

POLITECHNIKA WROCŁAWSKA
WYDZIAŁ ELEKTRONIKI

KIERUNEK: Automatyka i Robotyka
SPECJALNOŚĆ: Technologie informacyjne w systemach automatyki

PRACA MAGISTERSKA

Metody detekcji obiektów w sekwencjach
video z ruchomą kamerą

Object detection methods in video sequences
recorded by moving camera

AUTOR:
Karol Ankutowicz

PROWADZĄCY PRACĘ:
dr inż. Bartosz Jabłoński,
W-4/K-8

OCENA PRACY:

Niniejsza praca przedstawia jeden ze sposobów wykorzystania przetwarzania obrazów. Współcześnie mamy z nim do czynienia wszędzie tam gdzie konieczna jest analiza obrazów pod kątem konkretnej ich właściwości. Problem jakim zajmujemy się w tej pracy dotyczy wykrywania oraz śledzenia poruszających się obiektów w sekwencji wideo w sytuacji kiedy kamera się porusza. Sekwencję taką potraktujemy jako zbiór obrazów będących w uporządkowanej kolejności. Z pomocą wybranych metod detekcji obiektów będziemy mogli badać zachowanie wykrywanych obiektów. Z naszej perspektywy interesujące jest to czy jest możliwe skuteczne śledzenie obiektów uwzględniając nie tylko ich ruch ale również ruch kamery. Zajmiemy się zaprogramowaniem a następnie analizą wybranych metod w celu sprawdzenia ich przydatności do tego typu zadań.

Spis treści

1. Wstęp	4
1.1. Geneza pracy	4
1.2. Cel pracy	5
2. Przegląd metod detekcji obiektów	7
2.1. Cel i tematyka przeglądu	7
2.2. Metody wykrywania i śledzenia obiektów	9
2.2.1. Metody Kanade-Lucas-Tomasi i RANSAC	10
2.2.2. Segmentacja MAP (TS-MAP) z algorytmem Mean-Shift	12
2.2.3. Operator Kirsch wraz z LDP i LBP	12
2.2.4. Algorytm Canny oraz transformacja Hougha	13
2.3. Podsumowanie przeglądu literaturowego	14
3. Teoria przetwarzania wideo w kontekście detekcji obiektów	15
3.1. Cel przetwarzania	15
3.2. Metody wykrywania obiektów	16
3.2.1. Metody wykrywania krawędzi i kształtów	16
3.2.2. Metody segmentacji	19
3.3. Metody śledzenia obiektów	21
4. Projekt badawczy	22
4.1. Założenia projektowe	22
4.2. Dobór metod	22
4.3. Dobór materiału badawczego	23
5. Wyniki eksperymentów numerycznych	24
5.1. Przykładowe obrazy wynikowe	24
5.2. Analiza otrzymanych wyników	29
5.2.1. Progowanie z algorytmem SURF	30
5.2.2. Detektor krawędzi Canny z algorytmem SURF	35
5.2.3. Progowanie z metodą różnicową	37
5.2.4. Detektor krawędzi Canny z metodą różnicową	41
5.2.5. Nieparametryczne metody wykrywania krawędzi z algorytmem SURF	42
5.2.6. Nieparametryczne metody wykrywania krawędzi z metodą różnicową	43
5.3. Porównanie działania algorytmów	44
6. Podsumowanie i wnioski	51
7. Spis rysunków	54
8. Spis tabel	55
9. Bibliografia	57

1. Wstęp

1.1. Geneza pracy

Współczesne technologie, zarówno te sprzętowe jak i programowe, ulegają ciągłemu rozwojowi. Rozwój ten jest ze sobą ściśle związany – coraz nowsze i lepsze algorytmy i wykorzystujące je aplikacje wymagają do dobrego działania coraz lepszego sprzętu. Spośród wielu gałęzi technologii związanych z programowaniem skupmy się na jednej – przetwarzaniu obrazów. Najprościej mówiąc jest to uzyskanie z obrazów interesujących nas informacji. Uwypuklenie potrzebnych danych pozwala nam na dalsze działanie. Często jednak, mimo znacznego zmniejszenia ilości danych, nadal stoimy przed koniecznością wielokrotnego wykonywania ciągle tych samych operacji. Praca ta zlecona człowiekowi na pewno jest uciążliwa. Poza tym ludzkie oko często bywa zawodne. Po uwzględnieniu jeszcze subiektywnych odczuć możemy mieć spore obawy co do powtarzalności analizy. Wobec takiego obrotu sprawy jak najbardziej zasadne wydaje się posiłkowanie się algorytmami wspomagającymi analizę przez człowieka lub nawet ją zastępujący. Dążymy w ten sposób do powtarzalności i wysokiej skuteczności reagowania na rezultaty przetwarzania.

Przetwarzanie obrazów i automatyzacja jego analizy znajduje szerokie zastosowanie i spotykamy się z nim w życiu codziennym. Przykładem może być system kontroli natężenia ruchu drogowego na skrzyżowaniu. Odpowiednio ustawiona kamera może spokojnie zliczać obiekty – pojazdy możemy podzielić na kilka grup (np. samochody osobowe, ciężarowe) w ramach których różne modele są już do siebie bardzo zbliżone kształtem. Zatem mając pewną niedużą bazę wzorców możemy łatwo wykrywać określone obiekty i w tym przypadku dokładnie określić ilość wykrywanych pojazdów.

Przykładem który z kolei ciężko całkowicie zautomatyzować mogą być np. systemy monitorujące różne ważne obiekty (budynki) w celu wykrywania zagrożeń. Ich mankamentem jest to, że potencjalna liczba niebezpieczeństw jest bardzo duża. Zaliczyć się może do nich podejrzanе zachowanie jakiejś osoby lub wzniecony pożar. Oprogramowanie zachowania algorytmu w każdym z nich jest bardzo złożonym zadaniem. Dlatego nadal główną rolę w takim przypadku odgrywa człowiek, który kontroluje to co można zauważyć.

Widzimy zatem, że są zarówno takie problemy, które można całkowicie powierzyć dobrze napisanemu algorytmowi oraz takie gdzie jeszcze jest na czym popracować. Problem, jakiemu jest poświęcona ta praca ma nieco bardziej ogólny – nie mamy sprecyzowanego konkretnego rodzaju obiektu. Musimy mieć za to mieć pewność, że śledzimy ciągle ten sam obiekt. W tym

celu musimy w jakiś sposób wyróżnić go z tła. Możemy to osiągnąć poprzez metody bazujące np. na konturach lub barwach. Będziemy również badać jego zachowanie a dodatkowym aspektem będzie ruch kamery. Mając zatem określone zadanie do wykonania oraz znając sposoby przetwarzania informacji możemy skonkretyzować problem badawczy i wykorzystując znane nam metody – rozwiązać go.

1.2. Cel pracy

Celem pracy jest badanie metod detekcji obiektów w sekwencjach wideo z ruchomą kamerą. Właśnie ruch kamery może mieć tutaj decydujące znaczenie w analizie skuteczności znanych powszechnie metod. Powoduje to konieczność nieco innego podejścia względem statycznej kamery. Chodzi tu nie tylko o zmienność tła ale również większe możliwości zmiany położenia, wielkości i orientacji śledzonego obiektu. Analiza wybranych algorytmów pozwoli nam ocenić ich skuteczność właśnie w tym zastosowaniu. Są one przedstawione w kolejnych rozdziałach pracy. Finalnie będziemy mogli ocenić jak wybrane metody poradziły sobie z postawionym zadaniem.

Rozdział pt. *Przegląd metod detekcji obiektów* przedstawia nam na wstępie zapoznanie się z tematem poprzez zaprezentowanie przykładów z pozycji literaturowych. W poszukiwaniach odpowiednich przykładów istotne były: śledzenie obiektów w sekwencjach wideo oraz ruch kamery, którą nagrywamy. Jeśli przynajmniej jeden z tych aspektów był omawiany – była to podstawa do włączenia go do analizy. Same obiekty niekoniecznie należały do obiektów ruchomych (wśród tych statycznych możemy wyróżnić m. in. znaki drogowe). Część z nich zostało później wykorzystana do analizy skuteczności w eksperymentach numerycznych. Różnorodność przytoczonych tu metod oraz rodzajów obiektów jakie śledzimy pokazują jak szerokim zagadnieniem się zajmujemy.

Następny rozdział pt. *Teoria przetwarzania wideo w kontekście detekcji obiektów* jest prezentacją podstaw teoretycznych grup metod, które wykorzystamy w badaniach. W tej części zostaną przedstawione kroki algorytmów, wykorzystywane w nich wzory a także ich przykładowe działanie. Metody jakie wykorzystamy możemy podzielić na dwie grupy. Pierwsze z nich skupiają się na wykrywaniu obiektów z tła. Efekt ten można uzyskać poprzez detekcję krawędzi lub obszarów o zbliżonej barwie lub teksturze. Algorytmy z drugiej grupy z kolei służą do ich śledzenia wykrytych obiektów. Proces ten opiera się na badaniu zachowania (przesunięcia, orientacji, wielkości) tych obiektów lub wydzielonych z nich

charakterystycznych części. Wybrane i opisane w tym rozdziale metody będziemy wykorzystywać w dalszej części pracy

W kolejnych części pracy pt. *Projekt badawczy* skupimy się na wprowadzeniu do własnej analizy wyselekcjonowanych algorytmów. Znając podstawy teoretyczne z poprzedniego rozdziału wybierzemy kilka metod, które następnie będziemy badać. Część opisanych tam algorytmów pominiemy ze względu na duże podobieństwo w działaniu. Tutaj też zapoznamy się ze środowiskiem pracy, w jakich poszczególne metody zostały zaimplementowane. Istotną częścią tego rozdziału będzie dobór i specyfika dobranego materiału badawczego, wykorzystywanego w kolejnym etapie.

Najistotniejszą częścią pracy są rezultaty badań zgromadzone w rozdziale pt. *Wyniki eksperymentów numerycznych* wraz z *Podsumowaniem i wnioskami*. To właśnie tu sprawdzimy działanie wyselekcjonowanych metod w wybranym przez nas środowisku dla sekwencji wideo. Pozwoli nam to przeanalizować, które sposoby działania poradzą sobie najlepiej z postawionym zadaniem. Poprawność algorytmów sprawdzimy subiektywnie (wzrokowo) ale również zaproponujemy obiektywne (liczbowe) kryterium jakości. Część z tych metod będziemy mogli porównać z rezultatami uzyskanymi w analizie zrealizowanej przez innych autorów a opisanej w rozdziale pt. *Przeglądzie metod detekcji obiektów*.

W ten sposób, począwszy od przeglądu przykładów z literatury, poprzez zapoznanie się z podstawami teoretycznymi oraz sposobem działania wybranych algorytmów a skończywszy na sprawdzeniu ich skuteczności dokonamy analizy sposobów detekcji obiektów w sekwencjach wideo. Sprawdzimy, które metody poradziły sobie lepiej a które gorzej z postawionym problemem.

2. Przegląd metod detekcji obiektów

2.1. Cel i tematyka przeglądu

Zanim zabierzemy się za poznanie teorii i implementację wybranych metod pozwalających na wykrywanie i śledzenie obiektów, przyjrzyjmy się temu co już zostało zrobione. W poniższym opisie skupimy się na kilku wybranych metodach, które pomogą w zrozumieniu aspektów głównego zagadnienia pracy. Zaprezentowanie różnych podejść do tematu pozwoli poszerzyć wiedzę na ten temat oraz uzmysłowić wielorakość rozwiązań i ich skuteczność. Zostały również przedstawione trudności z jakimi zmierzili się autorzy oraz efekty ich poczynąń.

Istotą pracy jest badanie metod wykrywania oraz śledzenia obiektów w sekwencjach wideo z ruchomą kamerą. Możemy wyszczególnić kilka ważnych aspektów poruszanego problemu, a są nimi:

- ruch kamery;
- perspektywa nagrywania;
- rodzaje i ruch śledzonych obiektów.

Dwa spośród trzech wymienionych czynników: ruch kamery oraz perspektywa nagrywania ściśle się ze sobą łączą. Tutaj mamy możliwości, które możemy uprościć do dwóch przypadków: filmowanie z powietrza lub filmowanie z powierzchni ziemi. Powodem rozróżnienia jest różnica w ograniczeniach ruchu kamery (obiektów również). Do pierwszej grupy możemy też zaliczyć np. nagrywanie pod wodą i uogólnić to jako przypadek, w którym zakres ruchu kamery zawiera się w całej sferze. Związane jest to też z ruchem obiektów, które mogą się znaleźć w dowolnej pozycji względem nas. Zauważmy, że z kolei w drugim przypadku ruch wszystkich obiektów zawiera się w uproszczeniu w przestrzeni zawierającej połowę sfery. Wówczas pewna powierzchnia jak np. powierzchnia Ziemi jest tu ograniczeniem dla obiektów. Zarówno dla jednej jak i dla drugiej grupy możemy znaleźć sporo ciekawych przykładów zastosowania. W pierwszym przypadku możemy tu mieć do czynienia np. z obrazami satelitarnymi lub nagraniami z dronów [12] [16]. Jeśli chodzi o obrazy satelitarne to mają jedną, ustaloną perspektywę z góry, która może objąć swoim zasięgiem duże obszary. Teoretycznie ruch takiej kamery jest możliwy we wszystkich kierunkach ale w praktyce ruch sprowadza się do przesunięć równoległych do powierzchni Ziemi. Wynika to jednak tylko i wyłącznie z ich przeznaczenia. Do drugiej grupy możemy zaliczyć np.: o kamerę jaką obsługuje człowiek znajdujący się na powierzchni ziemi czy

kamerę zamontowaną na pojeździe kołowym [15]. Nagrywanie z takiego pojazdu daje możliwość nagrywania z określonej, małej wysokości. Zazwyczaj również obracać kamerę możemy jedynie w prawo lub w lewo, rzadziej dodatkowo w górę. Obrót w dół miałby sens jedynie przy nagrywaniu powierzchni drogi ale to daje nam już kamera ustawiona równolegle do jezdni. Czasem kamera jest ustawiona nieruchomo ale traktujemy ją jako jeden obiekt ruchomy ze względu na operatora lub pojazd na którym jest zamontowana.

Nie mniej istotną kwestią jest jednak to jakie obiekty chcemy wykrywać i śledzić. Mamy tutaj do czynienia zarówno z obiektami statycznymi jak i takimi które się poruszają. Do pierwszej grupy możemy zaliczyć np. drogę (jezdnię) [6] czy też znaki drogowe [15] wykrywane z kamery zamontowanej na pojeździe. Zastosowanie algorytmu śledzenia i wykrywania takich obiektów może np. pomóc w automatycznym sterowaniu samochodem. Oprócz obiektów statycznych możemy jeszcze wyróżnić śledzenie obiektów dynamicznych, często samochodów [11] [13] ale również ludzi [12]. W przypadku wykrywania samochodów możliwa jest np. kontrola przepływu ruchu na drogach, wykrywanie korków itp. W odniesieniu do ludzi może to być przydatne np. do poszukiwań z powietrza w trudnych terenach. U człowieka dodatkową kwestią jest ruch kończyn, a przy dużej dokładności kamery lub dużemu zbliżeniu również mimika twarzy. Może to pozwolić na wykrywanie emocji, a także niewłaściwych zachowań.

Zanim jednak przystąpimy do szukania rozwiązania naszego problemu detekcji to musimy się zapoznać z wszelkimi związanymi z tym utrudnieniami. Na pewno istotne są błędy spowodowane niedokładnością samej metody. Oprócz tego możemy spodziewać się kilku innych przeszkód. Wśród nich możemy wyróżnić takie, które dotyczą zarówno rodzajów obiektów, scenerii i warunków w jakich sekwencje są nagrywane czy możliwości techniczne sprzętu.

Ważną kwestią jest jakość badanego materiału. Chcąc mieć dobre wyniki należałoby zachować wysoką jakość wideo. Niestety czasem jest to trudne do zrobienia, ponieważ pociąga to za sobą transfer wielkiej ilości danych. To z kolei może powodować, że nawet dobrze działający algorytm może potrzebować dużo czasu na przetworzenie sporej ilości informacji i nie działa w czasie rzeczywistym [13] [14]. W tym celu możemy się ograniczyć chociażby do wstępnego odrzucenia części klatek [6] zwłaszcza jeśli różnice pomiędzy nimi są znikome. Pamiętajmy, żeby nie odrzucić ich zbyt dużej ilości bo wówczas działanie algorytmu może się w ogóle nie sprawdzić. Możemy też skupić się w danych klatkach sekwencji jedynie na tych regionach obrazu gdzie wykrywane są obiekty ruchome [13]. Ponadto istotną kwestią jest to co nagrywamy. Równoczesne złożenie perspektywy i rodzaju

obiektu powoduje, że niektóre zmieniają tylko swoje położenie i wielkość jak np. znaki drogowe [15]. Z kolei nagrywając ludzi musimy się liczyć ze zmianą pozycji (chód, bieg, leżenie, siedzenie), ruchem kończyn czy szybkością przemieszczania się [12]. Oczywiście w pierwszym z omawianych przypadków liczbę niezbędnych danych z filmu możemy ograniczyć w znacznie większym stopniu niż w przypadku drugim.

Inną kwestią jest sceneria i warunki pogodowe w trakcie nagrywania. Filmując w mocnym świetle musimy się liczyć z cieniami obiektów, które w takim przypadku powinniśmy umieć rozpoznawać. Również opady deszczu lub śniegu czy też mgła potrafią mocno zakłócić obraz. W tym przypadku gorsza widoczność może się odbić na poprawnym wyodrębnieniu poruszającego się obiektu ze sceny. Ponadto rodzaj podłoża może mieć wpływ na zakłócenia w nagrywaniu z perspektywy naziemnej, np. z kamery zamontowanej na pojeździe [15]. Poruszając się po nierównej drodze można spowodować mniejsze lub większe wstrząsy co może się odbić na wykryciu i śledzeniu obiektów. Podobnie w przypadku kiedy mając kamerę w rękach często trudno jest nam na tyle ją ustabilizować aby nie wprowadzać dodatkowych zakłóceń. Problem ten występuje również w przypadku filmowania z powietrza (domyślnie chodzi o filmowanie dronem), z tym, że wtedy zagrożeniem może być kolizja z innym obiektem latającym, z budynkiem czy też np. zderzenie z ptakiem.

Wszystkie poruszane powyżej okoliczności utrudniające badania powinniśmy zdecydowanie brać pod uwagę przy tworzeniu algorytmu. Na pewno zminimalizowanie wymienionych zakłóceń i zagrożeń wpłynie pozytywnie na jakość nagrania i w konsekwencji pozwoli skupić się na samym wykrywaniu i śledzeniu obiektów. Można zastosować dodatkowe metody, pozwalające wspomagać sam proces niwelowania trudności, jednak nie są one najważniejsze. Kluczowe jest tu poprawne rozpoznanie i śledzenie poruszającego się obiektu.

2.2. Metody wykrywania i śledzenia obiektów

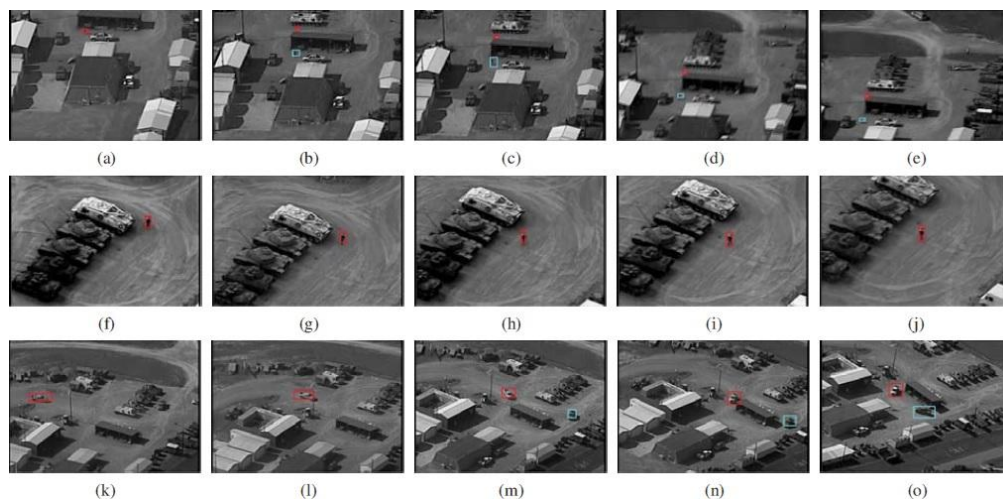
Na bazie wybranych pozycji literaturowych, poniżej zostały wypisane metody, które znalazły zastosowanie do detekcji obiektów. A są nimi:

- metoda śledzenia Kanade-Lucas-Tomasi wraz z metodą RANSAC [6] [12] [13];
- segmentacja MAP (TSMAP) wraz z algorytmem Mean Shift do analizy przemieszczeń [14];
- detektor krawędzi Kirsch z wykorzystaniem deskryptora lokalnej kierunkowości (LDP) oraz lokalnego wzoru binarnego (LBP) [10];
- detektor krawędzi Canny-Deriche wraz z transformacją Hougha i filtrem Ramer-Douglas-Peucker oraz metodą dopasowania Flann [15].

Pierwsze trzy znalazły zastosowanie przy wykrywaniu i śledzeniu obiektów z perspektywy lotu ptaka, natomiast ostatni z nich został użyty do wykrywania i rozpoznawania znaków drogowych z perspektywy kamery zamontowanej na pojeździe. Scharakteryzujemy ich działanie oraz sposób wykorzystania.

2.2.1. Metody Kanade-Lucas-Tomasi i RANSAC

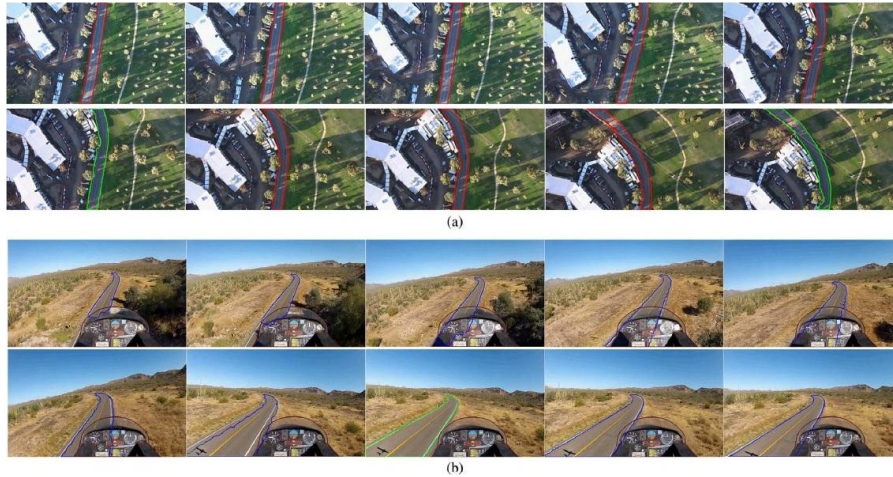
Metoda śledzenia Kanade-Lucas-Tomasi jest często wykorzystywaną metodą do wykrywania ruchu obiektów w sekwencjach wideo. Jest to metoda różnicowa polegająca na porównywaniu kolejnych klatek obrazu. Założeniem algorytmu jest to, że jasność lub kolor obiektów nie zmienia się znacząco z klatki na klatkę co umożliwia detekcję ich przesunięcia. Jest to metoda lokalna, odporna na szумы i drobne zniekształcenia obrazu. Równocześnie bardzo popularne jest użycie wraz z nią metody RANSAC. Algorytm ten wspomaga detekcję poprzez eliminację błędów w zbiorach gdzie większość próbek jest poprawna [7]. Użycie tych dwóch metod równocześnie może być pomocne przy wykrywaniu ruchu z perspektywy lotu ptaka. Jednym z zastosowań jest wykrywanie ruchu małych obiektów (np. ludzi lub samochodów) na podstawie znajdowanych punktów kluczowych [12]. Takimi punktami mogą być np. narożniki obiektów. Są to reprezentatywne punkty na obrazie, które pozwalają nam na detekcję ruchu. Ich przesunięcie możemy zauważyć biorąc pod uwagę klatki sekwencji z określonego okna czasowego co zostało pokazane na rysunku 2.1. W zależności od tego jaki obiekt śledzimy to jego skuteczność może sięgać nawet blisko 100%.



Rys 2.1. Wykrywanie kolejnych obiektów na klatkach, (a)-(e), (f)-(j), (k)-(o) [12].

Innym zastosowaniem jest wykrywanie drogi zarówno z perspektywy lotu ptaka jak i z perspektywy kamery zamontowanej na pojeździe [6]. W przypadku wykrycia drogi nasze działanie właściwie skupia się na detekcji jej krawędzi co zostało zobrazowane na rysunku

2.2. W pierwszym przypadku możemy badać np. natężenie ruchu na drodze a drugi może się przydać np. przy opracowywaniu automatycznego sterowania dla pojazdu kołowego. Wykrycie drogi w ten sposób pozwala się skupić tylko i wyłącznie na interesującym nas fragmencie obrazu, zawartym między brzegami drogi. Kolejne operacje np. wykrywanie pojazdów na drodze nie muszą być już wykonywane na całym obrazie.



Rys 2.2. Wykrycie jezdni i naniesienie jej krawędzi na obraz oryginalny: z perspektywy z lotu ptaka (a) i z perspektywy kamery zamontowanej na pojeździe (b) [6].

Stosując metody wykrywania i śledzenia obiektów w czasie rzeczywistym musimy uwzględnić problem przepustowości danych. Jednak jak już zostało wspomniane wcześniej nie ma konieczności badania całych obrazów ale np. na jezdnię. W ten sposób możemy skierować nasz algorytm na obszary, które nas interesują (ROI – Regions of Interest) [13]. Stanowi to kompromis pomiędzy dokładnością metody a ograniczeniami przesyłania danych. Właśnie w ten sposób możemy ułatwić sobie wykrywanie obiektów np. samochody będą wyszukiwane tylko na drodze.



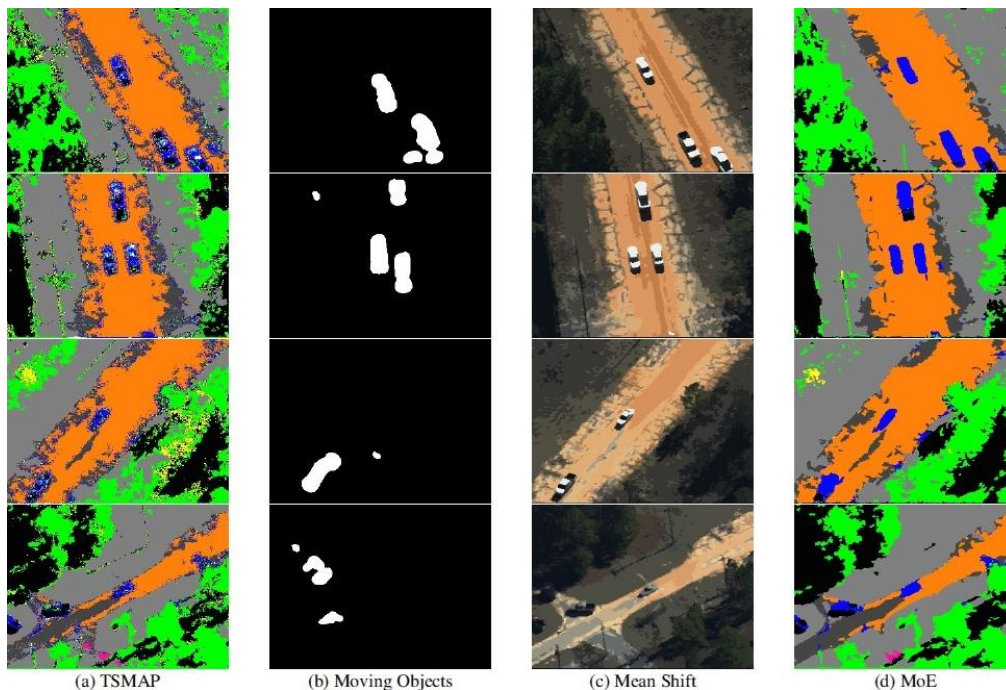
Rys 2.3. Wykrycie i wyodrębnienie z obrazu poruszających się samochodów [13].

2.2.2. Segmentacja MAP (TSMAP) z algorytmem Mean-Shift

Segmentacja MAP (TSMAP) wraz z użyciem średniego przesunięcia obrazu jest inną z metod pozwalającą na wykrycie poruszających się obiektów [14]. Uściślając – w przytoczonym przykładzie są używane metody segmentacji z różnymi modyfikacjami: segmentacja wykrywająca tylko obiekty poruszające się, segmentacja z wykorzystaniem metody Mean-Shift oraz segmentację TSMAP do rozróżnienia i rozgraniczenia obiektów znajdujących się na obrazie. Została też wykorzystana metoda złożona z trzech powyższych.

Segmentacja wykrywająca poruszające się obiekty (Moving Objects Detection) realizuje „wydobycie” interesujących nas fragmentów sekwencji. Badając kolejne klatki znajdujemy różnice pomiędzy nimi. W efekcie wyróżnione zostają jedynie interesujące nas obiekty. Na rysunku 2.4 w podpunkcie b) zostały przedstawione binarne obrazy wynikowe – poruszające się pojazdy zostały oznaczone kolorem białym.

Segmentacja opierająca się na metodzie Mean-Shift do swojego działania wykorzystuje kolor, profil krawędzi i teksturę obrazu.

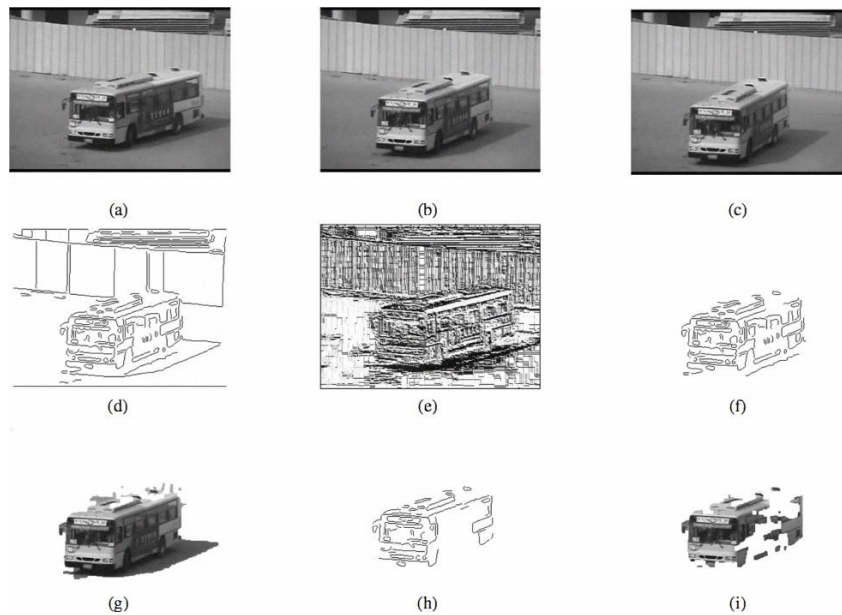


Rys 2.4. Różne metody segmentacji: a) TSMAP, b) segmentacja poruszających się obiektów (Moving Objects), c) Mean Shift, d) złożenie tych trzech metod (Mixture of Experts - MoE) [14].

2.2.3. Operator Kirsch wraz z LDP i LBP

Detektor krawędzi Kirsch jest z kolei jedną z metod opierających się na maskach wielkości 3x3 uwzględniające w środku interesujący nas obiekt (piksel). Podobnie jak inne tego typu

metody jak. Np. Prewitt, Sobel czy Scharr występują różne filtry dla różnych kierunków krawędzi. W odróżnieniu jednak od powyższych sposobów, detektor Kirsch w swoich obliczeniach uwzględnia wszystkie piksele sąsiadujące bezpośrednio z tym, który nas interesuje. Po wykryciu krawędzi wspomagamy się jeszcze lokalnym deskryptorem (LDP) [10]. W ten sposób uzyskujemy kompletne informacje o konturach na obrazie – te najbardziej wyraźne lokalnie na obrazie możemy pogrupować, tworząc zwarte struktury geometryczne (obiekty) i potraktować jako segmenty na naszym obrazie. Podobnie jak inne z przytoczonych tu metod, służy ona jedynie do wyszczególnienia tych struktur na obrazach. Do śledzenia ich musimy się posłużyć dodatkową metodą. W tym celu każdorazowo korzystamy z 3 ostatnich klatek sekwencji wideo i wykorzystujemy lokalny wzór binarny (LBP). Skuteczność tej metody dla badanego obiektu wynosi średnio ok. 90% co jest bardzo dobrym wynikiem.

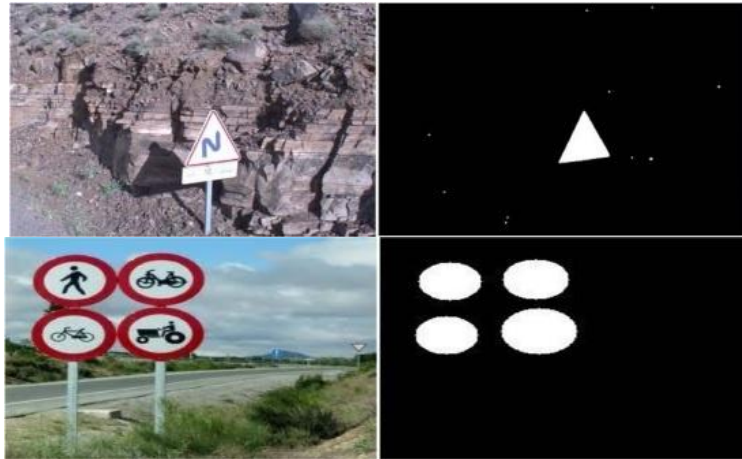


Rys 2.5. Wykrycie ruchu jadącego autobusu przy pomocy detektora Kirsch wraz z LDP i LBP[104].

2.2.4. Algorytm Canny oraz transformacja Hougha

Detektor Krawędzi Canny jest obecnie jedną z popularniejszych metod wykrywania krawędzi. Jest to wielostopniowy algorytm w którym zawiera się kolejno: wstępne przetwarzanie filtrem Gaussa (redukcja zakłóceń), szukanie natężenia gradientu obrazu (poszukiwanie kierunku przebiegu krawędzi), zapewnienie ciągłości znalezionych krawędzi oraz progowanie zapewniające usunięcie nieistotnych krawędzi. Wykrycie konturów obrazu w ten sposób może być przydatne np. w wykrywaniu znaków drogowych [15]. W tym celu do znalezionych w ten sposób krawędzi wykorzystamy transformację Hougha i filtr Ramer-Douglas-Peucker,

które pozwolą nam na wykrycie okręgów i prostych tworzących trójkąty na obrazie. W ten sposób mamy już wyodrębnione kształty, które możemy porównywać ze wzorcami w bazie.



Rys 2.6. Efekty działania transformacji Hougha i filtru Ramer-Douglas-Peucker [15].

W zależności od konkretnego znaku drogowego możemy zauważyć niewielkie rozbieżności w kwestii wykrywania i rozpoznawania znaków drogowych. O ile wszystkie kontury są dla każdej grupy znaków jednakowe o tyle ich centralna część jest zróżnicowana. Wykrywalność była na poziomie 90-93% natomiast rozpoznawanie na poziomie 80-87% co jest dobrym wynikiem. Należy tu również pamiętać o tym, że znaki mogą nie zostać rozpoznane np. ze względu na ich zniszczenie.

2.3. Podsumowanie przeglądu literaturowego

Powyższy przegląd literaturowy pozwala nam na zobrazowanie zagadnień poruszanych w tej pracy i ich użycie w wyszukanych pozycjach. Pozwala nam to na lepsze zrozumienie działania poszczególnych metod. Daje również możliwość zastąpienia wymienionych innymi, podobnymi algorytmami. Dzięki temu będziemy mieli możliwość ich poprawnej implementacji i porównanie ich działania w problemie wykrywania i śledzenia obiektów w sekwencjach wideo przy ruchomej kamerze.

3. Teoria przetwarzania wideo w kontekście detekcji obiektów

3.1. Cel przetwarzania

Przetwarzanie obrazów możemy określić jako wykonanie operacji mających na celu uwydatnienie istotnych dla nas cech danego obrazu. Obraz jest to nic innego jak rodzaj sygnału, a w tym przypadku jest to sygnał dwuwymiarowy. Znaczny rozrost jego wielkości powoduje, że liczba danych jakie możemy uzyskać może być bardzo duża. Część z nich jednak może być dla nas w ogóle nieprzydatna. Przetwarzanie obrazu pozwala nam na odrzucenie tych zbędnych i pozostawieniu tylko tych istotnych dla danego problemu.

Najmniejsze jednolite elementy obrazu to piksele. To one niosą nam informacje o obrazie takie jak kolor czy odcień szarości. Oprócz badania właściwości konkretnego piksela istotne mogą być też właściwości jego sąsiadów co pozwala na znacznie szerszą analizę.

Współcześnie przetwarzanie obrazów znajduje bardzo szerokie zastosowanie zarówno dla inżynierów jak i dla ludzi nie związanych na co dzień z techniką. Osoba z pierwszej grupy wykorzysta to np. do opracowania systemu automatycznej kontroli jakości wyrobów w fabryce na podstawie zdjęć produkowanych elementów. Z kolei osoba z drugiej grupy może być zachwycona nowym filtrem graficznym w aplikacji na portalu społecznościowym. Bez względu na to czy zastosowanie jest mniej lub bardziej poważne to możemy zauważyć ciągły rozwój tej gałęzi technologii.

My skupiamy się na tej bardziej poważnej części. Będziemy analizować sekwencje wideo pod kątem wykrycia i śledzenia poruszających się obiektów. Sekwencja wideo jest niczym innym jak zbiorem kolejnych obrazów (klatek). Problem nie dotyczy pojedynczych obrazów zatem całą naszą uwagę ukierunkujemy na większą ilość danych. W kolejnych klatkach będziemy analizować występowanie konkretnych obiektów opierając się głównie na ich kształcie. Tu możemy polegać na metodach wykrywania krawędzi lub metodach segmentacji. Następnie będziemy badać zmienność ruchu obiektów względem poprzednich klatek w danej sekwencji – zatem będziemy obserwować przemieszczanie się kształtów czy też konturów obiektów. Zatem nasze metody podzieliliśmy właśnie na takie dwie grupy – jedne skupiają się na wykrywaniu obiektów a pozostałe na ich śledzeniu. Poniżej przedstawiamy wybrane algorytmy. Planowo te opisane miały służyć do części badawczej – zostały wykorzystane te w całości poprawnie zaimplementowane. Dalszy podział tych metod na określone grupy według sposobu działania daje wielorakość możliwości realizacji rozwiązania określonego problemu.

3.2. Metody wykrywania obiektów

3.2.1. Metody wykrywania krawędzi i kształtów

Metody wykrywania krawędzi opierają swoje działanie na znajdowanie na obrazie punktów gwałtownej zmiany luminacji (intensywności). Można ze sporym prawdopodobieństwem przypuszczać, że właśnie takie punkty znajdują się na granicy dwóch różnych obszarów (obiektów) umożliwiając ich wyraźne rozgraniczenie.

W wykrywaniu krawędzi zastosowanie znajdują operatory przybliżające pochodne w określonych kierunkach. Rozmiar takiego operatora to zazwyczaj kwadrat 3x3 piksele, gdzie wartość jest obliczana dla piksela środkowego na podstawie wartości luminacji sąsiadów. Możemy w ten sposób uzyskać 8 różnych masek o wymiarach 3x3 – jesteśmy w stanie wyznaczyć krawędzie nachylone pod kątem od 0° do 315° z krokiem 45° . Operatory te przyjęły nazwy od nazwisk osób które jako pierwsze je zastosowały. A byli to: Judith Prewitt, Irwin Sobel, Hanno Scharr, Russell Kirsch. Różnice między nimi wynikają z innych wartości wag danych pikseli, co zostało zobrazowane na rys. 3.1. Miało to oczywiście wpływ na uzyskiwane efekty np. operator Sobela (wagi: 1,2,1) lepiej wygładza niż operator Prewitta (wagi: 1,1,1), z kolei maska Scharra (wagi: 3, 10, 3) została zoptymalizowana pod kątem właściwości kierunkowych. Operator Kirscha jako jedyny bierze pod uwagę wartości wszystkich ośmiu sąsiednich pikseli. Ponadto do tej grupy możemy jeszcze zaliczyć Krzyż Roberta, który właściwie możemy potraktować jako maskę 2x2. Dzieje się ze względu na rozmieszczenie pikseli o zerowych wagach (rys. 3.2).

+1	0	-1
+1	0	-1
+1	0	-1

a

+1	0	-1
+2	0	-2
+1	0	-1

b

+3	0	-3
+10	0	-10
+3	0	-3

c

+5	-3	-3
+5	0	-3
+5	-3	-3

d

Rys. 3.1. Przykładowe maski dla kąta 180° : Prewitt (a), Sobel (b), Scharr (c), Kirsch (d).

-1	0	0
0	+1	0
0	0	0

e

-1	0
0	+1

f

Rys. 3.2. Krzyż Roberta (e) możemy przedstawić w nieco uproszczonej formie (f).

Do innej grupy możemy zaliczyć algorytmy wykrywania krawędzi Canny’ego i Deriche’a [2] [8]. Nazwy pochodzą od osób, które je opracowały: Johna Canny’ego i Richarda Deriche’a. Są to kilkukrokowe algorytmy a w ich działaniu możemy wyróżnić cztery etapy. Różnic pomiędzy nimi możemy doszukać się jedynie w dwóch pierwszych krokach. Zasadniczo algorytm Deriche’a został stworzony na podstawie algorytmu Canny’ego, stąd w literaturze można się też spotkać z nazwą algorytmu Canny-Deriche. Kolejne etapy algorytmów to:

- redukcja szumu – wykorzystujemy filtr Gaussa (Canny) lub filtr o nieskończonej odpowiedzi impulsowej (Deriche) [4]. Jest to konieczne ze względu na dużą czułość pod względem obecności szumu. Wynikowy obraz jest nieco rozmazany, za to pozbawiony jest znaczących pojedynczych zakłóceń;

- szukanie gradientu obrazu – w tym etapie zajmujemy się wykryciem kierunku skierowania krawędzi. W praktyce sprowadza się to do zaokrąglania miar kątów i wynikiem jest ograniczenie ilości kierunków do czterech (w pionie, poziomie i po obu przekątnych). Potrzebujemy wartości pierwszych pochodnych dla kierunku poziomego (G_x) oraz pionowego (G_y). Możemy je uzyskać z detektorów krawędzi takich jak np. Krzyż Robertsa, Prewitt czy Sobel. Szybkość narastania krawędzi G oraz kierunek Θ możemy obliczyć za pomocą wzorów:

$$G = \sqrt{G_x^2 + G_y^2} \quad (3.1)$$

$$\Theta = \arctan\left(\frac{G_y}{G_x}\right) \quad (3.2)$$

- usuwanie pikseli o pośrednich wartościach – celem tego kroku jest zapewnienie ciągłości krawędzi przy równoczesnym maksymalnym jej zwężeniu. Po użyciu różnych detektorów możemy często zauważyć, że krawędzie i tło nie mają wyraźnego rozgraniczenia ze względu na stopniowanie pikseli. Wynikać to może np. z rozmazania obrazu pierwotnego. Naszym celem jest usunięcie tych „przejściowych” pikseli. Efektem jest linia ciągła złożona z pojedynczych pikseli;

- progowanie z histerezą – w ostatnim kroku skupiamy się na usuwaniu nieistotnych krawędzi. Do istniejących krawędzi dołączane są kolejne piksele aż do osiągnięcia dolnego progu wykrywania. Pozwala to na uniknięcie rozdzielania krawędzi w miejscu słabszego kontrastu.

Algorytm Canny’ego jest obecnie jednym z najczęściej stosowanych narzędzi do wykrywania krawędzi na obrazie. Mimo, że został opracowany już w 1986 r. (algorytm Deriche rok później) to nadal trudno doszukać się bardziej skutecznych metod.

Innym sposobem jest zastosowanie transformaty Haara [3] [5]. Wówczas Obraz dzielimy na bloki 2x2 piksele a ich wartości możemy obliczyć ze wzorów [17]:

$$h_0 = \frac{1}{2}(a_{11} + a_{10} + a_{01} + a_{00}) \quad (3.3)$$

$$h_x = \frac{1}{2}(a_{11} + a_{10} - a_{01} - a_{00}) \quad (3.4)$$

$$h_y = \frac{1}{2}(a_{11} - a_{10} + a_{01} - a_{00}) \quad (3.5)$$

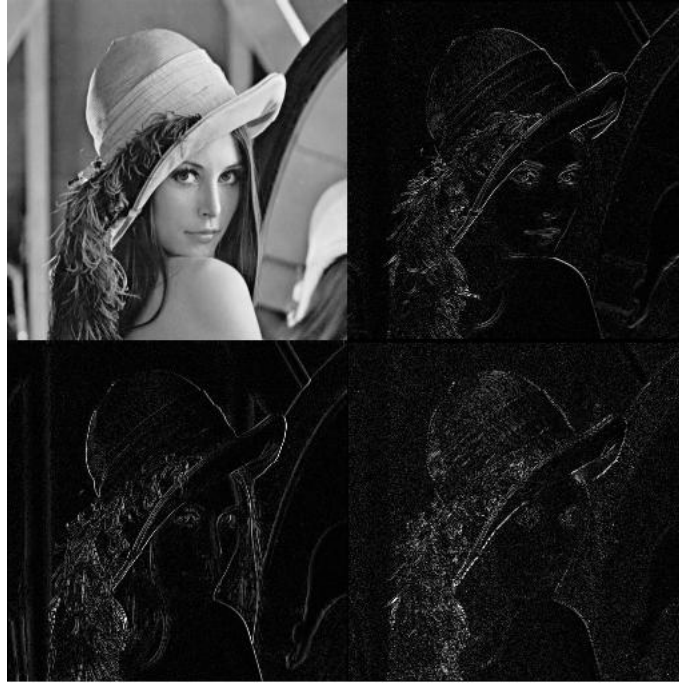
$$h_z = \frac{1}{2}(a_{11} + a_{10} - a_{01} + a_{00}) \quad (3.6)$$

Konsekwencją tych działań jest podzielenie obrazu pierwotnego (przykład na rys. 3.3) na cztery części odpowiadającej odpowiednim filtrom (efekt na rys. 3.4):

- lewa górna część – 2-D filtr dolnoprzepustowy;
- prawa górna część – filtr poziomy górnoprzepustowy i pionowy dolnoprzepustowy (uwydatnione krawędzie poziome);
- lewa dolna część – filtr poziomy dolnoprzepustowy i pionowy górnoprzepustowy (uwydatnione krawędzie pionowe);
- prawa dolna część – 2-D filtr górnoprzepustowy.



Rys. 3.3. Oryginalny obraz *Lena.png* [18].



Rys. 3.4. Obraz *Lena.png* po dekompozycji falką Haara [18].

3.2.2. Metody segmentacji

Segmentacja, w przeciwieństwie do wykrywania krawędzi, nie jest skupiona na konturach obiektów ale na ich wnętrzu. Celem jest połączenie w jeden obszar pikseli o podobnych właściwościach takich jak np. kolor czy poziom jasności. W ten sposób możemy wyodrębnić z tła poszczególne obiekty. Możemy wyróżnić kilka metod segmentacji uwzględniając konkretne działania na obszarach [1]:

- przez podział obszaru – metoda polega na dzieleniu dużych obszarów na coraz mniejsze, charakteryzujące się odmiennymi właściwościami od tego z którego zostają wydzielone;
- przez rozrost obszaru – metoda polega na dołączaniu pikseli o podobnych właściwościach. Piksele przed ewentualnym przyłączeniem są testowane pod względem stopnia podobieństwa. Kryteria przyłączenia mogą być statyczne (np. ustalenie stałego progu tolerancji) lub dynamiczne (np. wyznaczenie uśrednionych wartości z obszaru lub relacji przestrzennych);
- przez łączenie podziałów – sąsiadujące ze obszary o wystarczająco podobnych właściwościach zostają połączone w jeden większy;
- hybrydy powyższych metod – możliwe jest połączenie np. metody podziału z rozrostem - samo dzielenie kolejnych dużych obszarów na stałą ilość mniejszych może nie uwzględniać wystarczającego podobieństwa między sąsiednimi, oddzielonymi fragmentami;
- progowanie;

- segmentacja wododziałowa.

Najprostszą spośród metod segmentacji obrazu jest progowanie. Mając do dyspozycji obraz w odcieniach szarości otrzymujemy obraz binarny. Często jest wykorzystywany w innych algorytmach np. w detektorze krawędzi Canny'ego [2].

W najprostszym wydaniu każdy z pikseli obrazu ma określoną jasność $I_{i,j}$ i ustalamy określony próg T . Każdy piksel mający jasność mniejszą niż określony próg T stanie się pikselem czarnym a te z jasnością powyżej progu staną się pikselami białymi. Do rozwiązania pozostaje kwestia odpowiedniego ustawienia progu. Tu możemy wyróżnić sześć aspektów o które możemy się oprzeć dobierając próg:

- kształt histogramu, np. analiza szczyt, dolin i krzywizn w nim występujących;
 - bazowanie na klasteryzacji – próbki poziomu szarości są skupione w dwóch częściach: tło oraz pierwszy plan (obiekt);
 - bazowanie na entropii np. regionów tła i pierwszego planu;
 - bazowanie na atrybutach obiektu – wyszukiwanie miary podobieństwa między obrazem oryginalnym i binaryzowanym biorąc pod uwagę m. in. jego kształt;
 - zastosowanie wyższego rzędu rozkładu prawdopodobieństwa lub korelacji pomiędzy pikselami;
 - metody lokalne – dostosowanie wartości progu bazując na lokalnych cechach obrazu.
- W tych metodach wartości progu dla pikseli całego obrazu nie są jednakowe.

Innym rodzajem segmentacji obrazu, który warto wyróżnić jest segmentacja wododziałowa. W tej metodzie obraz traktowany jest jako powierzchnia topograficzna. Piksele o najwyższej intensywności są traktowane jako linie przełomowe i stanowią granice regionu. „Woda” umieszczona na pikselach wspólnej linii działu wodnego płynie w dół do wspólnej minimalnej lokalnej intensywności. Piksele odprowadzone do wspólnego minimum stanowią jeden segment.

Jednym z najbardziej powszechnych algorytmów segmentacji wododziałowej został wprowadzony przez Fernanda Meyera na początku lat 90-tych. Możemy wyróżnić kolejne kroki tego algorytmu:

- wybieramy punkty skąd będziemy zaczynać „zalewanie”, każdy z nich ma inną etykietę;
- wszystkich sąsiadów takich pikseli wprowadzamy do kolejki priorytetowej; ich priorytet odpowiada wielkość gradientu piksela;
- wydobywamy piksela o najniższym priorytecie; jeśli wszyscy jego sąsiedzi mają jednakową etykietę to jemu nadawana jest identyczna; Wszyscy nieoznakowani sąsiedzi niebędący w kolejce priorytetowej są w niej umieszczane;

- powtarzaj poprzedni krok aż kolejka priorytetowa będzie pusta.

3.3. Metody śledzenia obiektów

SURF

SURF (Speeded Up Robust Features) jest opatentowanym lokalnym detektorem i deskryptorem [9]. Jest inspirowany częściowo przez transformację o niezmienniej skali (Scale-Invariant Feature Transform - SIFT), jednak jest od niej kilka razy szybsza i bardziej odporna na różne transformacje obrazu.

Całość opiera się na punktach charakterystycznych, wykrywanych w drodze obliczeń składających się z obliczenia wyznacznika Hessianu. Funkcja deskryptora bazuje na sumie odpowiedzi falki Haara wokół interesującego nas punktu. Deskryptory mogą być użyte do rozpoznawania obiektów oraz śledzenia ich. Algorytm można podzielić na 3 etapy:

- wykrywanie punktów charakterystycznych;
- badanie sąsiedztwa;
- dopasowywanie i łączenie.

SURF wykorzystuje filtry w kształcie kwadratu jako przybliżenie wygładzania Gaussa. Filtrację obrazu przyspiesza użycie wzoru:

$$S(x, y) = \sum_{i=0}^x \sum_{j=0}^y I(i, j) \quad (4.1)$$

SURF wykorzystuje również detekcję „plam” (regionów punktów o zbliżonych właściwościach). Bazuje on na Hesjanie:

$$H(p, \sigma) = \begin{pmatrix} L_{xx}(p, \sigma) & L_{xy}(p, \sigma) \\ L_{yx}(p, \sigma) & L_{yy}(p, \sigma) \end{pmatrix} \quad (4.2)$$

Wyznacznik takiej macierzy służy do obliczeń zmienności otoczenia wokół danego punktu na obrazie. Punkty o największej wartości wyznacznika są właśnie poszukiwanymi przez nas punktami.

4. Projekt badawczy

4.1. Założenia projektowe

Badanie wytypowanych metod sprowadzi się do sprawdzenia ich działania na własnoręcznie nagrany materiał wideo. Metody te zawsze będą dobierane w pary (wykrywanie-śledzenie). My na podstawie obiektywnych kryteriów oraz subiektywnych odczuć ocenimy jak radzą sobie poszczególne algorytmy. Materiał wideo przedstawia proste obiekty poruszające się po drewnianym podłożu. Całość została napisana w jednym skrypcie, działającym w środowisku Matlab (w naszym przypadku wersja R2011b). Skrypt ten zawiera:

- część poświęconą wczytaniu sekwencji wideo, wydzieleniu osobnych klatek i przekonwertowanie ich do postaci w odcieniach szarości;
- zaimplementowane lub wywoływane metody wykorzystywane w badaniach (częściowe wykorzystanie gotowych funkcji);
- instrukcje pozwalające na wyświetlanie obrazów wynikowych oraz danych liczbowych pomocnych przy interpretacji wyników.

4.2. Dobór metod

Spośród metod wykrywania i śledzenia obiektów do analizy zostało wytypowanych kilka z nich. Ze względu na trudności z implementacją niektórych metod, zostały one ostatecznie odrzucone – pozostały jedynie te w pełni poprawnie działające. Spośród metod wykrywania obiektów wybieramy:

- krzyż Robertsa – metoda wykrywania krawędzi z wykorzystaniem masek 2x2;
- operator Prewitta – metoda wykrywania krawędzi z wykorzystaniem masek 3x3;
- algorytm Canny'ego – wieloetapowy detektor krawędzi;
- progowanie – przedstawiciel metod segmentacji obrazu.

Natomiast spośród metod śledzenia obiektów wytypowaliśmy:

- metodę różnicowa – najzwyczajniejsze wyznaczenie różnicy obrazów;
- algorytm SURF – śledzenie punktów charakterystycznych znajdujących przez ten algorytm.

4.3. Dobór materiału badawczego

Do analizy badania skuteczności wybranych metod zostały nagrane krótkie sekwencje wideo przedstawiające małe, proste elementy na nieco zróżnicowanym kolorystycznie tle. Obiektami tymi są: niebieska piłeczka do unihokeja oraz 2 sklezione ze sobą patyczkami lizaki w jaskrawych opakowaniach. Oba obiekty różnią się między sobą kolorem kształtem oraz możliwym zakresem ruchu w poszczególnych kierunkach. Tłem są panele podłogowe w odcieniach koloru brązowego (tło niestety nie jest jednolite).



Rys. 4.1. Klatka z pierwszego z użytych przykładów – „lizaki”.



Rys. 4.2. Klatka z drugiego z użytych przykładów – „piłka”.

5. Wyniki eksperymentów numerycznych

W tym rozdziale zajmiemy się analizą metod wykrywania i śledzenia obiektów. Zanim jednak przejdziemy do najważniejszej części czyli zbadania działań par metod (wykrywania + śledzenie) zajmijmy się analizą niektórych algorytmów. Ma to na celu sprawdzenie działania metod, które mają zmienne parametry.

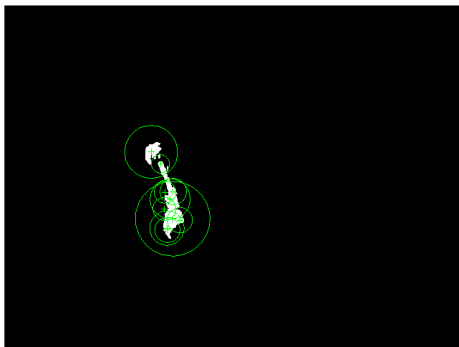
Do takich metod należą:

- progowanie – zakres progu będzie liczbą całkowitą z przedziału $[0; 255]$;
- detektor krawędzi Canny – tutaj będziemy badać zachowanie metody przy zmianie 2 parametrów, a są to: wartość progowania (istotne przy dołączaniu kolejnych krawędzi do najlepiej wyznaczonych) – ten parametr będzie z przedziału $(0; 1)$ oraz *sigma* (odchylenie standardowe funkcji Gaussa) – tu wybierzemy kilka liczb naturalnych dodatnich.

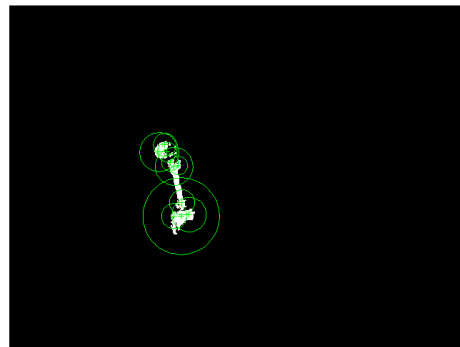
Ponadto w algorytmie SURF możemy określić ilość punktów charakterystycznych (minimów lokalnych). Dla łatwości obliczeń ustalamy ten parametr jako stały równy 10. Pozostałe metody (detektory krawędzi Roberts i Prewitt oraz metoda różnicowa) zostają w swoim podstawowym kształcie.

5.1. Przykładowe obrazy wynikowe

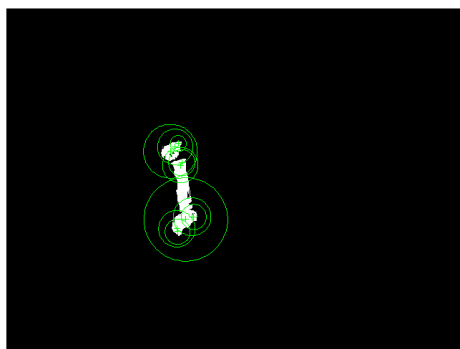
Po wybraniu kilku metod możemy rozpocząć generowanie obrazów, na podstawie których będziemy mogli sposób obiektywny ocenić skutecznych wyselekcjonowanych algorytmów. Poniżej przedstawionych zostało kilka przykładów ich działania na poszczególnych klatkach nagrywanych sekwencji wideo. Możemy zaobserwować zmiany jakie zachodzą na skutek poruszania się obiektu oraz kamery. Może to wpłynąć także na inne aspekty jak np. zmiana oświetlenia sceny. W kolejnym podrozdziale zajmiemy się szczegółową analizą.



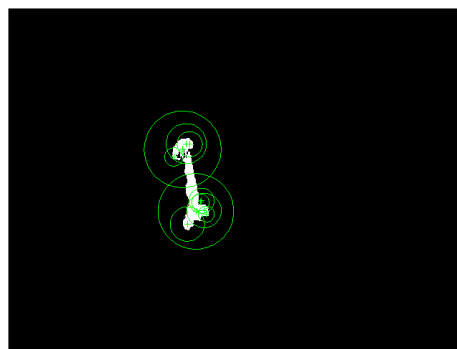
(a)



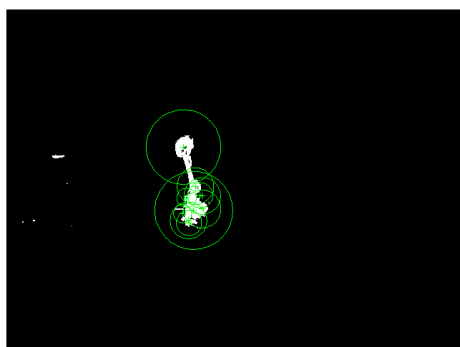
(b)



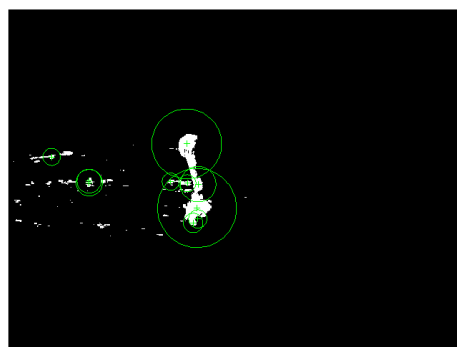
(c)



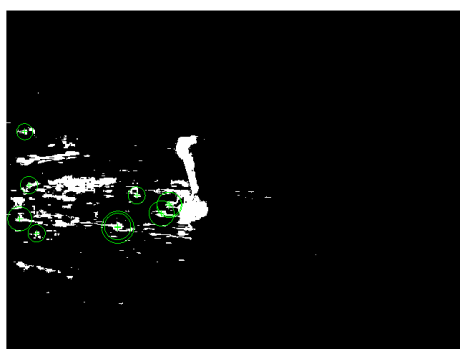
(d)



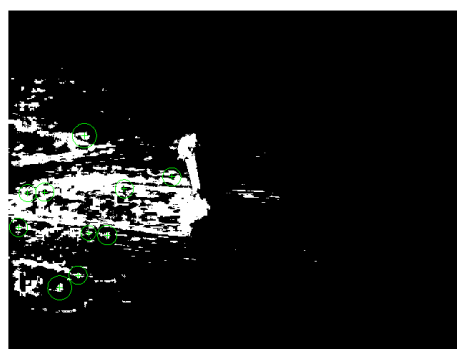
(e)



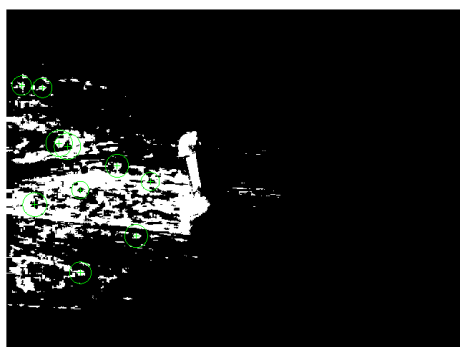
(f)



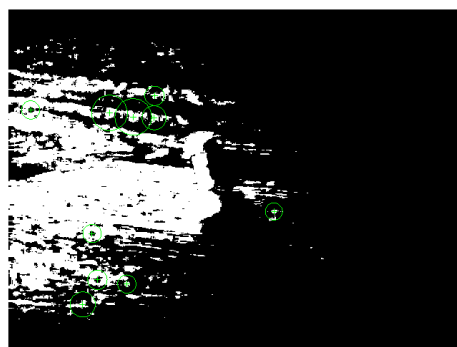
(g)



(h)

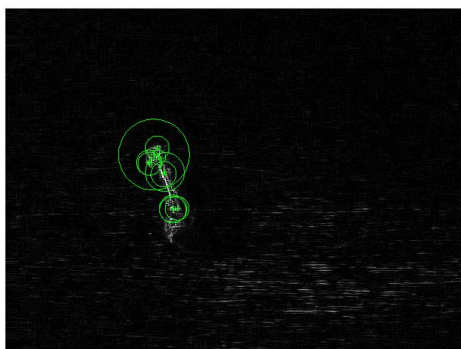


(i)



(j)

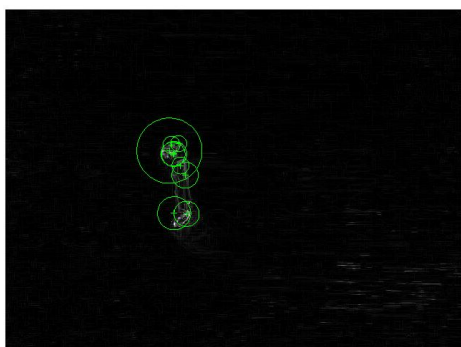
Rys. 5.1. Progowanie z algorytmem SURF (a)-(j).



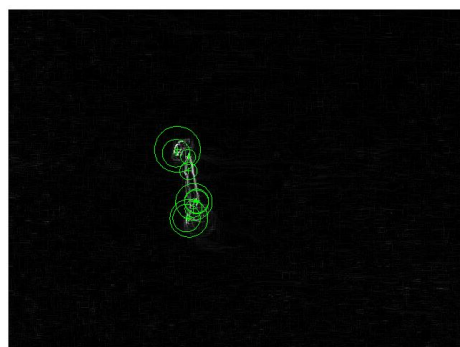
(a)



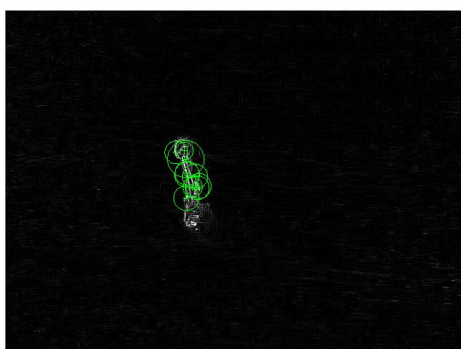
(b)



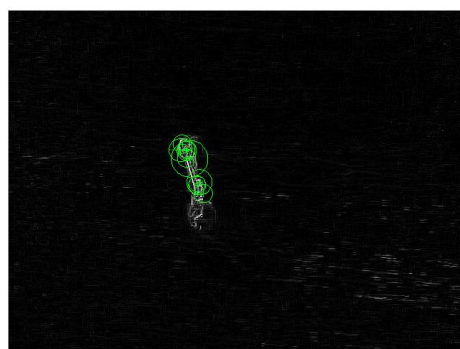
(c)



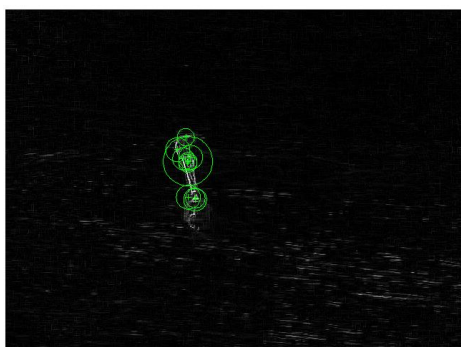
(d)



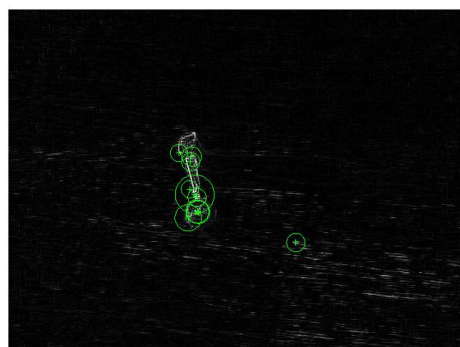
(e)



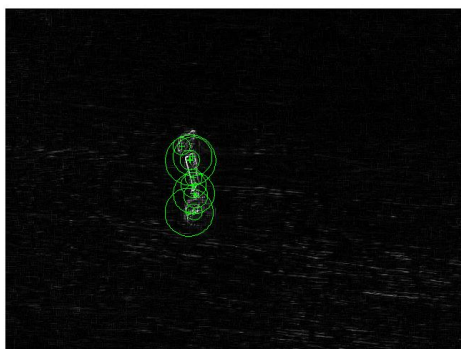
(f)



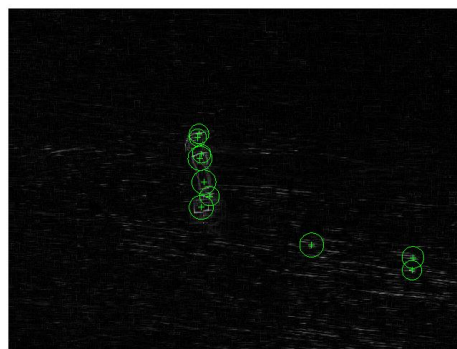
(g)



(h)

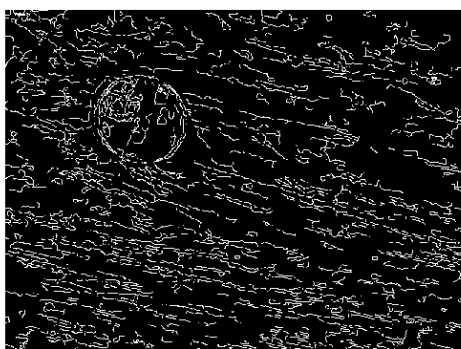


(i)



(j)

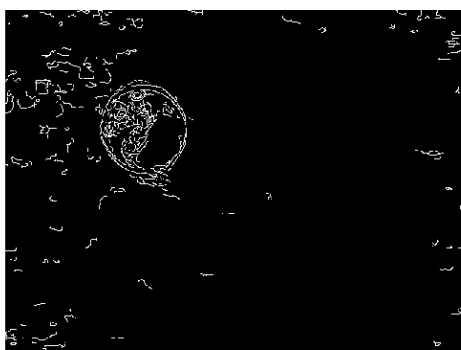
Rys. 5.2. Detektor krawędzi Roberts z algorytmem SURF (a)-(j).



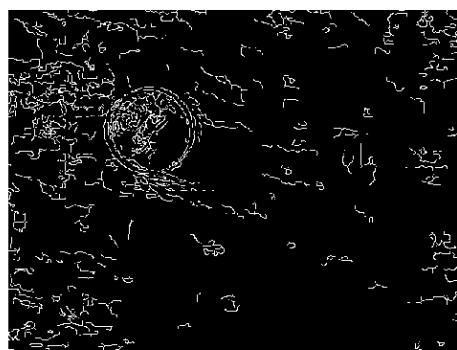
(a)



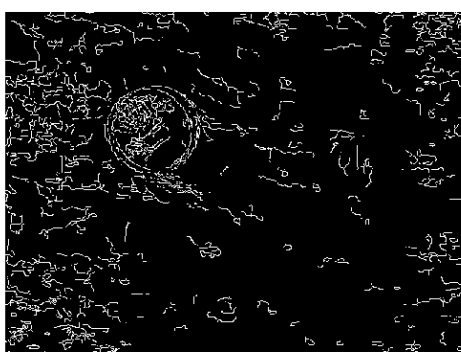
(b)



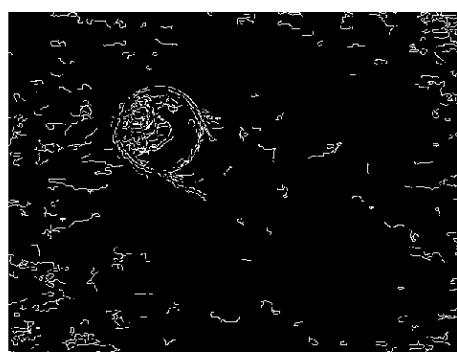
(c)



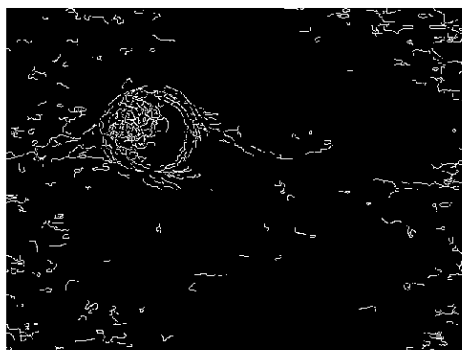
(d)



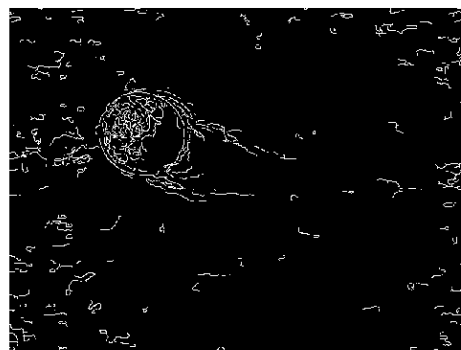
(e)



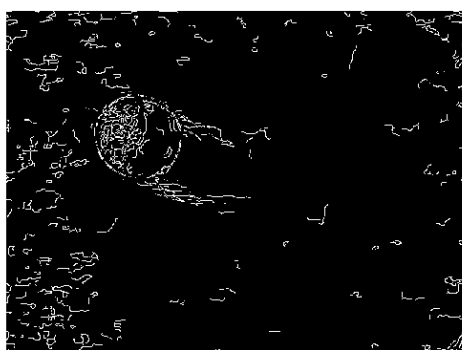
(f)



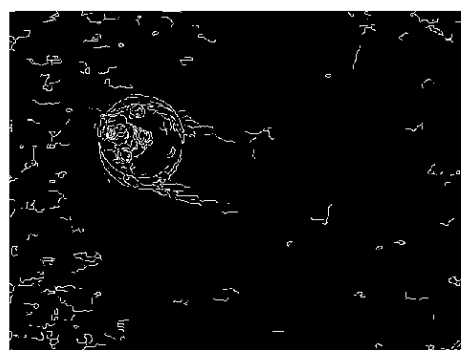
(g)



(h)

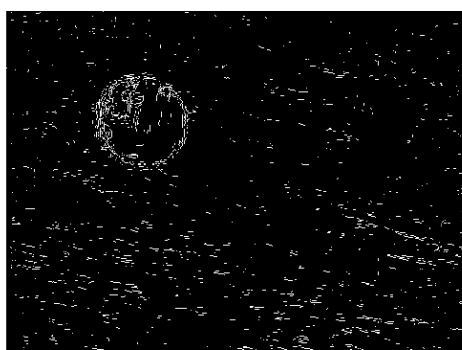


(i)

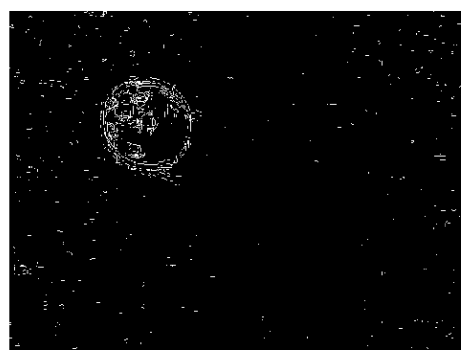


(j)

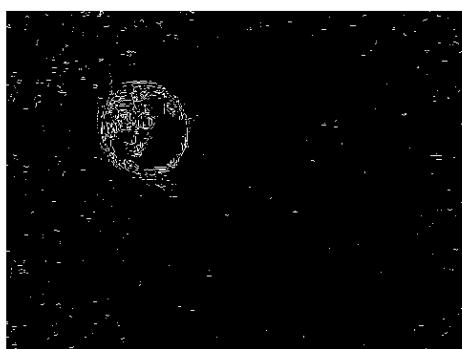
Rys. 5.3. Detektor krawędzi Canny z metodą różnicową (a)-(j).



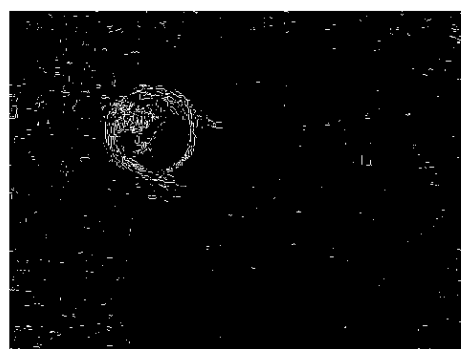
(a)



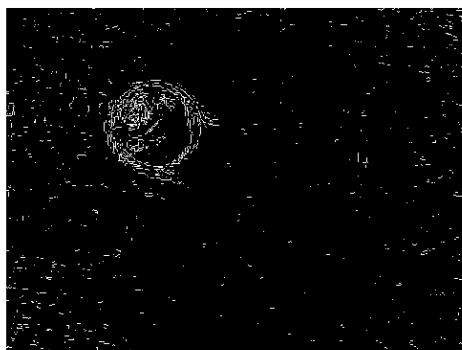
(b)



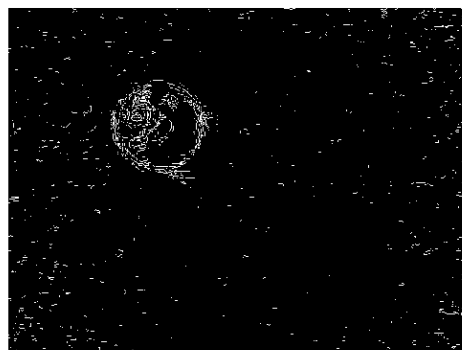
(c)



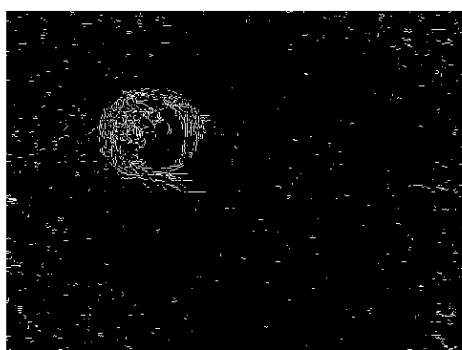
(d)



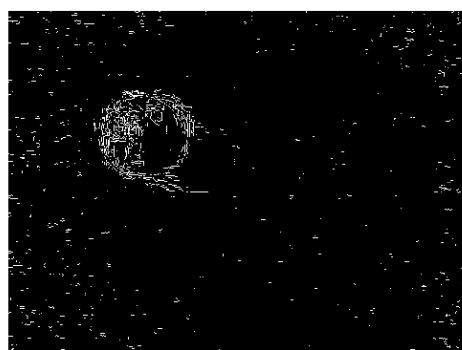
(e)



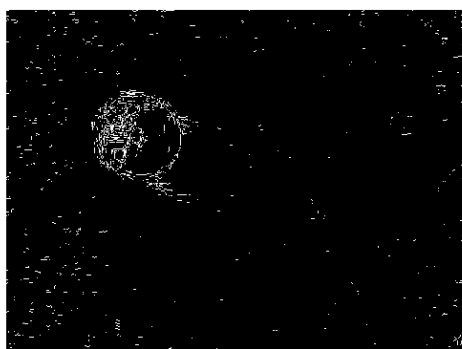
(f)



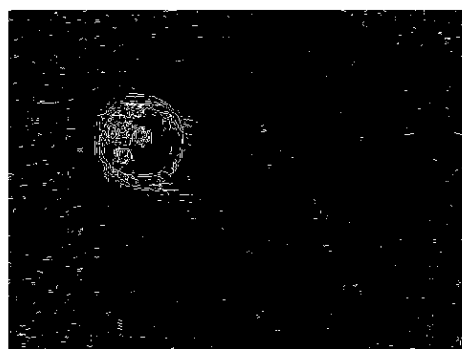
(g)



(h)



(i)



(j)

Rys. 5.4. Detektor krawędzi Prewitt z metodą różnicową (a)-(j).

5.2. Analiza otrzymanych wyników

W tej części skupimy się na analizie danych liczbowych, które można otrzymać na podstawie otrzymanych obrazów wynikowych. Poszczególne zestawy danych zostały przedstawione z wykorzystaniem tabel i wykresów.

5.2.1. Progowanie z algorytmem SURF

Progowanie samo w sobie jest zależne od pewnej ustalonej wartości, który spowoduje rozdzielanie pikseli obrazu do dwóch grup – względem wyznaczonego progu piksele zostaną pogrupowane na te o wyższym oraz na te o niższym poziomie jasności. Naszym celem jest odnalezienie wartości progu, który skupia wszystkie punkty wyznaczone przez algorytm SURF wyłącznie na obiekcie.

Aby mieć jakiś punkt zaczepienia, na samym początku wybieramy kilka wartości progu i badamy jego wpływ na procent punktów charakterystycznych skupionych na obiekcie. Zakres wartości progu należy do przedziału $[0; 255]$. My wybieramy na początku sam środek zakresu oraz kilka mniejszych i większych wartości z krokiem co 8 co wdaje krok w wysokości $1/32$ całego zakresu. Wyniki dla *lizaków* zawiera poniższa tabela:

	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)	(i)	(j)
112	100	100	10	10	10	0	0	0	0	0
120	100	100	100	100	20	10	0	0	0	0
128	100	100	100	100	100	80	10	0	0	0
136	100	100	100	100	100	100	100	50	70	0
144	100	100	100	100	100	100	100	100	100	80

Tab.5.1. Procentowe skupienie punktów wyznaczonych przez algorytm SURF na poruszającym się obiekcie (*lizaki*), cz. I.

Możemy zauważyć, że zmiana wartości progu ze 112 na 144 powoduje ponad 4-krotny skupienie istotnych punktów na obiekcie. Przy tak szybkim wzroście, równocześnie dla drugiej z tych wartości mamy niemal 100 % takich punktów zawierających się w obiekcie. Możemy zatem przypuszczać, że niewiele większe progi pozwolą nam już na 100-procentowe skupienie punktów charakterystycznych na poruszającym się przedmiocie:

	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)	(i)	(j)
145	100	100	100	100	100	100	100	100	100	100
146	100	100	100	100	100	100	100	100	100	100
147	100	100	100	100	100	100	100	100	100	100

Tab.5.2. Procentowe skupienie punktów wyznaczonych przez algorytm SURF na poruszającym się obiekcie (*lizaki*), cz. II.

Widzimy, że już dla progu 145 mamy już całkowite skupienie na nagrywanym przedmiocie. Kolejne wartości progu nie powodują znajdowania naszych minimów poza

obiektem. Powodem tego jest wyeliminowanie wszystkich nieistotnych i niepotrzebnych elementów, które znajdują się w tle. Jednak to nie koniec. Może się okazać, że pod wpływem ciągłego zwiększania wartości progu może zająć problem z wykryciem samego obiektu. Sprawdźmy dla kilku przykładowych wartości progów (co 1/8 całego zakresu uwzględniając maksymalną dopuszczalną wartość) czy nasze podejrzenia są słuszne.

	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)	(i)	(j)
159	100	100	100	100	100	100	100	100	100	100
191	100	100	100	100	100	100	100	100	100	100
223	60 (100)	20 (100)	100	100	100	100	100	100	100	50 (100)
255	10 (100)	0	50 (100)	0	60 (100)	30 (100)	20 (100)	50 (100)	50 (100)	0

Tab. 5.3. Procentowe skupienie punktów wyznaczonych przez algorytm SURF na poruszającym się obiekcie (lizaki), cz. III.

Dane z tabeli 5.3. potwierdzają nasze obawy – zakres wartości progu ze 100-procentowym wykryciem obiektu nie kończy się wraz z końcem zakresu. Widzimy również, że dla wartości progu 223 oraz 255 pojawił się problem z wyznaczeniem przez metodę SURF ustalonej liczby punktów (10) ale nadal w komplecie znajdują się w obrębie śledzonego obiektu – stąd wartość 100 [%] w nawiasie. Ułatwmy nieco dalsze: poszukiwania:

	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)	(i)	(j)
195	100	100	100	100	100	100	100	100	100	100
199	100	100	100	100	100	100	100	100	100	100
203	100	100	100	100	100	100	100	100	100	100
207	100	60 (100)	70 (100)	100	100	100	100	100	100	100
211	100	60 (100)	80 (100)	100	100	100	100	100	100	100
215	100	70 (100)	80 (100)	100	100	100	100	100	100	100
219	70 (100)	50 (100)	60 (100)	100	100	100	100	100	100	70 (100)

Tab. 5.4. Procentowe skupienie punktów wyznaczonych przez algorytm SURF na poruszającym się obiekcie (lizaki), cz. IV.

Na podstawie danych zebranych w tabeli 5.4. widzimy, że między wartościami progu wynoszącymi 203 i 207 znajduje się nasze szukane maksimum. Możemy już dokładnie wyznaczyć jego wartość:

	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)	(i)	(j)
204	100	100	100	100	100	100	100	100	100	100
205	100	100	100	100	100	100	100	100	100	100
206	100	100	90 (100)	100	100	100	100	100	100	100

Tab. 5.5. Procentowe skupienie punktów wyznaczonych przez algorytm SURF na poruszającym się obiekcie (lizaki), cz. V.

Dane z tabeli 5.5. pozwalają nam na wyznaczenie maksymalnego progu pozwalającego wykryć w każdej klatce sekwencji określoną ilość punktów (10). Wszystkie wyznaczone przez algorytm SURF znajdują się wewnątrz badanego obiektu. Wartość maksymalna wynosi 205.

W podobny sposób wyznaczymy odpowiednie progi dla sekwencji wideo z poruszającą się piłką. Początkowe wartości progu ustala się na identyczne jak w przypadku nagrywania ruchu sklejonych ze sobą lizaków. Wyniki przedstawia poniższa tabela:

	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)	(i)	(j)
112	0	0	10	10	10	10	10	10	10	10
120	0	0	10	10	10	10	10	0	10	10
128	0	0	0	10	10	10	0	0	0	10
136	0	0	0	10	0	0	0	0	0	10
144	0	0	0	0	0	0	0	0	0	0

Tab. 5.6. Procentowe skupienie punktów wyznaczonych przez algorytm SURF na poruszającym się obiekcie (piłka), cz. I.

Na podstawie zebranych danych widzimy, że średnia różnica między w kolejnych klatkach między progami o wartościach 112 i 144 nie przekracza 10 %. Ponadto przy progu 144 ani jeden punkt nie znajduje się w obrębie naszej nagrywanej piłki. Na podstawie zebranych rezultatów możemy przypuszczać, że poszukiwana wartość progu będzie znacznie poniżej środka zakresu. Sprawdzamy to dla wielkości progu od 48 do 96 z krokiem 16 (1/16 całego zakresu). Wyniki przedstawia kolejna tabela”

	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)	(i)	(j)
48	100	100	100	100	100	100	100	100	100	100
64	100	100	100	100	100	100	100	100	100	100
80	20	60	70	90	80	90	80	90	80	80
96	10	10	10	10	10	30	30	20	40	10

Tab. 5.7. Procentowe skupienie punktów wyznaczonych przez algorytm SURF na poruszającym się obiekcie (piłka), cz. II.

Możemy zaobserwować, że nasz próg będzie mniejszy niż 80 ale nie mniejszy niż 64. Zatem szukamy dalej pomiędzy tymi dwiema liczbami z krokiem co 4:

	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)	(i)	(j)
68	80	100	90	100	100	100	100	100	100	100
72	40	90	90	100	90	100	100	100	100	90
76	30	50	80	90	100	90	100	90	90	100

Tab. 5.8. Procentowe skupienie punktów wyznaczonych przez algorytm SURF na poruszającym się obiekcie (piłka), cz. III.

W ten sposób zawężaliśmy poszukiwania do liczby mniejszej niż 68 ale nadal nie mniejszej niż 64. Możemy już wyznaczyć dokładną wartość progu:

	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)	(i)	(j)
65	100	100	100	100	100	100	100	100	100	100
66	90	100	100	100	100	100	100	100	100	100
67	90	90	100	100	100	100	100	100	100	100

Tab. 5.9. Procentowe skupienie punktów wyznaczonych przez algorytm SURF na poruszającym się obiekcie (piłka), cz. IV.

W ten sposób wyznaczyliśmy wartość progu dla którego obserwujemy 100 % skupienia na obiekcie punktów wyznaczonych przez algorytm SURF. W tym przypadku wynosi ona 65.

Podobnie jak w przypadku *lizaków*, może się zdarzyć sytuacja w której wartość progu nie tylko pozwala pominąć elementy tła ale może również zaważyć na wykryciu samego obiektu. Sprawdźmy zatem dla kilka przykładowych wartości progu począwszy od 0 (początek zakresy) idąc co 1/16 całości przedziału (co 16) a skończywszy na wartości mniejszej niż 48. Próg 48 jest na tym etapie najmniejszym znanym progiem dającym 100-procentowe wyznaczenie punktów i zawieranie się ich wszystkich w granicach śledzonego obiektu. Wyniki przedstawia poniższa tabela:

	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)	(i)	(j)
0	0	0	0	0	0	0	0	0	0	0
16	10 (100)	20 (100)	0	0	0	0	0	0	20 (100)	60 (100)
32	100	100	100	100	100	100	90 (100)	100	100	100

Tab. 5.10. Procentowe skupienie punktów wyznaczonych przez algorytm SURF na poruszającym się obiekcie (piłka), cz. V.

Możemy zauważyć, że dla progu o wartości 32 mamy niemal 100 % wykrycia punktów więc sprawdzimy wartości nieco większe:

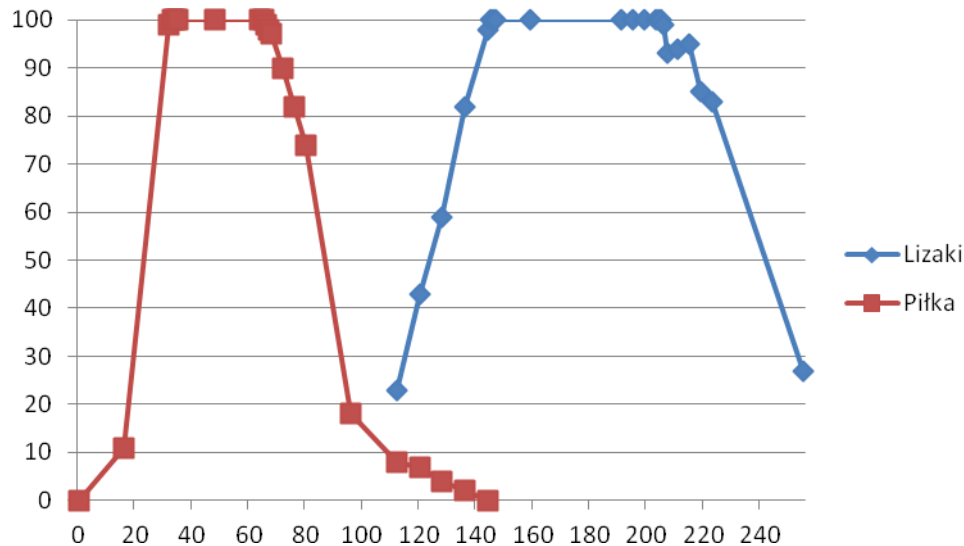
	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)	(i)	(j)
33	100	100	100	100	100	100	100	100	100	100
34	100	100	100	100	100	100	100	100	100	100
35	100	100	100	100	100	100	100	100	100	100

Tab. 5.11. Procentowe skupienie punktów wyznaczonych przez algorytm SURF na poruszającym się obiekcie (piłka), cz. VI.

Widzimy, że próg o wartości 33 jest minimalną wartością jaka pozwala nam na wyznaczenie wszystkich 10 punktów charakterystycznych w każdej i te punkty są skupione na obiekcie.

Na podstawie danych zebranych w tabelach od 5.1. do 5.11. możemy sporządzić wykres zależności procentu wykrycia punktów z algorytmu SURF od wartości dobranego progu. W tym celu dla każdego uwzględnionego w badaniach progu wyznaczono średni procent wykrycia ze wszystkich uwzględnianych klatek sekwencji wideo. Na jednym wykresie zaznaczymy wartości zebrane na podstawie nagrań obu przedmiotów.

Możemy zauważyć, że dla *lizek* minimalny próg dający pełne skoncentrowanie punktów charakterystycznych wynosi 145 a maksymalny 205. Dla *piłki* z kolei wartość progu spełniająca ten warunek waha się między 33 a 65. Biorąc pod uwagę, że mamy 256 możliwości, dla *lizek* optymalne jest 23,83 % zakresu (61 dopuszczalne wartości) a dla *piłki* 12,89 % zakresu (33 dopuszczalnych wartości).



Rys. 5.5. Wykres zależności procentu skoncentrowania punktów z algorytmu SURF na obiekcie od wartości progu.

Na podstawie powyższego wykresu możemy zauważyć, że zawartość nagrywanej sekwencji może mieć spory wpływ na znajdowanie optymalnej wartości progu. Dla jednego przykładu (*lizaki*) możemy zaobserwować większe prawdopodobieństwo wykrycia obiektu dla coraz wyższej wartości progu (do pewnego momentu) z kolei dla innego przykładu (*piłka*) może stać się dokładnie odwrotnie. Sporą różnicę między tymi przykładami uzmysławia nam fakt, że oba obiekty były nagrywane na tym samym tle. Jedynym utrudnieniem mogła być perspektywa ruszającej się kamery (co mogło też spowodować np. różne wstrząsy a także z innej perspektywy pokazać światłocienie na nagraniu) ale i ta była podobna przy nagrywaniu ruchu obu przedmiotów.

5.2.2. Detektor krawędzi Canny z algorytmem SURF

Detektor krawędzi Canny, podobnie jak progowanie, nie daje jednoznacznych rezultatów a wszystko przez parametry, które można zmieniać. Tutaj mamy do czynienia z dwoma wielkościami, które będziemy modyfikować. Pierwsza to wartość progowania (dotyczy przedłużania i łączenia krawędzi), która musi być liczbą z przedziału (0; 1). My do badań wybieramy liczby od 0,1 do 0,9 z krokiem co 0,1. Drugą z wielkości jest *sigma* (σ) czyli odchylenie standardowe funkcji Gaussa. Tu Ninie mamy takiego ograniczenia jak w przypadku pierwszego z parametrów. Wybieramy liczby naturalne z przedziału [1; 5].

W poniższych tabelach przedstawione są średnie procentowe koncentracji punktów wyznaczonych algorytmem SURF na śledzonych obiektach. W kolejnych wierszach mamy

wartości w zależności od wartości sigmy a w kolumnach są rozpisane kolejne wartości progowania.

	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9
1	58	93	100	100	100	100	100	100	100
2	60	94	99	99	100	100	100	100	100
3	66	93	98	100	100	100	100	100	100
4	70	95	98	100	100	100	100	100	100
5	69	93	97	100	100	100	100	100	100

Tab. 5.12. Średnie procentowe skupienie punktów wyznaczonych przez algorytm SURF na poruszającym się obiekcie w zależności od wartości σ i progowania (lizaki).

	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9
1	12	98	99	100	100	100	100	94 (100)	74 (100)
2	27	92	98	97	100	100	100	100	100
3	37	97	99	100	96	100	99	98	99
4	42	90	97	98	99	99	96	97	96
5	33	85	93	99	94	96	96	95	97

Tab. 5.13. Średnie procentowe skupienie punktów wyznaczonych przez algorytm SURF na poruszającym się obiekcie w zależności od wartości σ i progowania (piłka).

Podobnie jak w przypadku progowania, również i dla detektora krawędzi Canny jesteśmy w stanie tak dobrać parametry aby wszystkie punkty wyznaczone z algorytmu SURF były skupione na śledzonym obiekcie. Dla *lizaków* stało się tak w aż 30 przypadkach na 45 (66,67 %) a dla *piłki* w 11 na 45 przypadków (24,44 %). W tabeli dla *piłki* znalazły się 2 wyniki gdzie wartość 100 była w nawiasie. W tych przypadkach również wszystkie istotne punkty znajdowały się w obrębie śledzonych obiektów ale nie zawsze było ich 10. Wynikać to może z takiej kombinacji parametrów, która sama w sobie bardzo w ogóle ogranicza wykrywanie krawędzi.

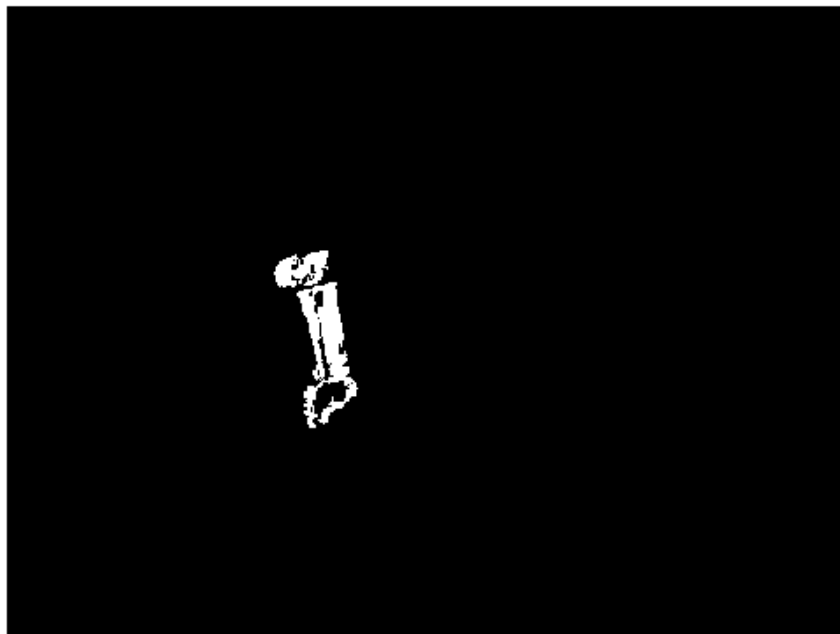
Na podstawie powyższych wyników możemy stwierdzić, że detektor krawędzi Canny bardzo dobrze radzi sobie z wykryciem krawędzi i jest to dobrą podstawą do dalszego działania w kwestii śledzenia ruchomych obiektów. Dla *lizaków* otrzymujemy średnią 95,16 % punktów z metody SURF w obrębie wykrytego i śledzonego przedmiotu oraz 89,29 % dla *piłki*. W tym drugim przypadku wartość wzrasta do 90 % po uwzględnieniu

jedynie tych wykrytych punktów. Jak już zostało wspomniane dla par parametrów (1; 0,8) i (1; 0,9) łącznie wykryto mniej punktów ale żaden z nich nie znajdował się poza obrębem nagrywanej piłki.

5.2.3. Progowanie z metodą różnicową

W tej części będziemy zajmować się znalezieniem optymalnego progu do wykrycia obiektu, a następnie wyznaczymy różnice kolejnych obrazów (klatek). Przy wyznaczaniu różnicy obrazu może wystąpić problem z dobraniem obiektywnego (liczbowego) kryterium, które pozwoli stwierdzić czy dana grupa pikseli wytypowana przez progowanie będzie należała wyłącznie do obiektu. Może to wynikać chociażby z niedokładności wykrycia samych obiektów. Możemy uzyskać to liczbę pikseli wyznaczonych do dwóch grup, zgodnie z ustalonym progiem ale w tym przypadku istotne jest ich położenie .

Przyjrzyjmy się dokładniej z czym przyjdzie nam się zmierzyć. Na rysunkach 5.6.-5.8. przedstawiono kolejne wybrane klatki z nagrywania *lizaków*. Są to kolejno różnice między klatkami: (b) i (c), (f) i (g) oraz (i) i (j). Próg jest równy połowie zakresu i wynosi 128. Na pierwszej z nich poza różnicą wynikającą z przesunięcia obiektu nie doszukamy się żadnych dodatkowych zmian. Problemy tego typu zaczynają się mniejszym (rys. 5.7.) lub większym (rys. 5.8.) stopniu uwidaczniać na kolejnych klatkach. Oprócz przesunięcia obiektu nasz obraz wynikowy zostaje „zaśmiecony” przez różnicę tła w kolejnych klatkach.



Rys. 5.6. Progowanie z metodą różnicową dla „lizaków” (klatki „b”, „c”).



Rys. 5.7. Progowanie z metodą różnicową dla „lizaków” (klatka „f”-„g”).



Rys. 5.8. Progowanie z metodą różnicową dla „lizaków” (klatka „i”-„j”).

Możemy jednak zauważyć pewną prawidłowość. Sięgnijmy do pewnych danych, które już wyznaczyliśmy, mianowicie chodzi o procentowe wykrycie punktów algorytmem SURF na obrazach poddanych progowaniu. Wyniki dla progu o wartości 128 przedstawia tabela 5.14.

	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)	(i)	(j)
128	100	100	100	100	100	80	10	0	0	0

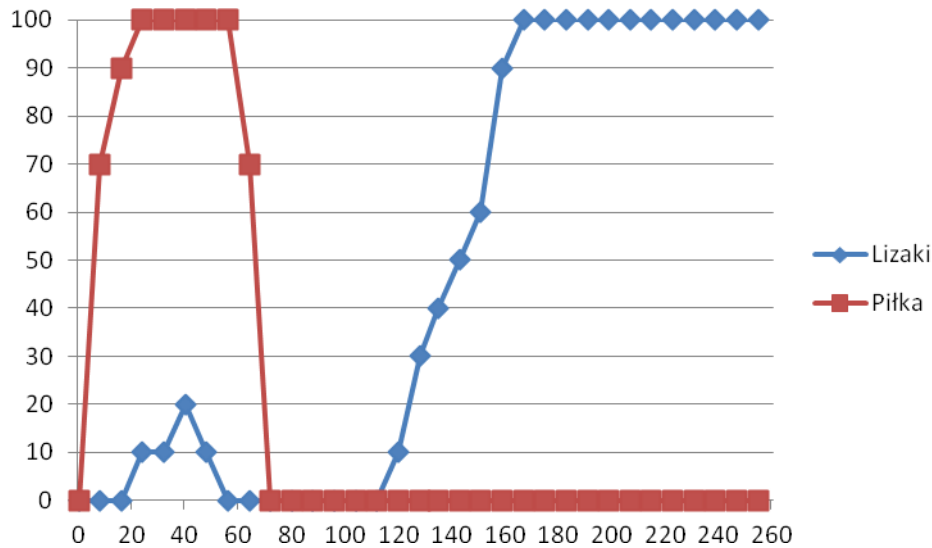
Tab.5.14. Procentowe skupienie punktów wyznaczonych przez algorytm SURF na poruszającym się obiekcie (lizaki) dla progu o wartości 128.

Możemy zauważyć pewną prawidłowość. Otóż w klatkach gdzie 100 % punktów charakterystycznych było skupionych na obiekcie różnica obrazów zawierała tylko piksele znajdujące się w obrębie obiektu (rys. 5.6.). Tam gdzie procent był mniejszy pojawiały się w różnicy piksele z poza śledzonego obiektu. Widzimy zatem każda najmniejsza odchyłka od 100 % dla danego progu wprowadzi nam niechciane piksele do naszej różnicy. Pomocne będzie tu wyznaczenie zakresów progów wyznaczonych przez współdziałania progowania i algorytmu SURF. Dla *lizaków* zakres progu wynosi od 145 do 205 a dla *piłki* od 33 do 65.

Aby potwierdzić nasze przypuszczenia sprawdźmy je wyznaczając procentowe skoncentrowanie pikseli na obiekcie na wyświetlanych różnicach. Jeśli wyznaczane piksele znajdowały się wyłącznie na obiekcie na wszystkich klatkach (czasem może być ich naprawdę mało) to dla takiego progu mamy 100 % skuteczności. Jeśli wszystkie piksele są jednego koloru (według progu wszystkie piksele przynależą do jednej grupy) lub daje się zauważyć piksele nie należące do obiektu na wszystkich klatkach to wówczas dla takiego progu mamy 0 % skuteczności. Jeśli na różnych klatkach mieliśmy pośrednią sytuację (np. w połowie sprawdzanych klatek zaczęły się pojawiać piksele z różnicy tła) to podano odpowiedni pośredni wynik równy procentowi klatek, na których był sam obiekt. Dane zostały zebrane w tabeli 5.15. oraz przedstawione na wykresie 5.9.

	0	8	16	24	32	40	48	56	64	72	80
Lizaki	0	0	0	10	20	10	10	0	0	0	0
Piłka	0	70	90	100	100	100	100	100	70	0	0
	88	96	104	112	120	128	135	143	151	159	167
Lizaki	0	0	0	0	10	30	40	50	60	90	100
Piłka	0	0	0	0	0	0	0	0	0	0	0
	175	183	191	199	207	215	223	231	239	247	255
Lizaki	100	100	100	100	100	100	100	100	100	100	100
Piłka	0	0	0	0	0	0	0	0	0	0	0

Tab.5.15. Procentowe skupienie pikseli na poruszającym się obiekcie w zależności od wskazanego progu (cz. I).



Rys. 5.9. Wykres zależności procentowego skupienia pikseli na poruszającym się obiekcie w zależności od wskazanego progu.

Na podstawie powyższych danych możemy określić zachowanie poszczególnych obiektów i tła na jakim były nagrywane. Dla *lizaków* początkowo nie mamy rozróżnienia obiektu od tła (progowanie łączy to w jednolity obraz), następnie dla wartości progu ok. zaczyna pojawiać się obiekt, choć początkowo jest on bardzo słabo widoczny. Później jednak zaczyna być wykrywana różnica w tle. Dopiero później ono zanika i między wartościami progów 159 i 167 jest całkowicie niewidoczne. Od tego momentu jedyną wykrywaną różnicą w obrazie jest przesunięcie się obiektu, jednak wraz ze wzrostem progu sam obiekt jest coraz słabiej widoczny. Z kolei dla *piłki* już od niemal samego początku widoczna jest różnica w przesunięcie obiektu (choć początkowo bardzo słabo.), tło „do głosu” zaczyna dochodzić między progami o wartościach 56 i 64. Później co prawda zanika ale znacznie wcześniej zanika nasz badany obiekt. Od pewnego momentu wyznaczenie różnicy traci sens, ponieważ nie zarówno obiekt jak i tło jest przez progowanie przyłączone do jednej grupy.

Wartości progów dobieraliśmy co 8. Zatem maksymalne procentowe skupienie pikseli dla *lizaków* rozpoczyna się między progiem o wartości 159 a 167 i kończy się wraz z końcem zakresu (255), a dla *piłki* rozpoczyna się między wartościami 16 a 24 i kończy między 56 a 64. Wyznamy dokładne graniczne wartości progu. Średnia procentowa skuteczność dla *lizaków* wyniosła 46,36 % a dla *piłki* 22,12 %.

	9	10	11	12	13	14	15
Piłka	70	80	80	80	80	90	90
	57	58	59	60	61	62	63
Piłka	100	100	100	100	80	80	80
	160	161	162	163	164	165	166
Lizaki	100	100	100	100	100	100	100

Tab.5.16. Procentowe skupienie pikseli na poruszającym się obiekcie w zależności od wskazanego progu (cz. II).

Na podstawie tabeli 5.16. możemy podać dokładne przedziały progów. Dla *lizaków* wynosi on od 159 do 255 a dla *piłki* od 16 do 60. Daje to 97 możliwości dla *lizaków* (38,04 %) i 45 możliwości dla *piłki* (17,65 %).

5.2.4. Detektor krawędzi Canny z metodą różnicową

Sprawdzenie działania progowania z metodą różnicową zostało oparte m. in. na wnioskach i wynikach otrzymanych na podstawie badań progowania z algorytmem SURF. Również i my odniesiemy się do tego co już udało nam się przeanalizować. Sięgnijemy do wyników jakie otrzymaliśmy wyznaczając optymalne parametry detektora krawędzi Canny do późniejszego wykorzystania w algorytmie SURF. Sprawdzimy rezultaty uzyskane dokładnie dla takich samych wartości σ (od 1 do 5 co 1) oraz progowania (od 0,1 do 0,9 co 0,1) w detektorze Canny

Dla poszczególnych wartości parametrów detektora Canny wyznaczymy procentową ilość klatek na których możemy wyróżnić grupę pikseli przynależącą jedynie do obiektu (bez tła).

	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9
1	0	50	80	100	100	100	100	100	100
2	0	10	80	80	100	100	100	100	100
3	0	0	80	100	100	100	100	100	100
4	0	0	60	100	100	100	100	100	100
5	0	0	40	100	100	100	100	100	100

Tab. 5.17. Średnie procentowe skupienie wyznaczonych pikseli na różnicach obrazów (klatek) w zależności od wartości σ i progowania (lizaki).

	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9
1	0	0	50	40	70	100	100	100	100
2	0	0	70	70	80	100	100	100	100
3	0	0	60	60	60	60	60	100	100
4	0	0	60	80	80	80	80	80	100
5	0	0	10	60	60	60	60	60	90

Tab. 5.18. Średnie procentowe skupienie wyznaczonych pikseli na różnicach obrazów (klatek) w zależności od wartości σ i progowania (piłka).

Na podstawie danych w tabelach 5.17. i 5.18. widzimy, że jest możliwe takie dobranie parametrów aby na różnicach kolejnych obrazów (klatek) były wykryte tylko krawędzie śledzonego obiektu. Dla *lizaków* mamy 29 z 45 możliwości (64,44 %) a dla *piłki* 11 z 45 możliwości (24,44 %). Dla wyników innych niż 100 % problemem było w mniejszym lub większym stopniu wykrycie krawędzi z tła. Dawały się one jednak częściowo lub całkowicie zniwelować wzrostem wartości progowania w detektorze Canny. Z kolei dla *piłki* większa szansa na wykrycie jedynie obiektu rośnie dla mającej wartości parametru σ . Dobrane parametry dały średnią skuteczność 75,11 % dla *lizaków* oraz 58,67 % dla *piłki*.

5.2.5. Nieparametryczne metody wykrywania krawędzi z algorytmem SURF

W tej części badań skupimy się na sprawdzeniu działania detektorów krawędzi Prewitt oraz Roberts wspólnie z metodą SURF. W przeciwieństwie do progowania czy detektora Canny metody te nie mają na żadnym etapie działania wielkości, którą można zmieniać. Wobec tego dają nam one jednoznaczne rozwiązanie. Przyjrzyjmy się jak prezentuje się procentowe skupienie punktów wyznaczonych przez algorytm SURF na śledzonym obiekcie. Badanie dotyczy ciągle tego samego fragmentu sekwencji wideo.

	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)	(i)	(j)
Prewitt	0	30	30	100	90	0	0	0	0	10
Roberts	100	100	90 (100)	100	100	100	100	90	100	70

Tab. 5.19. Procentowe skupienie punktów wyznaczonych przez algorytm SURF na poruszającym się obiekcie (*lizaki*).

	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)	(i)	(j)
Prewitt	20	100	90	90	70	80	60	50	90	100
Roberts	10	100	50	60	20	30	10	30	50	100
	(100)		(100)	(100)	(100)	(100)	(100)	(100)	(100)	

Tab. 5.20. Procentowe skupienie punktów wyznaczonych przez algorytm SURF na poruszającym się obiekcie (piłka).

Możemy zauważyć, że nie zawsze udaje się wyznaczyć ustaloną liczbę punktów co widać zwłaszcza przy metodzie Roberta. Pozytywem w tej sytuacji na pewno jest całkowite znajdowanie się takich punktów w obrębie obiektu (wartości 100 [%] w nawiasach). Otrzymane wyniki nie są zadowalające, zwłaszcza, że nie ma możliwości ingerencji w działanie metod Prewitt i Roberts (brak parametrów) co mogłoby wpłynąć na ich poprawę. Średnie wykrycie punktów przedstawia tabela 5.21. Widzimy, że dla dwóch różnych przykładów metody Prewitt i Roberts dały zupełnie odmienne rezultaty.

	Lizaki	Piłka
Prewitt	26	75
Roberts	95	46

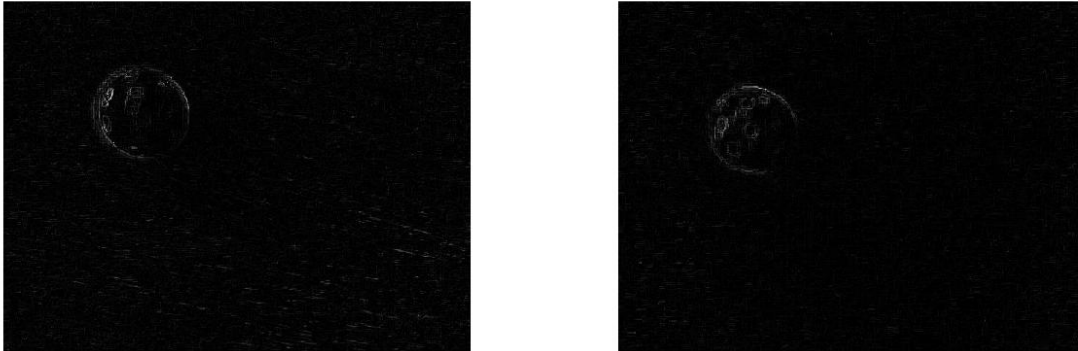
Tab. 5.21. Średnie procentowe skupienie punktów wyznaczonych przez SURF na badanych obiektach.

5.2.6. Nieparametryczne metody wykrywania krawędzi z metodą różnicową

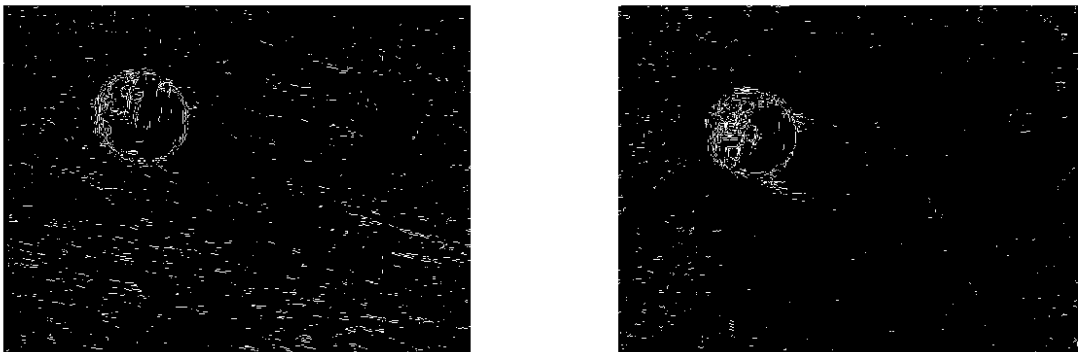
W tej części badań będziemy zajmować się działaniem detektorów krawędzi Prewitt i Roberts z metodą różnicową. Niestety tu mamy najmniejsze pole manewru spośród wszystkich możliwych rozwiązań. Nie mamy możliwości ingerencji w żadną z przytoczonych metod ze względu na brak jakichkolwiek zmiennych wielkości w ich działaniu. Możemy jedynie wzrokowo ocenić ich działanie. Przyjrzyjmy się zatem wybranym klatkom z nagrywania *piłki*. Zostały one zaprezentowane na rysunkach 5.10. oraz 5.11.

Brak liczbowego kryterium powoduje, że ciężko nam obiektywnie stwierdzić skuteczność na tle innych metod. Wzrokowo możemy jednak ocenić, że detektory krawędzi Prewitt i Roberts z metodą różnicową są raczej mało skuteczne. Problemem może być scena jaką nagrywamy. Może to wprowadzić wykrycie wielu niepotrzebnych nam krawędzi. Wzrokowo co prawda jesteśmy w stanie zauważyć obiekt ale to głównie dlatego, że wiemy w jakiej

części obrazów się ich spodziewać. Dla innego obrazu, z dużą ilością szczegółów mogłoby to już nie być takie oczywiste.



*Rys. 5.10. Detektor krawędzi Roberts z metodą różnicową dla „pilki”
(klatki „a”-„b” „i”-„j”).*



*Rys. 5.11. Detektor krawędzi Prewitts z metodą różnicową dla „pilki”
(klatki „a”-„b” „i”-„j”).*

Zatem ze względu na brak całkowitej ingerencji i brak możliwości znalezienia mierzalnego kryterium, pary tych metod (Prewitt + metoda różnicowa oraz Roberts + metoda różnicowa) nie zostaną uwzględnione w podsumowującym porównaniu działania wszystkich metod.

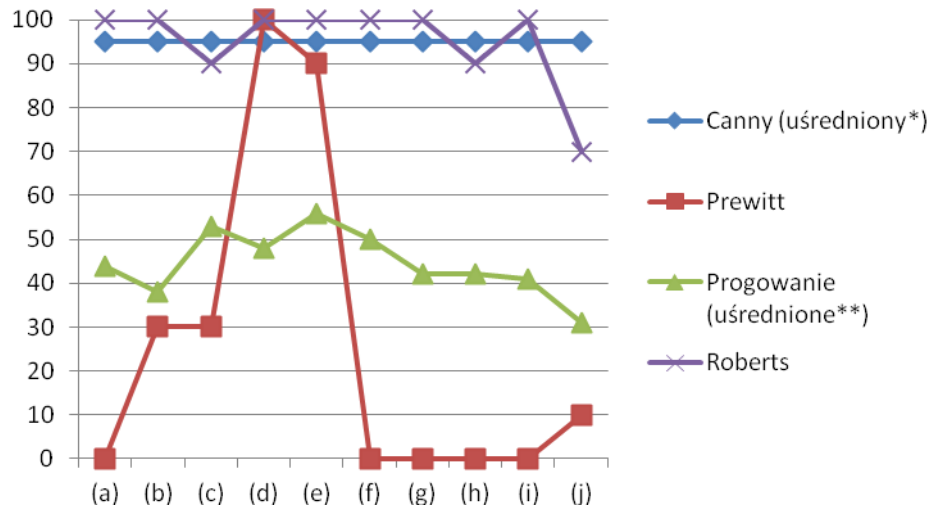
5.3. Porównanie działania algorytmów

Finalną częścią analizy wyselekcjonowanych algorytmów jest bezpośrednie porównania ich działania. Tabele 5.22 oraz 5.23 Przedstawiają procentowe skoncentrowanie punktów charakterystycznych wyznaczonych przez algorytm SURF dla obu nagrywanych sekwencji. O

ile dane wyznaczone przez detektory krawędzi Prewitt i Roberts są wyznaczone jednoznacznie (tabela 5.19.) o tyle progowanie i Canny dały nam już wiele wyników zależnych od parametrów tych metod. Zdecydowano się zastosować uśrednienie uzyskanych wcześniej wyników. Dla detektora krawędzi Canny jest to średnia jaką uzyskano na podstawie danych z tabel 5.12. (*lizaki*) i 5.13 (*piłka*). Ze względu na to, że w tych tabelach nie zdecydowano się jeszcze na rozbieżności i zaprezentowanie wyników na poszczególnych klatkach to średnią z całości przypisano do każdej klatki. Natomiast przy progowaniu wyznaczono już średnie z kilku wybranych progów (0, 32, 64, 96, 128, 159, 191, 223 i 255) z całego zakresu. Wyniki przedstawiono w tabelach i wykresach.

	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)	(i)	(j)	Śr.
Canny (uśredniony*)	95	95	95	95	95	95	95	95	95	95	95
Prewitt	0	30	30	100	90	0	0	0	0	10	26
Progowanie (uśrednione**)	44	38	53	48	56	50	42	42	41	31	45
Roberts	100	100	90	100	100	100	100	90	100	70	95

Tab. 5.22. Procentowe skupienie punktów wyznaczonych przez SURF (*lizaki*).



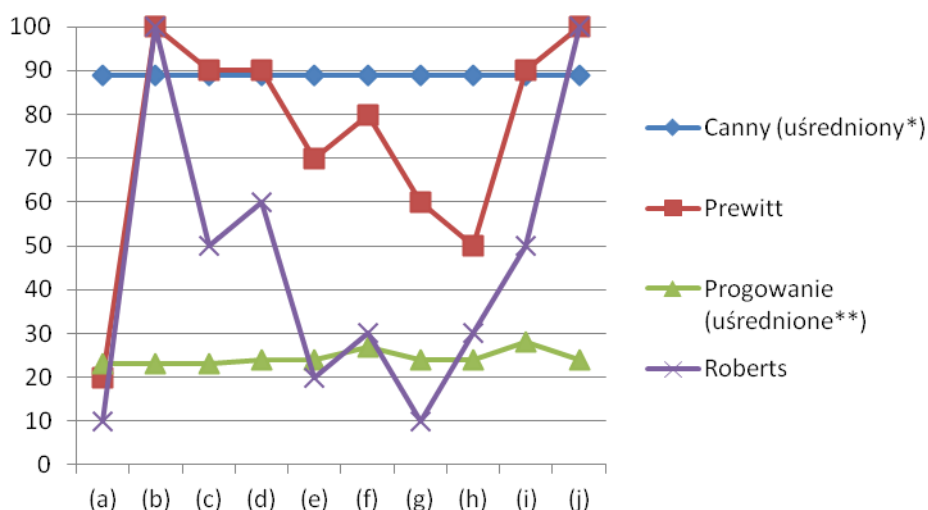
Rys. 5.12. Procentowe skupienie punktów wyznaczonych przez SURF (*lizaki*).

Na podstawie powyższych danych możemy stwierdzić, że dla *lizaków* najlepiej w algorytmem SURF współdziałały detektory krawędzi: Canny oraz Roberts. Dały one taki sam średni wynik z całej próby (95 % skuteczności) znacznie pozostawiając w tyle pozostałe metody. Detektor Canny daje nam duże możliwości odnośnie zmiany parametrów działania. Niestety nie mamy takiej możliwości przy metodzie Roberta ale mimo to wyniki otrzymano na tym samym, bardzo dobrym poziomie. W obu tych metodach było możliwe uzyskanie

100 % koncentracji punktów wyznaczonych przez SURF na obiekcie. Dla detektora krawędzi Prewitt możemy zauważyć bardzo dużą rozbieżność dla średnich z poszczególnych klatek. W zależności od klatki obiekt czasem nie był w ogóle zauważalny (0 %) a czasem cała uwaga była skupiona na nim (100 %). Duża podatność na filmowaną scenę oraz brak możliwości ingerencji w działanie metody poprzez parametry daje najmniejszą średnią skuteczność spośród badanych metod (zaledwie 26 %). Trochę lepiej prezentuje się progowanie – średnie z poszczególnych klatek generalnie nie odbiegają dużo od średniej z całości ale i ta nie jest imponująca i wynosi zaledwie 46 %. Na plus progowania jednak należy zaliczyć możliwość zmiany progu, co w części przypadków dawało nam całkowite wykluczenie wykrycia i śledzenia tła.

	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)	(i)	(j)	Śr
Canny (uśredniony*)	89	89	89	89	89	89	89	89	89	89	89
Prewitt	20	100	90	90	70	80	60	50	90	100	75
Progowanie (uśrednione**)	23	23	23	24	24	27	24	24	28	24	24
Roberts	10 (100)	100	50 (100)	60 (100)	20 (100)	30 (100)	10 (100)	30 (100)	50 (100)	100	46

Tab. 5.23. Procentowe skupienie punktów wyznaczonych przez SURF (piłka).



Rys. 5.13. Procentowe skupienie punktów wyznaczonych przez SURF (piłka).

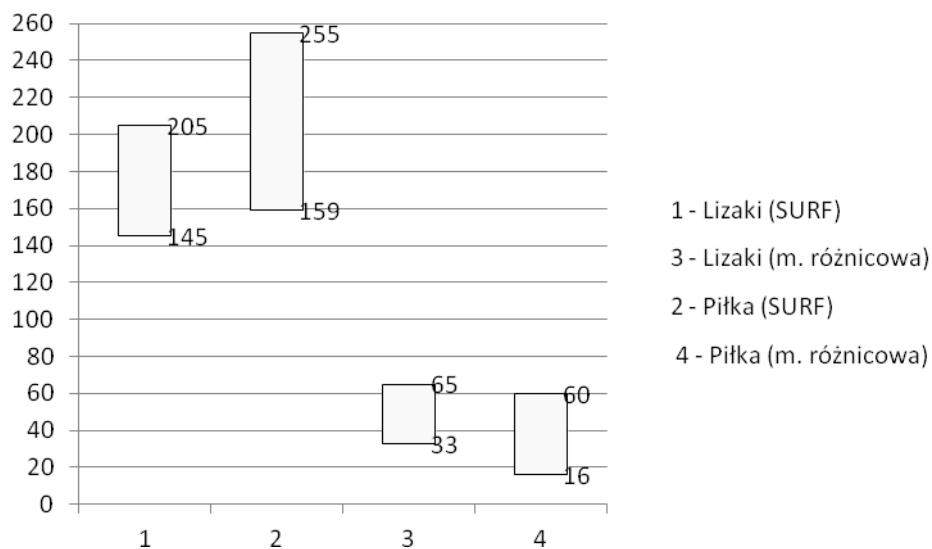
Nieco inaczej prezentują się wyniki analizy wykrycia i śledzenia dla *piłki*. Bez zmian pozostaje wysoka średnia skuteczność detektora Canny (89 %) oraz nienajlepszą średnią dla progowania (tym razem średnio zaledwie 24 %). W obu tych metodach, przy określonych

zestawach parametrów jesteśmy jednak w stanie uzyskać wyznaczenie 100 % punktów z algorytmu SURF w obrębie obiektu. Udało się też uzyskać po 100 % dla detektorów Prewitt i Roberts, jednak były to szczególne przypadki dla pojedynczych klatek. Tym razem metoda Prewitt była lepsza – średnia skuteczność wyniosła 75 %, podczas gdy Roberts może się pochwalić średnią jedynie 46 %. Możemy stwierdzić, że cały wpływ na zmianę wyników miała tylko zmiana śledzonego obiektu biorąc pod uwagę jej kształt i kolor. Utwierdza nas w tym fakt zapewnienia jednakowych warunków dla obu nagrań – takie samo tło, ta sama pora nagrywania (jasność, oświetlenie) oraz specyfika poruszania się kamery.

Kolejną rzecz jaką możemy przeanalizować to wpływ wartości progu w metodzie progowania na skuteczność algorytmu SURF oraz metody różnicowej. Wykorzystamy dane z podrozdziałów 5.2.1. oraz 5.2.3. Skupmy się jedynie na wartościach progów, które dają jednoznaczne wykrycie i śledzenie nagrywanych obiektów z całkowitym pominięciem tła. Zakresy tych progów możemy zobaczyć w tabeli 5.24 oraz wykresie 5.14.

	Metoda śl.	Min. próg	Maks. próg	Il. Progów	% zakresu
Lizaki	SURF	145	205	61	23,83
	M. różnicowa	159	255	97	37,89
Piłka	SURF	33	65	33	12,89
	M. różnicowa	16	60	45	17,58

Tab. 5.24. Progowanie w odniesieniu do algorytmu SURF i metody różnicowej.



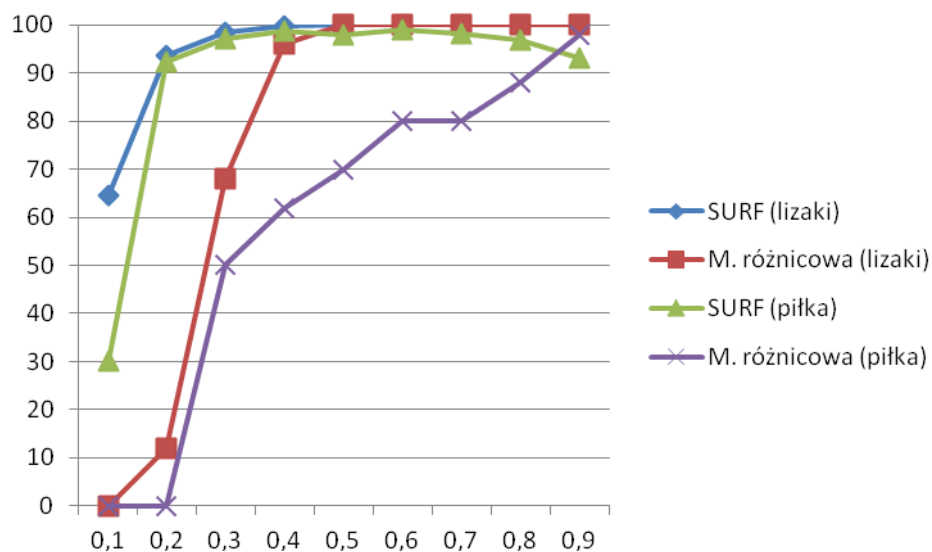
Rys. 5.14. Zakresy progów dających 100 % skuteczności wykrycia i śledzenia obiektów.

Widzimy, że zakresy progów dla obu tych metod nie są jednakowe. Dla metody różnicowej możemy zaobserwować więcej skutecznych możliwości niż dla algorytmu SURF. Kluczowe dla takiego obrotu sprawy jest odstępstwo obiektu śledzonego pod względem koloru i jasności od tła. Im dalej idziemy z wartością do początku lub końca zakresu tym łatwiejsze jest pominięcie tła. Zwłaszcza możemy to zaobserwować dla *piłki*. Dla *lizaków* obie metody działały dobrze dla wartości progów od 159 do 205 (27 możliwości co daje 18,36 %) a dla *piłki* od 33 do 60 (28 możliwości co daje 10,94 %).

Inny aspekt jaki możemy zbadać to sprawdzenie skuteczności algorytmu SURF oraz metody różnicowej wspólnie z detektorem krawędzi Canny. Będzie opierać się o wyniki badań z podrozdziałów 5.2.2. oraz 5.2.4. W tych podrozdziałach rozważaliśmy zmienność dwóch parametrów detektora Canny – σ oraz wartość progowania w tym detektorze. Z osobna dla każdej wartości poszczególnych parametrów wyciągnęliśmy średnią dla każdego z nagrywanych przedmiotów.

	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	Śr.
SURF (lizaki)	65	94	98	100	100	100	100	100	100	95
M. różnicowa (lizaki)	0	12	68	96	100	100	100	100	100	75
SURF (piłka)	30	92	97	99	98	99	98	97	93	89
M. różnicowa (piłka)	0	0	50	62	70	80	80	88	98	59

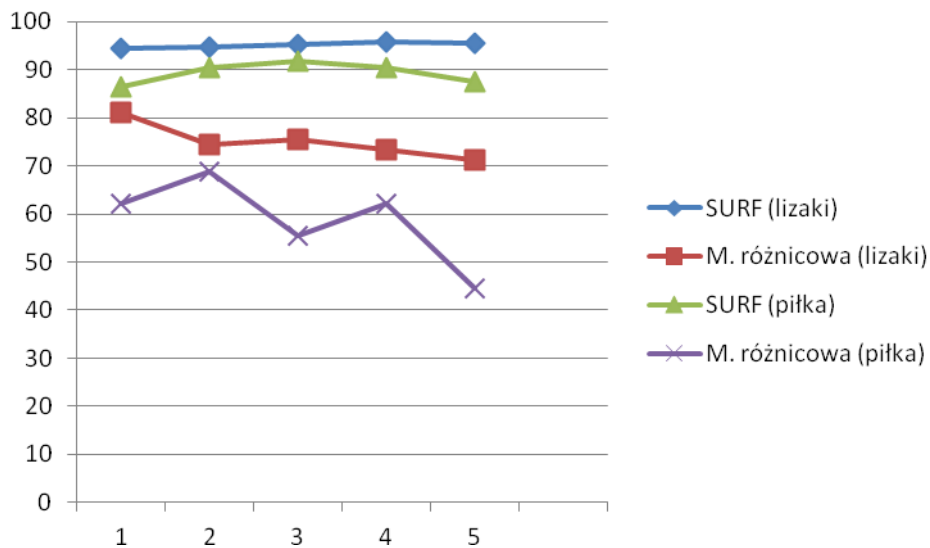
Tab. 5.25. Średnie procentowe skuteczności w zależności od wartości parametru progowania w detektorze krawędzi Canny.



Rys. 5.15. Średnie procentowe skuteczności w zależności od wartości parametru progowania w detektorze krawędzi Canny.

	1	2	3	4	5	Śr.
SURF (lizaki)	95	95	95	96	95	95
M. różnicowa (lizaki)	81	74	76	73	71	75
SURF (piłka)	86	90	92	90	88	89
M. różnicowa (piłka)	62	69	56	62	44	59

Tab. 5.26. Średnie procentowe skuteczności w zależności od wartości parametru σ w detektorze krawędzi Canny.



Rys. 5.16. Średnie procentowe skuteczności w zależności od wartości parametru σ w detektorze krawędzi Canny.

Na podstawie tabeli 5.25. oraz wykresu z rysunku 5.15. widzimy, że generalnie zwiększanie wartości parametru progowania w detektorze Canny powoduje zwiększenie skuteczności wykrycia i śledzenia obiektu. Możemy też zauważyć większą skuteczność algorytmu SURF w porównaniu do metody różnicowej, zwłaszcza dla małej wartości badanego parametru. Mimo to obie mogą osiągnąć średnią dla danej wartości parametru nawet 100 % ale jedynie dla *lizaków*. Porównywalnie radzi sobie algorytm SURF dla *piłki* (dla małych wartości progowania nawet lepiej niż metoda różnicowa dla *lizaków*) natomiast najsłabiej wypada metoda różnicowa dla *piłki*.

Znacznie mniejszą zmienność możemy zaobserwować zmieniając wartość parametru σ . Jedynie dla metody różnicowej dla *piłki* notuje się większe odchyłki od średniej. Widzimy, że algorytm SURF poradził sobie lepiej niż metoda różnicowa bez względu na nagrywany przedmiot.

Na sam koniec zbieramy do jednej tabeli średnie skuteczności wszystkich badanych algorytmów pod względem dobieranych kryteriów. Rezultaty przedstawia tabela 5.27.

		Canny	Prewitt	Progowanie	Roberts
<i>Lizaki</i>	SURF	95	26	45	95
	M. różnicowa	75	Nieklas.	46	Nieklas.
<i>Piłka</i>	SURF	89	75	24	46
	M. różnicowa	59	Nieklas.	22	Nieklas.
<i>Lizaki i piłka</i>	SURF	92	51	35	71
	M. różnicowa	67	Nieklas.	34	Nieklas.

Tab. 5.27. Podsumowanie procentowych skuteczności badanych algorytmów.

Tabela 5.27. daje nam pełny obraz średniej skuteczności metod. Dodatkowo możemy pokusić się o średnie poszczególnych metod biorąc pod uwagę współdziałanie z innymi metodami. Średnia dla algorytmu SURF wynosi 62 % podczas gry dla metody różnicowej 51 %. Dla detektorów krawędzi Prewitt i Roberts średnia wynika tylko na skutek działania z algorytmem SURF (odpowiednio 51 % i 71 %) a dla pozostałych metod wynosi: 80 % dla detektora krawędzi Canny oraz 35 % dla progowania.

6. Podsumowanie i wnioski

Niniejsza praca poruszyła ciekawe zagadnienie jakim jest badanie metod wykrycia i śledzenia obiektów. Oprócz oczywistego ruchu nagrywanych obiektów istotne było uwzględnienie również ruchu kamery. Celem było sprawdzenie w jakim stopniu jest to możliwe i w jak sobie z tym radzą różne algorytmy. Wybraliśmy kilka metod wykrywania obiektów, do których zaliczyliśmy detektory krawędzi (Canny, Prewitt, Roberts) oraz progowanie jako przedstawiciela metod segmentacji obiektów. Do śledzenia obiektów wybraliśmy algorytm SURF oraz metodę różnicową.

Materiałem badawczym były nagrane sekwencje wideo przedstawiające poruszanie się prostych i niedużych obiektów. Były to 2 sklejone ze sobą patyczkami lizaki w jaskrawych papierkach oraz niebieska piłeczka do unihokeja. Różnice w kształcie, kolorze i możliwym ruchu tych obiektów pomogły wyznaczyć jednakowe warunki nagrywania dla obu przedmiotów. Należały do nich: takie samo tło, naświetlenie oraz ta sama specyfika ruchu kamery.

Znajdowanie obiektywnych kryteriów porównania było kluczowe dla całego procesu badań skuteczności wybranych metod na analizowanym materiale wideo. Niestety nie dla każdej kombinacji metod udało się to uzyskać. Wobec takich przypadków zdecydowano się wstrzymać z oceną działania i nie klasyfikować względem pozostałych metod. Kilka spośród wyselekcjonowanych algorytmów miały możliwość zmiany swojego działania poprzez zmianę wartości wewnętrznego parametru lub parametrów. Dla takich przypadków zdecydowano się sprawdzić w jaki sposób miało to wpływ na końcowe rezultaty. Wszystkie algorytmy badane były osobno w poszczególnych parach (metoda wykrywania + metoda śledzenia) a na końcu otrzymane wyniki zostały skonfrontowane ze sobą. Dało to nam pełen obraz umożliwiający znalezienie najlepszej kombinacji metod spośród badanych.

Podczas badania kolejnych metod dało się zauważyć jak duży wpływ na wyniki może mieć badanie zachowania się różnych obiektów. Nie pomogło tu stanowiące zdecydowaną większość tło nagrywanej sekwencji ani inne aspekty mające na celu nadanie jednakowych warunków nagrywania dla obu obiektów. Rozbieżności te w kilku przypadkach można zniwelować zmianą parametrów, jednak nie każda wybrana do analizy metoda nam to umożliwiała.

Na podstawie danych i ich analizy zamieszczonych w rozdziale 5. możemy wytypować najlepszą kombinację metod a jest to detektor krawędzi Canny z algorytmem SURF. Głównym powodem osiągnięcia najwyższej oceny z analizowanych par algorytmów (92 %) jest to, że obie metody wykrywania i śledzenia obiektów działają bardzo dobrze na tym materiale wideo.

jest bardzo wysoka skuteczność wykrywania krawędzi obiektu dla różnych wartości parametrów detektora krawędzi Canny. W takich warunkach algorytm SURF wskazywał punkty charakterystyczne znacznie częściej w obrębie obiektu niż poza nim. Kolejne pary metod były już znacznie mniej skuteczne. Na podobnym poziomie średniej skuteczności znajdowały się: detektor krawędzi Roberts z algorytmem SURF (średnie 71 %) oraz detektor krawędzi Canny wraz z metodą różnicową. W pierwszym przypadku detektor Roberts okazał się gorszy niż Canny. W drugim metoda różnicowa spisała się gorzej niż SURF. Jeden wspólny czynnik z najlepszym zestawem algorytmów (Canny + SURF) pozwolił być na „podium” w analizowanych metodach ale wyniki mogą już nie być zadowalające.

Pozostałe pary metod były już znacznie mniej efektywne w porównaniu do trzech najlepszych. SURF z algorytmem Prewitt dały nam średnio 51 % skuteczności. Najslabiej pod tym względem wypadły pary metod z progowaniem. Progowanie z algorytmem SURF miało 35 % średniej skuteczności a ta sama metoda z metodą różnicową 34 %. Wpływ na tak słabe wyniki miały przede wszystkim nieduże zakresy dobrze wykrywającej obiekt wartości progu. Stąd pojawiało się wiele rozwiązań nie dających dobrych efektów co odbiło się na średnim wyniku. Pod względem procentowej średniej skuteczności nie brane były pod uwagę 2 pary: detektor krawędzi Prewitt z metodą różnicową oraz detektor krawędzi Roberts, również z metodą różnicową. Tutaj problemem było dobranie obiektywnego (liczbowego) kryterium, które pozwoliłoby nam zmierzyć jakość działania w wykryciu oraz śledzeniu obiektów.

Analizując wszystkie wybrane algorytmy na nagranych sekwencjach wideo możemy wysunąć wniosek, że metody, które mają możliwość zmiany wewnętrznych parametrów dadzą lepsze efekty niż te, w których nie mają takiej możliwości. Spośród metod wykrycia obiektów są takie dwie: detektor krawędzi Canny oraz progowanie. Chociaż średnie rezultaty mocno się od siebie różnią (Canny był średnio a progowanie najgorsze) to były to jedyne, które umożliwiały uzyskanie maksymalnego wyniku na całej długości trwania sekwencji wideo. Pozostałe dwie metody: Prewitt i Roberts nie dawały nam takiej możliwości. Powodowało to duży rozrzut skuteczności dla różnych filmowanych przedmiotów. Daje to mniejszą uniwersalność stosowania w różnych uwarunkowaniach. Ponadto wynik wykrywania automatycznie odbijał się na wyniku jego śledzenia. Im gorzej były one wykryte tym trudniej było wskazać jego ruch. Przeprowadzone doświadczenia pokazały, że przy dobrym wyodrębnieniu obiektu dobrze sprawdzi się nawet zwykła metoda różnicowa, które idealnie nadaje się do badania ruchu przy statycznej kamerze.

W ten sposób dokonaliśmy analizy metod, które w różnym stopniu sprawdziły się w wykrywaniu i śledzeniu obiektów. Była to jednak zaledwie część tego co można

przeanalizować. Powodem tego jest chociażby zbadanie działania zaledwie kilku metod, podczas gdy można wykorzystać jeszcze wiele innych. Można również zająć się analizą materiału wideo, który będzie przedstawiał znacznie więcej szczegółów co może znacznie utrudnić wykrycie obiektu. Innym aspektem jest także ruch kamery, która mogłaby się poruszać w zupełnie inny sposób niż przytoczony w tej pracy. Istotne mogą też być zakłócenia jakie mogą z tego wyniknąć (np. poprzez wstrząsy). Generalnie zakres problemu jaki się kryje pod pojęciem „wykrywania i śledzenia obiektu przy jednoczesnym ruchu kamery” na pewno się nie wyczerpał. Poszukiwanie innych rozwiązań niż te przedstawione w niniejszej pracy jest jak najbardziej wskazane, równocześnie mogą one sobie znacznie lepiej poradzić z postawionym problemem.

7. Spis rysunków

Rys. 2.1. Wykrywanie kolejnych obiektów na klatkach, (a)-(e), (f)-(j), (k)-(o) [102]	10
Rys. 2.2. Wykrycie jezdni i naniesienie jej krawędzi na obraz oryginalny: z perspektywy z lotu ptaka (a) i z perspektywy kamery zamontowanej na pojeździe (b) [6].....	11
Rys. 2.3. Wykrycie i wyodrębnienie z obrazu poruszających się samochodów [110]	11
Rys. 2.4. Różne metody segmentacji a) TSMAP, b) segmentacja poruszających się obiektów (Moving Objects), c) Mean-Shift, d) złożenie trzech powyższych metod (Mixture of Experts - MoE) [109].....	12
Rys. 2.5. Wykrycie ruchu jadącego autobusu przy pomocy detektora Kirsch wraz z LDP i LBP [104].....	13
Rys. 2.6. Efekty działania transformacji Hougha i filtru Ramer-Douglas-Peucker [105]	14
Rys. 3.1. Przykładowe maski dla kąta 180°: Prewitt (a), Sobel (b), Scharr (c), Kirsch (d)	16
Rys. 3.2. Krzyż Roberta (e) możemy przedstawić w nieco uproszczonej formie (f)	16
Rys. 3.3. Oryginalny obraz Lena.png [18]	18
Rys. 3.4. Obraz Lena.png po dekompozycji falką Haara [18].	19
Rys. 4.1. Klatka z pierwszego z użytych przykładów – „lizaki”	23
Rys. 4.2. Klatka z drugiego z użytych przykładów – „piłka”	23
Rys. 5.1. Progowanie z algorytmem SURF (a)-(j).	25
Rys. 5.2. Detektor krawędzi Roberts z algorytmem SURF (a)-(j).	27
Rys. 5.3. Detektor krawędzi Canny z metodą różnicową (a)-(j).	28
Rys. 5.4. Detektor krawędzi Prewitt z metodą różnicową (a)-(j).	29
Rys. 5.5. Wykres zależności procentu skoncentrowania punktów z algorytmu SURF na obiekcie od wartości progu.	35
Rys. 5.6. Progowanie z metodą różnicową dla „lizaków” (klatki „b”, „c”).....	37
Rys. 5.7. Progowanie z metodą różnicową dla „lizaków” (klatki „f”, „g”).....	38
Rys. 5.8. Progowanie z metodą różnicową dla „lizaków” (klatki „i”, „j”).....	38
Rys. 5.9. Wykres zależności procentowego skupienia pikseli na poruszającym się obiekcie w zależności od wskazanego progu.	40
Rys. 5.10. Detektor krawędzi Roberts z metodą różnicową dla „piłki” (klatki „a”, „b”, „i”, „j”).....	44
Rys. 5.11. Detektor krawędzi Prewitt z metodą różnicową dla „piłki” (klatki „a”, „b”, „i”, „j”).....	44
Rys. 5.12. Procentowe skupienie punktów wyznaczonych przez SURF (lizaki).	45
Rys. 5.13. Procentowe skupienie punktów wyznaczonych przez SURF (piłka).	46
Rys. 5.14. Zakresy progów dających 100 % skuteczności wykrycia i śledzenia obiektów.	47
Rys. 5.15. Średnie procentowe skuteczności w zależności od wartości parametru progowania w detektorze krawędzi Canny.	48
Rys. 5.16. Średnie procentowe skuteczności w zależności od wartości parametru σ w detektorze krawędzi Canny.	49

8. Spis tabel

Tab.5.1. Procentowe skupienie punktów wyznaczonych przez algorytm SURF na poruszającym się obiekcie (lizaki), cz. I	30
Tab.5.2. Procentowe skupienie punktów wyznaczonych przez algorytm SURF na poruszającym się obiekcie (lizaki), cz. II.	30
Tab.5.3. Procentowe skupienie punktów wyznaczonych przez algorytm SURF na poruszającym się obiekcie (lizaki), cz. III.	31
Tab.5.4. Procentowe skupienie punktów wyznaczonych przez algorytm SURF na poruszającym się obiekcie (lizaki), cz. IV.	31
Tab.5.5. Procentowe skupienie punktów wyznaczonych przez algorytm SURF na poruszającym się obiekcie (lizaki), cz. V.	32
Tab. 5.6. Procentowe skupienie punktów wyznaczonych przez algorytm SURF na poruszającym się obiekcie (piłka), cz. I	32
Tab. 5.7. Procentowe skupienie punktów wyznaczonych przez algorytm SURF na poruszającym się obiekcie (lizaki), cz. II	33
Tab. 5.8. Procentowe skupienie punktów wyznaczonych przez algorytm SURF na poruszającym się obiekcie (lizaki), cz. III.....	33
Tab. 5.9. Procentowe skupienie punktów wyznaczonych przez algorytm SURF na poruszającym się obiekcie (lizaki), cz. IV	33
Tab. 5.10. Procentowe skupienie punktów wyznaczonych przez algorytm SURF na poruszającym się obiekcie (lizaki), cz. V	34
Tab. 5.11. Procentowe skupienie punktów wyznaczonych przez algorytm SURF na poruszającym się obiekcie (lizaki), cz. VI	34
Tab. 5.12. Średnie procentowe skupienie punktów wyznaczonych przez algorytm SURF na poruszającym się obiekcie (lizaki).	36
Tab. 5.13. Średnie procentowe skupienie punktów wyznaczonych przez algorytm SURF na poruszającym się obiekcie (piłka).	36
Tab.5.14. Procentowe skupienie punktów wyznaczonych przez algorytm SURF na poruszającym się obiekcie (lizaki) dla progu o wartości 128.....	39
Tab.5.15. Procentowe skupienie pikseli na poruszającym się obiekcie w zależności od wskazanego progu (cz. I).	39
Tab.5.16. Procentowe skupienie pikseli na poruszającym się obiekcie w zależności od wskazanego progu (cz. II).	41
Tab. 5.17. Średnie procentowe skupienie wyznaczonych pikseli na różnicach obrazów (klatek) w zależności od wartości σ i progowania (lizaki).....	41
Tab. 5.18. Średnie procentowe skupienie wyznaczonych pikseli na różnicach obrazów (klatek) w zależności od wartości σ i progowania (piłka).....	42
Tab. 5.19. Procentowe skupienie punktów wyznaczonych przez algorytm SURF na poruszającym się obiekcie (lizaki).	42
Tab. 5.20. Procentowe skupienie punktów wyznaczonych przez algorytm SURF na poruszającym się obiekcie (piłka).	43
Tab. 5.21. Średnie procentowe skupienie punktów wyznaczonych przez SURF na badanych obiektach.	43
Tab. 5.22. Procentowe skupienie punktów wyznaczonych przez SURF (lizaki).	45

Tab. 5.23. Procentowe skupienie punktów wyznaczonych przez SURF (piłka).	46
Tab. 5.24. Progowanie w odniesieniu do algorytmu SURF i metody różnicowej.	47
Tab. 5.25. Średnie procentowe skuteczności w zależności od wartości parametru progowania w detektorze krawędzi Canny.	48
Tab. 5.26. Średnie procentowe skuteczności w zależności od wartości parametru σ w detektorze krawędzi Canny.	49
Tab. 5.27. Podsumowanie procentowych skuteczności badanych algorytmów.....	50

9. Bibliografia

- [1] Korohoda P., Tadeusiewicz R., *Komputerowa analiza i przetwarzanie obrazów*, Wydawnictwo Fundacji Postępu Telekomunikacji, Kraków 1997
- [2] Malina W., Smiatacz M., *Cyfrowe przetwarzanie obrazów*, Akademicka oficyna edukacyjna, Warszawa 2008
- [3] Mokrzycki W., *Wprowadzenie do przetwarzania informacji – II. Dyskretyzacja obrazu, operacje pikslowe, morfologiczne i przekształcenia obrazowe*, Akademicka Oficyna Wydawnicza Exit, Warszawa 2012
- [4] Oppenheim A.V., Schafer R. W., *Cyfrowe przetwarzanie sygnałów*, Wydawnictwo Komunikacji i Łączności, Warszawa 1979
- [5] Zieliński T. P., *Cyfrowe przetwarzanie sygnałów – od teorii do zastosowań*, Wydawnictwa Komunikacji i Łączności, Warszawa 2005
- [6] Creighton D., Kong H., Nahavandi S., Wei L., Zhou H., *Efficient Road Detection and Tracking for Unmanned Aerial Vehicle*, IEEE Transactions on intelligent transportation systems, Vol. 16 No 1, February 2015
- [7] Goulette F., Poręba M., *Automatyczna detekcja płaszczyzn w chmurze punktów w oparciu o algorytm RANSAC i elementy teorii grafów*, Archiwum Fotogrametrii, Kartografii i Teledetekcji Vol. 24, 2012, s. 301-310
- [8] Szymczyk M., Szymczyk P., *Analiza sceny przy użyciu deskryptorów punktów charakterystycznych*, Automatyka 2015, Tom 15, Zeszyt 3
- [9] Szymczyk M., Szymczyk P., *Aspekty czasowe algorytmu SURF w wersji sekwencyjnej i równoległej zaimplementowanej w technologii CUDA*, Automatyka i Robotyka 12/2011, s. 241-243
- [10] Abdullah-Al-Wahid M., Ahsan T., Chae O., Jabid T., Mohammad T., *An Edge-Texture Based Moving Object Detection for Video Content Based Application*, Proceedings of 14th International Conference on Computer and Information Technology (ICCIT 2011) 22-24 December, 2011, Dhaka, Bangladesh
- [11] Castelli T., Dinet E., Konik H., Tremeu A., *Moving Object Detection for Unconstrained Low-Altitude Aerial Videos, a Pose-Independent Detector Based on Artificial Flow*, 9th International Symposium on Image and Signal Processing and Analysis (ISPA 2015)
- [12] DeGol J., Nam M., *A Clustering Approach for Detecting Moving Objects Captured by a Moving Aerial Camera*, IEEE 2014 International Conference on Acoustic, Speech and Signal Processing (ICASSP)
- [13] Jachalsky J., Meuel H., Ostermann J., Reso M., *Superpixel-based Segmentation of Moving Objects for Low Bitrate ROI Coding Systems*, 2013 10th IEEE International Conference on Advanced Video and Signal Based Surveillance
- [14] Butler D., Cheng H., *Segmentation of Aerial Surveillance Video Using a Mixture of Experts*, Processing of the Digital Computing Techniques and Applications (DICTA 2015)
- [15] Chiheb R., Oulah Haj Thami R., Romadi M., Romadi R., *Detection and Recognition of Road Signs in Video Stream Based on the Shape of The Panels*, IEEE 2014
- [16] Hongguang L., Wenrui D., Zhe J., *Aerial Video Image Object Detection and Tracing Based on Motion Vector Compensation and Statistic Analysis*, IEEE 2009

- [17] Ankutowicz K., *Zastosowanie transformat Falkowych w zadaniach skalowania seam carving*, Projekt inżynierski, Politechnika Wrocławska, Wrocław 2015
- [18] <http://homepage.math.uiowa.edu/~jorgen/Haar.html>, Jorgensen P., Prof., Department of Mathematics, The University of Iowa, dostęp: 22.07.2016, godz. 12:15