tło. Takie podejście jest użyteczne [9] w stacjonarnym otoczeniu (na przykład w biurze, z jedną tylko poruszająca się osobą).

4.2. Mieszanina rozkładów Gaussa – MOG

Algorytm mieszanin rozkładów Gaussa (MOG, ang. *Mixture of Gaussians*) [5] reprezentuje jedna z bardziej popularnych metod stosowanych do usuwania tła z obrazu – można też spotkać się z określeniem GMM (ang. *Gaussian Mixture Modeling*), tj. modelowaniem mieszanina rozkładów Gaussa.

W algorytmie MOG/GMM każdy piksel modelowany jest za pomocą mieszaniny k rozkładów Gaussa (gdzie liczba k zwykle wynosi 3–5), co oznacza, że algorytm przewiduje występowanie w danym miejscu obrazu k różnych typów obiektów. W przypadku basenu oznaczać to może np. wodę, również odrobinę spienioną, albo fragment liny rozdzielającej tory. Przy czym każdy z tych trzech obiektów też może być interpretowany jako tło. Podobnie jak w prostej metodzie Gaussa, jeśli wartość danego piksela jest oddalona od wartości średniej mniej niż 2,5-krotność odchylenia standardowego dla któregokolwiek z rozkładów, piksel ten zostaje uznany za istotny. Wszystkie inne punkty są klasyfikowane jako tło. Aktualizowane są tylko te rozkłady, dla których znaleziono dopasowanie.

W bibliotece OpenCV znajduje się ulepszona wersja tego algorytmu [12], w której dobiera się dynamicznie liczbę k. Dzięki temu zapewnia się większa szybkość, jeśli tło jest proste, oraz większą dokładność w skomplikowanych przypadkach. Oznacza to, że dostarczony przez bibliotekę OpenCV algorytm MOG jest stosunkowo szybki i bardzo dobrze dostosowuje się do nowych warunków.

4.3. K-najbliższych sąsiadów – KNN

Metoda K-najbliższych sąsiadów (KNN, ang. *K-nearest neighbours*) stosowana jest m.in. przy klasteryzacji [2]. Algorytm potrzebuje jakiejś miary odległości między określonymi próbkami, które należy sklasyfikować. Dogodnie wykorzystuje się do tego celu metrykę euklidesową, czyli miarę odległości między dwoma punktami w dowolnej przestrzeni, w której każdy wymiar pozwala na reprezentację jednego z parametrów próbki.

MOST WIEDZY Pobrano z mostwiedzy.pl

Potrzebny jest zbiór uczący, który ma już odpowiednio przypisane klastry. Mając takie dane, dla każdej nowej próbki podlegającej klasyfikacji określa się zbiór najbliższych K próbek z zestawu treningowego. Następnie na podstawie liczby przedstawicieli poszczególnych klas wśród znalezionych K-najbliższych sąsiadów, próbka przyporządkowywana jest do najbliższej (podobnej) grupy.

Dostosowanie tej metody do zadania eliminacji tła z obrazu zostało opisane w pracy [13]. Dodatkowo, do tego algorytmu wprowadzono też adaptację polegającą na tym, że każda nowa próbka (po sklasyfikowaniu) dopisywana jest do tła.

5. Badania eksperymentalne

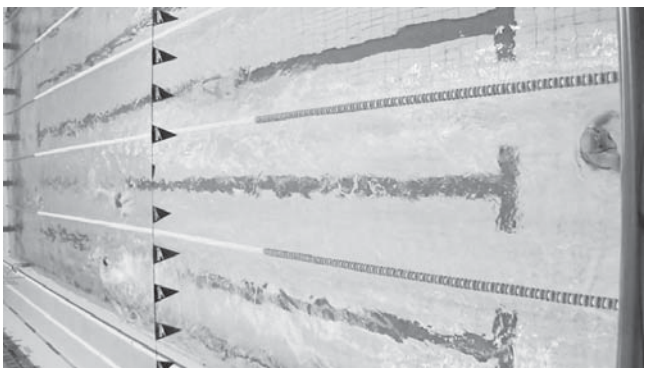
Przedstawione tu wyniki reprezentują czarno-białe maski. Białe punkty odpowiadają pierwszemu planowi, natomiast czarne są interpretowane jako tło. Testy prowadzone na jednym nagraniu, zaprojektowano w taki sposób, aby w określonym momencie (około setnej sekundy nagrania) zapisać na dysku maskę wynikową. Oryginalną ramkę, do której odnoszą się te wyniki, pokazano na rys. 4, gdzie widać pięciu pływaków, z których najmniejszy i najmniej widoczny znajduje się daleko z lewej strony. Algorytmy MOG i KNN miały wystarczająco dużo materiału, aby dokonać poprawnej adaptacji.

Przy różnych otworowaniach kolorów, algorytmy MOG i KNN mają inną wrażliwość, którą można regulować przeznaczonym do tego parametrem. Aby porównać ze sobą wszystkie warianty, dopasowano poziomy czułości w taki sposób, aby na masce wynikowej każdy z algorytmów miał tyle samo białych pikseli (z dokładnością do 1%). Ze względu na to, że algorytm opcjonalny BOT nie ma takiej możliwości regulacji, traktujemy go jako rozwiązanie porównawcze (ang. *benchmark*). W algorytmie KNN domyślny poziom czułości wynosi 400, natomiast w metodzie MOG jest to 16 (oczywiście parametru stosowanego w różnych algorytmach nie należy bezpośrednio porównywać).

5.1. Zrealizowane testy

Przeprowadzono operacje usuwania tła z wykorzystaniem algorytmów MOG i KNN. Korzystając przy tym z uwag poczynionych w części 4.4, zastosowano różne metody odwzorowania kolorów. Uzyskane wyniki porównano z osiągnieciami metody opcjonalnej BOT opisanej w 4.5.

Maski stworzone przez poszczególne algorytmy odpowiadające ramce z rys. 4 zilustrowano na rys. 5 w konfiguracji macierzowej – w taki sposób, aby w kolumnie lewej były maski pozyskane z algorytmu MOG, zaś w prawej wyniki działania algorytmu KNN. Wiersze w tej macierzy odpowiadają różnym odwzorowaniom kolorów, kolejno RGB, HSV, oraz H (sama barwa). W ostatnim wierszu (po środku) znajduje się maska uzyskana metoda BOT. Widać, że na wszystkich uzyskanych obrazach znajduje się sporo małych obiektów, które można łatwo usunąć za pomocą operacji morfologicznych.



Rys. 4. Ramka – nagranie z około setnej sekundy
Fig. 4. Frame recorded about a hundredth second

Zestawienie wyników na rys. 6 podane jest w identycznej konfiguracji, jednak tym razem maski uzyskano po przeprowadzeniu odpowiednich operacji morfologicznych. Aby ułatwić porównanie, zastosowano takie same morfologie w odniesieniu do wszystkich algorytmów (MOG, KNN, BOT). Zastosowano przy tym operacje zamknięcia z jadrem o wielkości 3×3 oraz operacje otwarcia jadrem wielkości 5×5.

5.2. Analiza wskaźnikowa

W tabeli 1 przedstawiono wartości uwzględnionych wskaźników i parametrów działania testowanych algorytmów.

Tabela 1. Parametry i wskaźniki algorytmów odcinania tła z użyciem morfologii
Table 1. Parameters and indicators of background removal algorithms using morphology

nazwa	α	τ(ms)	N	β
BOT	–	20	18	86
MOG RGB	66	25	21	0
MOG HSV	104	29	11	114
MOG H	28	25	12	112
KNN RGB	2880	43	18	0
KNN HSV	5376	47	11	0
KNN H	512	45	7	36

Nazwy scenariuszy wycinania tła wskazane w tej tabeli są następujące:

BOT – dla algorytmu opcjonalnego BOT opisanego w punkcie 4.5,

MOG RGB – z wykorzystaniem algorytmu MOG na ramce RGB,

MOG HSV – z wykorzystaniem algorytmu MOG na ramce HSV,

MOG H – z wykorzystaniem algorytmu MOG na ramce z barwa (ang. *hue*),

KNN RGB – z wykorzystaniem algorytmu KNN na ramce RGB,

KNN HSV – z wykorzystaniem algorytmu KNN na ramce HSV,

KNN H – z wykorzystaniem algorytmu KNN na ramce z barwa H.

Uwzględniono następujące parametry oraz wskaźniki działania testowanych algorytmów:

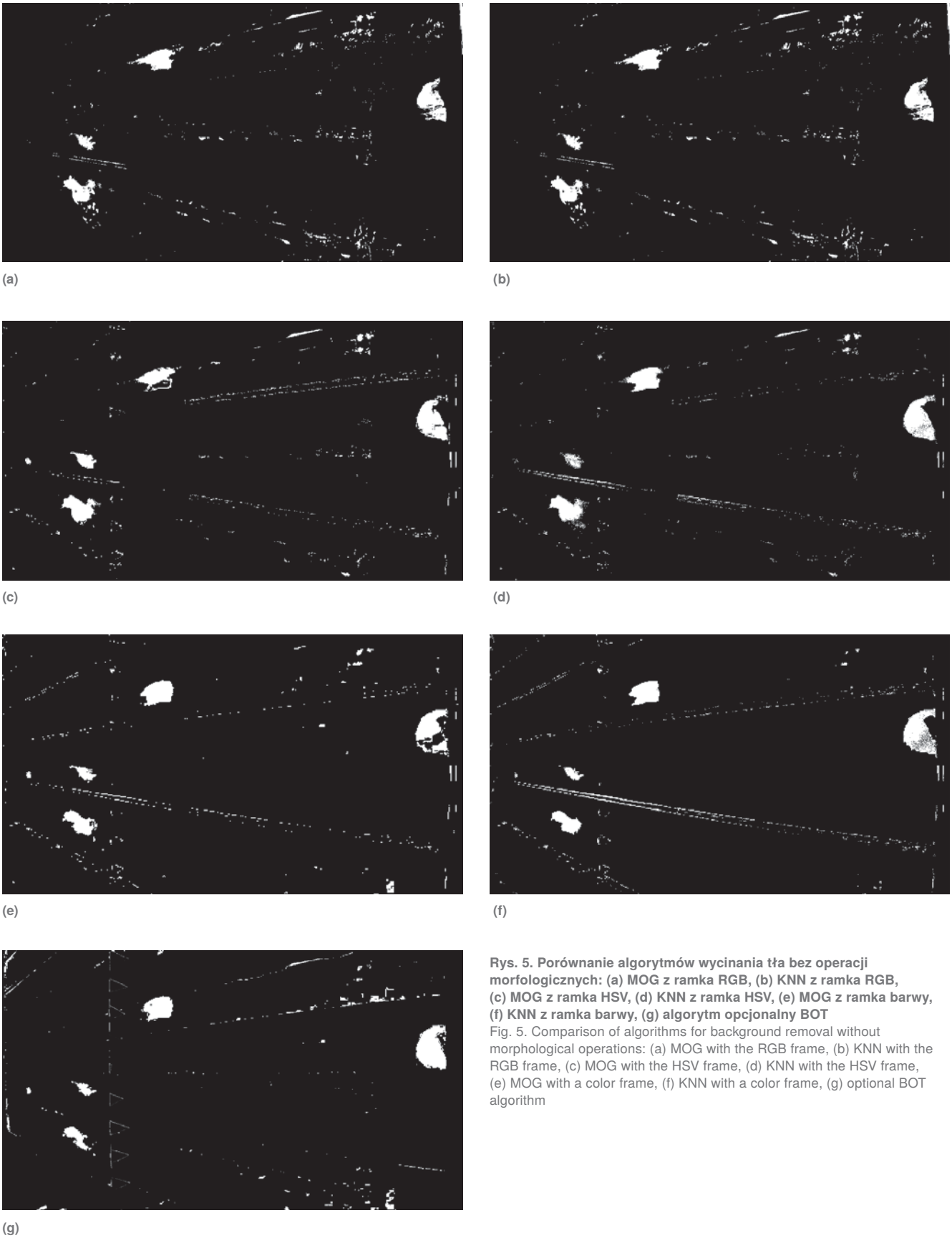
α – poziom czułości algorytmu ustawiony tak, aby maska wynikowa zawierała określoną liczbę pikseli pierwszego planu (im wyższa liczba, tym mniejsza wrażliwość),

τ – średni czas potrzebny na analizę jednej ramki bez operacji morfologicznych, mierzony w milisekundach z dokładnością do pełnych milisekund (im mniej, tym lepiej),

N – liczba konturów wykrytych na masce wyjściowej (rys. 6a–d) (im mniej, tym lepiej),

β – liczba pikseli zinterpretowanych jako pierwszoplanowe w okolicach miejsca, gdzie znajduje się najmniejszy pływak (idealnie powinno być około 160).

MOST WIEDZY Pobrano z mostwiedzy.pl



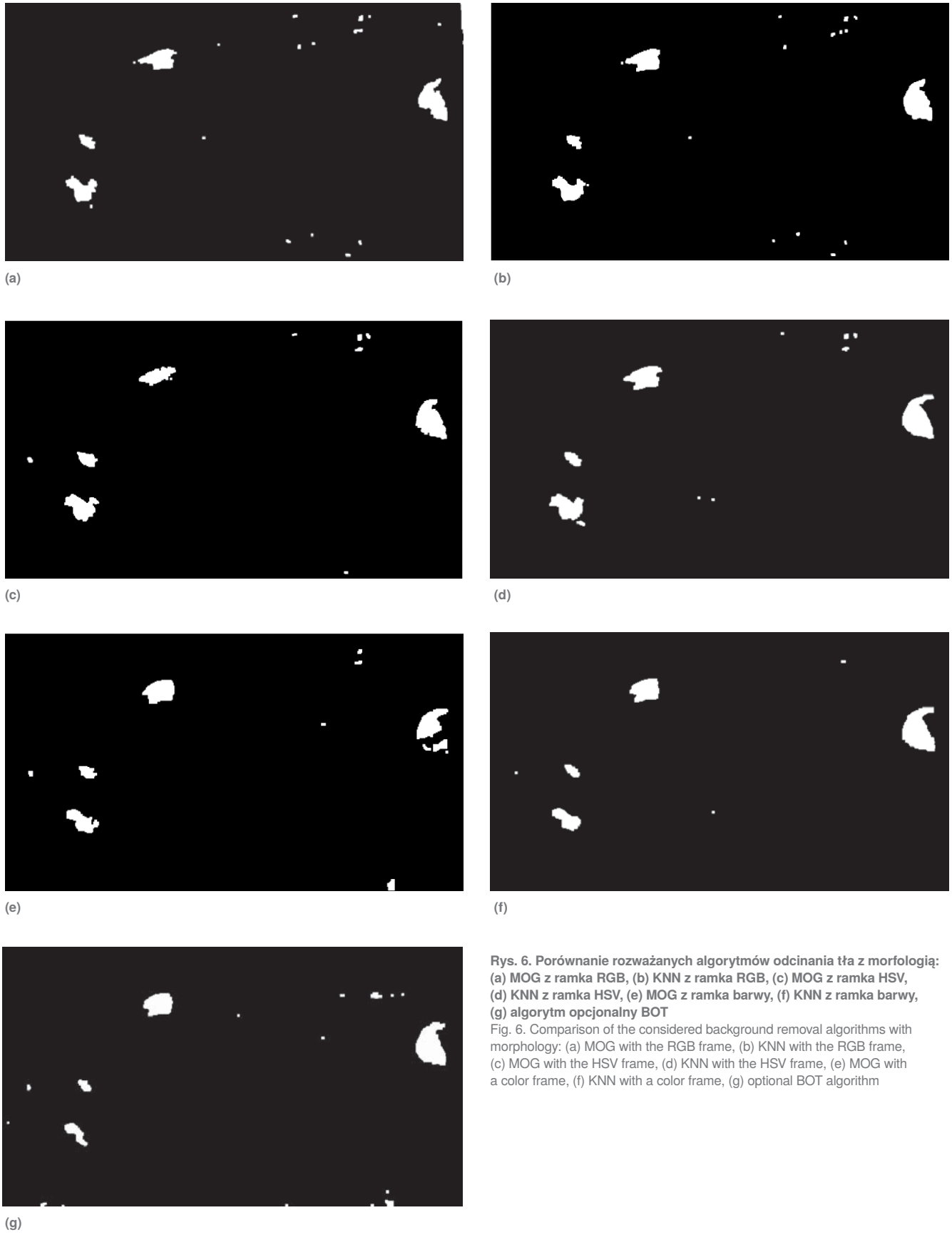
Rys. 5. Porównanie algorytmów wycinania tła bez operacji morfologicznych: (a) MOG z ramka RGB, (b) KNN z ramka RGB, (c) MOG z ramka HSV, (d) KNN z ramka HSV, (e) MOG z ramka barwy, (f) KNN z ramka barwy, (g) algorytm opcjonalny BOT
Fig. 5. Comparison of algorithms for background removal without morphological operations: (a) MOG with the RGB frame, (b) KNN with the RGB frame, (c) MOG with the HSV frame, (d) KNN with the HSV frame, (e) MOG with a color frame, (f) KNN with a color frame, (g) optional BOT algorithm

Wszystkie próby zostały przeprowadzone wykorzystując nagranie 960 × 540 pikseli, na komputerze z dwurdzeniowym procesorem Intel Pentium P6200 z taktowaniem o częstotliwości 2,13 GHz.

5.3. Analiza wyników

Tabela 1, w której użyto symboli opisanych w punkcie 5.2, prezentuje podsumowanie wyników przeprowadzonych testów. W systemach automatycznego śledzenia ważne jest, aby algo-

MOST WIEDZY Pobrano z mostwiedzy.pl



Rys. 6. Porównanie rozważanych algorytmów odcinania tła z morfologią: (a) MOG z ramka RGB, (b) KNN z ramka RGB, (c) MOG z ramka HSV, (d) KNN z ramka HSV, (e) MOG z ramka barwy, (f) KNN z ramka barwy, (g) algorytm opcjonalny BOT

Fig. 6. Comparison of the considered background removal algorithms with morphology: (a) MOG with the RGB frame, (b) KNN with the RGB frame, (c) MOG with the HSV frame, (d) KNN with the HSV frame, (e) MOG with a color frame, (f) KNN with a color frame, (g) optional BOT algorithm

MOST WIEDZY Pobrano z mostwiedzy.pl

rytm działał szybko oraz aby wykrywane były tylko te obiekty, które podlegają śledzeniu. W przypadku liczby konturów oczekiwano wyniku, jak najbliższego pięciu (faktycznej liczbie obiektów). Oczywiście wszystkie osoby, które znajdują się w basenie powinny mieć swoją wyraźną reprezentację w masce wyjściowej obrazu, to zaś zadanie jest weryfikowane pośrednio przez parametr β.

Analizując uzyskany poziom czułości zauważyć można, że przy odwzorowaniu HSV, algorytmy muszą być możliwie mało czułe (czyli mieć wysoki parametr α). Natomiast wskaźnik poziomu czułości przy ramce opartej tylko na barwie (H) powinien być znacznie niższy.

Pod względem średniego czasu potrzebnego do analizy jednej ramki widać, że warianty z wykorzystaniem algorytmu KNN są prawie dwukrotnie wolniejsze (od 43 ms do 47 ms) od wariantów opartych na mieszaninie rozkładów Gaussa MOG (od 25 ms do 29 ms). Algorytm opcjonalny BOT jest najszybszy (20 ms). Różnice czasowe między reprezentacjami RGB i HSV, zarówno w przypadku KNN, jak i MOG, wynoszą zaledwie około 4 ms (prawdopodobnie wynikają z operacji konwersji). W przypadku ramki z samą barwą (H) czas stracony na konwersję jest częściowo rekompensowany przez działanie na mniejszej liczbie danych.

W kategoriach jakości odcinania tła w tych konkretnych warunkach, zróżnicowanie między poszczególnymi metodami również jest znaczne. Trzy z siedmiu przetestowanych metod (MOG RGB, KNN RGB oraz KNN HSV) nie wykryły najmniejszego pływaka, tj. parametr β dla tych rozwiązań wynosi 0. Ponadto metody oparte na odwzorowaniu RGB charakteryzowały się największa liczbą wykrytych konturów ($N = 21$ i $N = 18$), co oczywiście utrudnia dalszą ich analizę.

W przypadku KNN H wykryty obiekt był bardzo mały ($\beta = 36$), jednak na korzyść tego wariantu warto dodać, że ma on również najmniejszą liczbę pozostawionych na obrazie konturów $N = 7$, co oznacza, że tylko dwa z nich nie reprezentują pływaka. W przypadku MOG H wyniki są dobre, najmniejszy pływak został wykryty; obiekt ten ma 112 pikseli, co zapewnia, że nie będzie zignorowany podczas dalszej obróbki (jak może się zdarzyć w przypadku KNN H). Liczba konturów też nie jest duża, jednak jest to jedyna metoda, przy której jeden z obiektów został podzielony na kilka niezależnych konturów (to również może utrudniać analizę).

Algorytm opcjonalny BOT prowadzi do dużej liczby konturów na obrazie ($N = 18$). Wielkość najmniejszego pływaka ($\beta = 86$) jest wystarczająca na potrzeby analizy. Wariant MOG HSV pozostawił na obrazie 11 konturów, czyli o 6 więcej niż potrzeba, zaś wielkość najmniejszego pływaka jest bardzo dobra ($\beta = 114$), podobna jak w przypadku MOG H ($\beta = 112$).

Warto zwrócić uwagę, że w przypadku przeanalizowanych rozwiązań, poza algorytmem opcjonalnym BOT, błędnie interpretowane obiekty pozostawione na obrazie dotyczą głównie prawej połowy obrazu. Nie jest to sytuacja wyjątkowa dla tej konkretnej ramki, a raczej ogólna właściwość związana z ustawieniem kamery. Ponieważ po prawej stronie obiekty są generalnie większe, wykluczenie szumów po tej stronie jest łatwiejsze. Jednym z możliwych rozwiązań jest zwiększanie wielkości jądra przekształceń morfologicznych wraz ze wzrostem odległości od lewej krawędzi obrazu, tak jak zaproponowano w pracy [7], gdzie podzielono obraz na kilka części, w których zastosowano jądra różnej wielkości. Dzięki takiemu rozwiązaniu w szerszym zakresie usuwa się niepotrzebne informacje z obrazu, bez ryzyka usunięcia ważnego obiektu. Takie podejście powinno poprawić rezultaty również w przypadku algorytmu BOT, jednak nadal pozostałyby obiekty położone z lewej strony. W przypadku tych obiektów należałoby zmienić sposób wykluczania fałszywych obiektów – zwłaszcza w części, gdzie analizowana jest przestrzeń poza basenem.

6. Podsumowanie i wnioski

Przeanalizowano rozmaite sposoby usuwania tła w nagraniach pochodzących z monitoringu basenu publicznego. Przetestowano siedem różnych wariantów realizacji tego zadania. Konstruktywnym wnioskiem z tej analizy jest wskazanie algorytmu MOG jako najbardziej efektywnego, opartego na ramkach z odwzorowaniem kolorów HSV. Zaleca się przy tym dodatkowe dostrojenie samych operacji morfologicznych (ważnych dla skuteczności całości monitoringu), aby osiągnąć dalszą poprawę skuteczności działania mieszaniny rozkładów Gaussa (MOG).

Zaproponowany algorytm opcjonalny BOT, opisany w punkcie 4.5, również zaleca się poddać usprawnieniu działania w odniesieniu do przestrzeni poza nisza basenowa. Wówczas algorytm BOT będzie mógł prawdopodobnie konkurować ze 'zwycięskim' algorytmem MOG HSV. Algorytm BOT jest znacznie szybszy (rzędu 50%), ale nie ma możliwości adaptacji (nawet do niewielkich zmian w analizowanym środowisku). Z kolei wadą algorytmów adaptacyjnych jest to, że pływacy pozostający w bezruchu są stopniowo ignorowani (aż do momentu, kiedy zaczną się ponownie poruszać).

Bibliografia

1. Baldini G., Campadelli P., Cozzi D., Lanzarotti R., *A simple and robust method for moving target tracking*, Proceedings of the International Conference Signal Processing Pattern Recognition and Applications (SPPRA 2002) 2002, 108–112.
2. Bishop C.M., Neural Networks for Pattern Recognition, chap. 2, 51–57. Clarendon Press. 1995.
3. Dai J., Li Y., He K., Sun J., *R-FCN: Object detection via region-based fully convolutional networks*, https://arxiv.org/pdf/1605.06409.pdf.
4. Gonzalez R.C., Woods R.E., *Image segmentation*. [in:] *Digital Image Processing* (2nd Ed), chap. 10, 567–635. Prentice Hall 2002.
5. KaewTraKulPong P., Bowden R., *An improved adaptive background mixture model for real-time tracking with shadow detection*. The 2nd European Workshop on Advanced Video-based Surveillance Systems. 1–5. Kluwer Academic Publishers 2001.
6. Piccardi M., *Background subtraction techniques: A review*. 2004 IEEE International Conference on Systems, Man and Cybernetics. Vol. 4, 3099–3104. IEEE (2004), DOI: 10.1109/ICSMC.2004.1400815.
7. Reiter K., *Monitoring basenu publicznego*. Praca magisterska nr pg/weti/ksdir/zk263**m**/09/2017 (promotor: prof. Z. Kowalczuk), Wydział ETI, Politechnika Gdańska, Gdańsk 2017.
8. Rypniewski J.J., *Monitoring wizyjny w pracy ratownika wodnego* (2014), http://plywalnieibaseny.pl/monitoring-wizyjny-w-pracy-ratownika-wodnego.
9. Wren C.R., Azarbayejani A., Darrell T., Pentland A.P., *Pfinder: Real-time tracking of the human body*. IEEE Trans. Pattern Anal. Mach. Intell. 19(7), 780–785, 1997.
10. Yilmaz A., Javed O., Shah M., *Object tracking: A survey*. ACM Computing Surveys 38(4), 13:1–45 (2006)
11. Zhu X., Zhu J.D.X., Yuan Y.W.L., *Towards high performance video object detection for mobiles* 2018, https://arxiv.org/pdf/1804.05830.pdf
12. Zivkovic Z., *Improved adaptive gaussian mixture model for background subtraction*. Proceedings of the 17th International Conference on Pattern Recognition, 2004. ICPR 2004. Vol. 2, pp. 28–31. IEEE 2004.
13. Zivkovic Z., Heijden, F.v.d.: *Improved adaptive gaussian mixture model for background subtraction*. Pattern Recognition Letters 27(7), 773–780 (2006) 15

MOST WIEDZY Pobrano z mostwiedzy.pl

Background removal in video recordings from swimming pool monitoring

Abstract: Automatic real-time image processing is crucial for many (video surveillance) monitoring solutions used, among others for security purposes. Often one of the most important stages of computer vision processing is separating the background from the objects in the foreground, so as to exclude all irrelevant information from the image. The aim of this work is to summarize the experience gained while tracking swimmers and to show the possibility of effective automatic video surveillance of people using a swimming pool. The effectiveness of two selected algorithms (MOG and KNN) is compared using different color mappings and the advantages and disadvantages of the analyzed methods are discussed

Keywords: Video surveillance, background subtraction, video processing, OpenCV, noise cancellation

prof. dr hab. inż. Zdzisław Kowalczuk
kova@pg.gda.pl

Prof. zw. dr hab. inż. (2003, 1993, 1986, 1978). Od 1978 uczestniczy w życiu naukowym na Wydziale Elektroniki, Telekomunikacji i Informatyki PG, gdzie jest profesorem zwyczajnym w dziedzinie Automatyka i Robotyka oraz szefem (założonej w 2006 r.) Katedry Systemów Decyzyjnych i Robotyki. Składał wizyty naukowe i wykładał na Uniwersytecie w Oulu (1985), Australijskim Uniwersytecie Narodowym (1987), Politechnice w Darmstadt (1989) oraz na Uniwersytecie George'a Mansona (1990–1991). Zajmuje się sterowaniem adaptacyjnym, identyfikacją systemów, detekcją błędów i diagnostyką przemysłową, przetwarzaniem sygnałów, sztuczną inteligencją oraz inżynierią sterowania i informatyką. Jako autor i współautor wydał 16 książek (w tym WNT 2002, oraz Springer 2004 i 2014), opracował ponad 100 artykułów oraz ponad 300 referatów konferencyjnych i rozdziałów. Wyróżniony w 1990 i 2003 Nagrodami Ministra Edukacji Narodowej (MENIS) oraz w 1999 Nagrodą Fundacji na Rzecz Nauki Polskiej w zakresie sterowania. Prezes POLSPAR i wiceprezes Towarzystwa Konsultantów Polskich, redaktor wydawnictwa PWNT.

mgr inż. Konrad Reiter
konradrekon@gmail.com

Absolwent Wydziału Elektroniki, Telekomunikacji i Informatyki Politechniki Gdańskiej. Studia ukończył we wrześniu 2017 r. w dyscyplinie Automatyka i Robotyka w Katedrze Systemów Decyzyjnych i Robotyki. Od tegoż roku podjął pracę w firmie Enbio przy projekcie autoklawu medycznego, gdzie może łączyć swoje zainteresowanie programowaniem z wiedzą zdobytą podczas studiów. W wolnym czasie gra w siatkówkę i lata na paralotni.

MOST WIEDZY Pobrano z mostwiedzy.pl