

Konrad Jędrzejewski
Wydział Elektroniki i Informatyki
Politechnika Koszalińska
Śniadeckich 2
75-453 Koszalin

Porównanie metod filtracji obrazów z kamer monitorujących

Słowa kluczowe: monitoring, filtracja, pikselizacja, filtr medianowy statystyczny, filtr otwierający.

Keywords: monitoring, filtering, pixelization, statistical median filter, the filter opening.

1. Wstęp

Wraz z wprowadzeniem kamer monitorujących do systemów bezpieczeństwa w latach 60. XX wieku, w systemach tego typu nastąpił przełom. Użycie kamer wideo w latach 60. i 70. szybko wzrastało z powodu zwiększenia ich niezawodności, obniżenia kosztów i technologicznych ulepszeń w kamerach opartych na lampach analizujących. Najbardziej znaczącym postępowaniem w technologii wideo w latach 80. było wynalezienie i wprowadzenie kamer półprzewodnikowych. Lata 90. ujrzały integrację technologii komputerowej z technologią kamer bezpieczeństwa [1]. W latach 80. usprawniono system bezpieczeństwa poprzez dodanie wykrywacza ruchu (VMD), który dokonywał analizy obrazu z kamery i informował o ruchu na obszarze monitorowania.

Do końca lat 80. wykrywacz ten był urządzeniem analogowym (AVMD). Analogowy wykrywacz ruchu nadawał się jedynie do stosowania w pomieszczeniach, gdzie poziom światła mógł być kontrolowany. W przypadku wykorzystania AVMD na zewnątrz, często występowały fałszywe alarmy czyniące system bezpieczeństwa bezużytecznym. W latach 90. wprowadzono cyfrowy wykrywacz ruchu (DVMD). Urządzenie to okazało się skuteczne do stosowania w pomieszczeniach i na zewnątrz. Dziś ciężko wyobrazić sobie skuteczny system bezpieczeństwa bez użycia kamery.

Systemy bezpieczeństwa służą do monitorowania fabryk, przedsiębiorstw, sklepów. Są szeroko stosowane nie tylko do wykrywania ruchu, ale także do ostrzegania o podejrzanych zachowaniach, torbach pozostawionych na dworcach, ruchu w nieodpowiednim kierunku (na przykład na autostradach). Potrafią nie tylko wykryć zagrożenie, ale także poddać je analizie i w razie potrzeby odpowiednio zareagować. Współczesne kamery przemysłowe zaopatrzone są najczęściej w wykrywacz ruchu

umieszczony w obudowie razem z kamerą. W zależności od specyfiki miejsca, w którym mają zostać wykorzystane, stosuje się inne typy kamer. Przemysłowe kamery zaopatrzone są często w reflektor składający się z diod podczerwonych, aby mogły monitorować także w trybie nocnym. Kamery mogą działać w sieci, gdzie każda posiada swój adres IP. Możliwe jest wydzielenie obszaru, w którym system ma wykrywać ruch. Jeżeli obraz z kamery obejmuje na przykład kawałek ruchliwej ulicy, można wyłączyć ją z obszaru monitorowania (obraz nadal będzie rejestrowany w całości). Ważną cechą systemów bezpieczeństwa jest możliwość kalibracji, czyli określenia czułości systemu na ruch. Obraz w systemach monitoringu nagrywany jest często dopiero w momencie wykrycia ruchu. Stosowanie wykrywacza ruchu w każdej kamerze osobno pozwala uniknąć gromadzenia się ogromnej ilości danych, zwłaszcza jeśli kamer jest dużo.

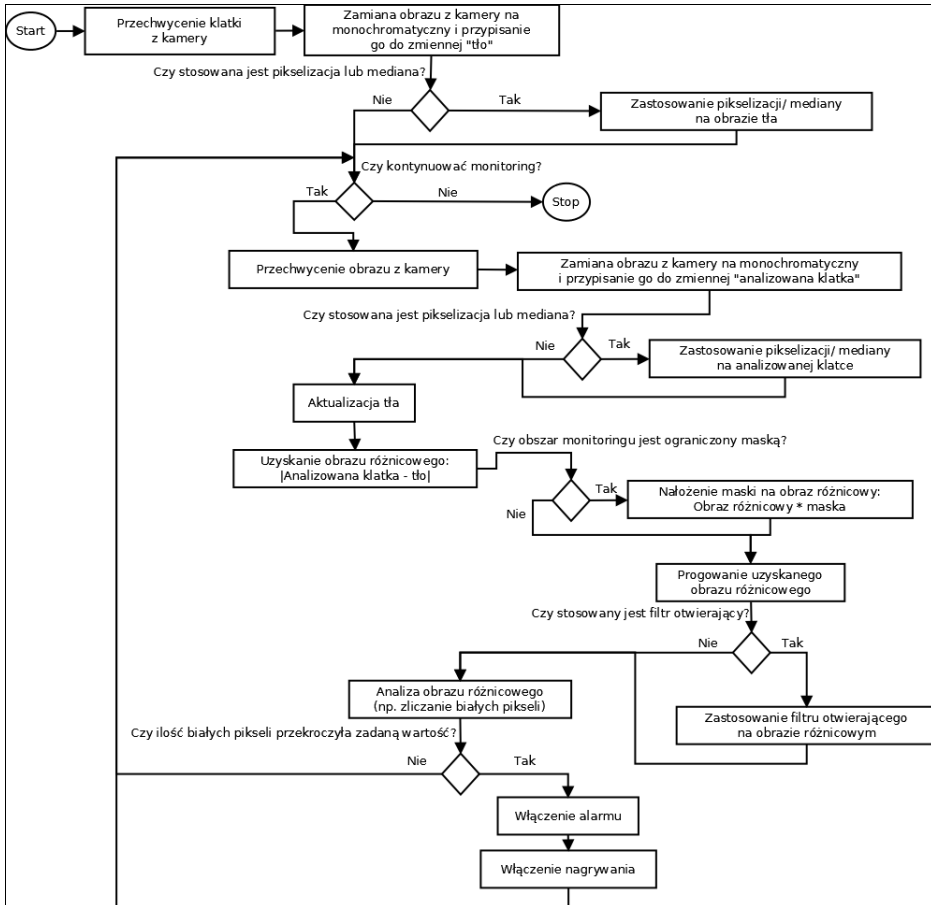
Obecnie ceny kamer przemysłowych zaczynają się od 200 złotych i mogą sięgać nawet kilku tysięcy złotych. Powstaje pytanie czy istnieje jakaś tańsza alternatywa. Okazuje się, że dla niektórych zastosowań można uzyskać zadowalający poziom bezpieczeństwa wykorzystując zwykłą kamerę internetową podłączoną do komputera, w którym zainstalowana jest aplikacja z zaimplementowanym algorytmem dokonującym analizy przechwytywanych klatek i alarmującym użytkownika o wykrytym ruchu.

2. Algorytm detekcji ruchu

Algorytm detekcji ruchu można zrealizować na wiele sposobów. Różnice mogą wynikać z wyboru klatek, które decydujemy się porównywać, wyboru metody usuwania zakłóceń, czy też metod związanych z analizą obrazu różnicowego. Algorytm operuje na czarno-białej kopii przechwyconego obrazu i w żaden sposób nie ingeruje w obraz oryginalny (oprócz ewentualnej ramki dodanej do obrazu, która może otaczać ruchomy obiekt w celu wyraźnego oznaczenia potencjalnego „intruza”). Jedną z najbardziej typowych metod jest porównanie bieżącej klatki z poprzednią. Jest to użyteczne w kompresji wideo, gdzie istnieje potrzeba przewidywania zmian w obrazie i zapisania tylko tych zmian, nie całej klatki [2]. Metoda ta, mimo że wydaje się najbardziej oczywista, nie sprawdza się najlepiej w systemach monitoringu. Różnice pomiędzy rejestrowanym klatkami mogą być bowiem tak znikome, że algorytm może w ogóle nie zadziałać.

Najbardziej efektywne algorytmy bazują na budowaniu tzw. tła sceny i porównywaniu bieżącej klatki z tłem [2]. Tło sceny uzyskuje się poprzez stałą aktualizację polegającą na nakładaniu na poprzednio uzyskane tło, bieżącej klatki. Można to zrobić za pomocą morphingu i podobnych algorytmów. Zabieg ten, przy doborze odpowiednich parametrów, dodatkowo pozwala pozbyć się problemu wykrywania regularnych ruchów (np. delikatnego ruchu liści na wietrze). Od bieżącej klatki odejmuje się tło, piksel po pikselu (co do wartości bezwzględnej) i uzyskuje obraz różnicowy. Obraz ten poddaje się następnie progowaniu, które binaryzuje obraz zamieniając piksele, których jasność przekracza zadaną wartość, na piksele białe a pozostałe na piksele czarne. Pozwala to pozbyć się już części zakłóceń. Na uzyskany obraz różnicowy nakłada się często maskę,

która ogranicza obszar monitoringu. W końcu tak przetworzony obraz może zostać poddany analizie. Najprostszą metodą jest zliczanie białych pikseli. Jeżeli ich liczba przekroczy zadaną wartość, system reaguje stanem alarmowym. Istnieją też bardziej wyrafinowane metody np. rozpoznawanie kształtów, określanie kierunku, w którym przemieszcza się obiekt, klasyfikacja ruchu.



Rys. 1. Przykładowy algorytm detekcji ruchu

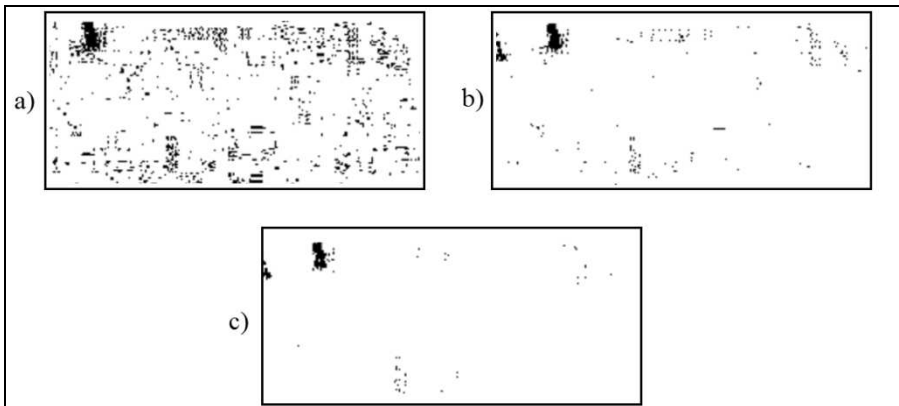
Na Rys. 1 zaznaczono miejsca, w których wykorzystuje się metody usuwania zakłóceń. Łatwo zauważyć, że pikselizacja i filtr medianowy operują na obrazach tła i bieżącej klatki, filtr otwierający wykorzystuje się zaś bezpośrednio na obrazie różnicowym po prognowaniu. Metody te, w połączeniu z odpowiednim stopniem prognowania, pozwalają się pozbyć zakłóceń, ale w szerszym kontekście- oprócz szumu wynikającego z niedoskonałości kamery, można pozbyć się innych niepotrzebnych informacji np.

ruchu bardzo małych obiektów. Można także sparametryzować np. ilość białych pikseli, bądź rozmiar obszarów białych pikseli wykrywanych na obrazie różnicowym, które mają powodować stan alarmu i połączyć wymienione powyżej parametry w stopień dokładności z jaką ma działać monitoring. Jak widać system zapewnia część funkcjonalności, którą można uzyskać dzięki kamerom przemysłowym i specjalizowanemu oprogramowaniu. W dalszej części artykułu chciałbym opisać wybrane przeze mnie algorytmy filtracji i dokonać ich porównania.

3. Porównanie sposobów działania wybranych metod filtracji

3.1. Progowanie

Progowanie jest kluczowym punktem algorytmu wykrycia ruchu. Operacji tej poddaje się obraz różnicowy. Progowaniu można poddać obrazy monochromatyczne. Dzięki temu otrzymujemy obraz zbinaryzowany (składający się tylko z białych i czarnych pikseli). Progowanie charakteryzuje się poziomem progowania. Jest to najczęściej wartość od 0 do 255 określająca stopień jasności piksela w przetwarzanym obrazie. Jeżeli wartość przekroczy zadany próg jasności, odpowiedniemu pikselowi przypisywany jest kolor biały, w innym wypadku piksel staje się czarny. W zależności od stosowanego algorytmu filtracji stosuje się różne poziomy progowania, im algorytm bardziej ingeruje w obraz różnicowy, tym mniejsze progowanie. Metoda ta służy jedynie do wstępnego przetwarzania obrazu, z racji że ciężko byłoby usunąć zakłócenia samym progowaniem bez utraty skuteczności wykrycia ruchu, dlatego stosuje się ją w parze z filtrami, które są względem niej komplementarne.



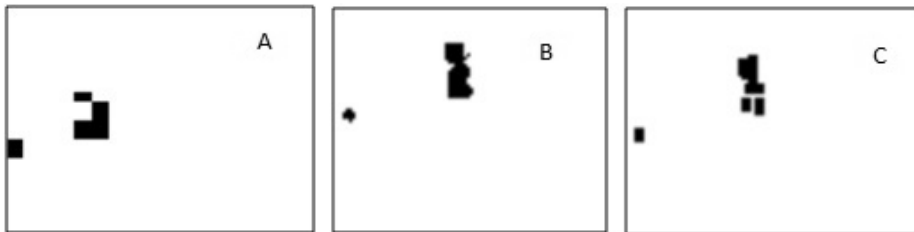
Rys. 2. Wpływ progowania na obraz różnicowy (na potrzeby wydruku odwrócono kolory):
a) próg = 1; b) próg = 5; c) próg = 8

3.2. Pikselizacja

Metoda polega na wstawieniu wartości średniej pikseli z zadanego obszaru (np. 8 x 8) w miejsce oryginalnych pikseli. Działa podobnie jak filtr dolnoprzepustowy uśredniający (gdzie w masce są tylko wartości 1), obejmuje jednak większe obszary pikseli. Jest to zabieg dość radykalny dla obrazu różnicowego, dlatego stosuje się go z niskim poziomem progowania. Pikselizacji używa się na obrazie tła i przechwyconej klatce. Przykładowy wynik pikselizacji ilustruje Rys. 3A.

3.3. Filtr statystyczny medianowy

Metoda polega na posortowaniu pikseli w masce (3 x 3) rosnąco, lub malejąco i przypisanie wartości środkowej do piksela w środku maski. Filtr medianowy jest bardzo dokładny i nie ingeruje mocno w obraz różnicowy, wymaga zatem wyższego progu niż pozostałe przeze mnie opisane. Metoda skutecznie pozwala się pozbyć z obrazu zakłóceń typu pieprz i sól zachowując niemal w pełni jakość przetwarzanego obrazu. Podobnie jak pikselizację filtr medianowy stosuje się na obrazie tła i przechwyconej klatce. Przykładowy rezultat obraz różnicowy powstały w wyniku pikselizacji przedstawiono na Rys. 3B.



Rys. 3. A - obraz różnicowy powstały w wyniku pikselizacji (odwrócone kolory); B - obraz różnicowy powstały w wyniku zastosowania filtru medianowego (odwrócone kolory); C - obraz różnicowy powstały w wyniku zastosowania filtru otwierającego (odwrócone kolory)

3.4. Filtr otwierający

Filtr ten jest w rzeczywistości złożeniem dwóch innych filtrów: erozji i dylatacji. W przeciwieństwie do pozostałych opisanych przeze mnie metod usuwania zakłóceń, metoda ta nie ingeruje w obraz tła i bieżącej klatki. Operuje bezpośrednio na obrazie różnicowym po progowaniu. Najpierw przeprowadzona jest erozja (filtr minimalny), która przypisuje pikselowi wartość minimalną sąsiadujących pikseli. Pozwala to na pozbycie się z obrazu różnicowego małych grup pikseli. Erozja narusza jednak także większe obiekty, dlatego obraz poddaje się następnie dylatacji (filtr maksymalny), która przypisuje pikselowi wartość maksymalną sąsiadujących pikseli. Pozwala to przywrócić rozmiar dużym obiektom. Małe obiekty, które zniknęły z obrazu różnicowego w wyniku

erozji, nie zostaną odtworzone. Filtr ten ingeruje w obraz różnicowy słabiej niż pikselizacja, ale mocniej niż filtr medianowy. Nie wymaga stosowania wysokiego progowania. Przykładowy obraz różnicowy powstały w wyniku zastosowania filtru otwierającego przedstawiono na Rys. 3B.

4. Porównanie efektywności wybranych metod filtracji

Badając skuteczność algorytmów starałem się stworzyć warunki sprzyjające i mniej sprzyjające dla metod filtracji. Za sprzyjające warunki przyjąłem dużą różnicę poziomu jasności poruszającego się obiektu w stosunku do jasności jego otoczenia. Dobrym rozwiązaniem okazało się zastosowanie wskaźnika laserowego na powierzchni białej ściany w ciemnym i doświetlonym pomieszczeniu. Badanie powtórzono dla 3 kamer różnej klasy w odległościach 2-10 metrów od ściany (wskaźnik umieszczony był zawsze w odległości 10 metrów - zmieniła się tylko odległość kamery). Wskaźnikiem świecono w losowy punkt na białej ścianie (w obrębie pola widzenia kamery). Badanie powtórzono po 10 razy w ciemnym i oświetlonym pomieszczeniu odnotowując czy nastąpiło wykrycie ruchu. Każda kamera została odpowiednio skalibrowana, a wartość progowania dobrana odpowiednio do algorytmu (była stopniowo zwiększana aż do momentu kiedy nie występowały żadne zakłócenia). Do testów użyto napisanej przeze mnie aplikacji testującej algorytmy utworzonej w języku C# z wykorzystaniem biblioteki AForge.NET i środowiska Visual Studio 2010. Aplikacja umożliwiła dynamiczny wybór metody filtracji, wybór kamery, ustalanie progu i stopnia aktualizacji tła. Na ekranie widoczne były: obraz różnicowy i obraz oryginalny, na którym, w przypadku wykrycia ruchu pojawiał się napis „ALARM”, a poruszający się obiekt był otaczany ramką. Wykorzystano sprzęt:

- kamera Logitech V-UAP41 (rozdzielczość 352 x 228 pikseli);
- kamera Creative VF0470 (rozdzielczość 640 x 480 pikseli);
- kamera Microsoft LifeCam HD-3000 (rozdzielczość 1280 x 720 pikseli);
- wskaźnik laserowy;
- notebook Toshiba Satellite L670-16M (procesor: Intel Core i3-330M 2,13 GHz,
- pamięć RAM: 4GB, karta graficzna: ATI Mobility Radeon HD 5650).

W wyniku przeprowadzonych badań uzyskano następujące rezultaty. Litera „O” oznacza oświetlone pomieszczenie a „C” ciemne pomieszczenie.

Tabela 1. Skuteczność wykrycia światła lasera dla kamery Logitech V-UAP41

Odległość [m]	Ilość wykryć (na 10 prób)					
	Pikselizacja		Filtr medianowy		Filtr otwierający	
	O	C	O	C	O	C
2	10	10	10	10	10	10
3	10	10	10	10	10	10
4	10	10	10	10	10	10
5	7	10	7	10	8	10
6	4	6	5	7	6	8
7	0	4	0	4	0	5
8	0	0	0	0	0	2
9	0	0	0	0	0	0
10	0	0	0	0	0	0

Tabela 2. Skuteczność wykrycia światła lasera dla kamery CreativeVF0470

Odległość [m]	Ilość wykryć (na 10 prób)					
	Pikselizacja		Filtr medianowy		Filtr otwierający	
	O	C	O	C	O	C
2	10	10	10	10	10	10
3	10	10	10	10	10	10
4	10	10	10	10	10	10
5	10	10	10	10	10	10
6	10	10	10	10	10	10
7	7	10	7	10	10	10
8	0	10	4	10	10	10
9	0	10	0	10	8	10
10	0	10	0	10	0	10

Tabela 3. Skuteczność wykrycia światła lasera dla kamery Microsoft LifeCam HD-3000

Odległość [m]	Ilość wykryć (na 10 prób)					
	Pikselizacja		Filtr medianowy		Filtr otwierający	
	O	C	O	C	O	C
2	10	10	10	10	10	10
3	10	10	10	10	10	10
4	10	10	10	10	10	10
5	10	10	10	10	10	10
6	10	10	10	10	10	10
7	8	10	10	10	10	10
8	2	10	6	10	10	10
9	0	10	2	10	9	10
10	0	10	0	10	6	10

Z powyższych wyników można wywnioskować, że pod względem skuteczności wykrycia ruchu przy użyciu badanych metod filtracji najlepszą metodą jest filtr otwierający, mniej skuteczna jest mediana, a najmniej skuteczna pikselizacja. Jednakże wyciągnięcie tylko takich wniosków mogłoby okazać się niezasadne, ponieważ należy także wziąć pod uwagę fakt, że w sprzyjających warunkach ilość wykryć w kamerach Creative VF0470 i Microsoft LifeCam HD-3000 wynosiła 10 na 10 dla każdej z badanych odległości. Filtr medianowy, pomimo dużej dokładności w przetwarzaniu, wymagał większego progowania niż filtr otwierający, co skutkowało spadkiem skuteczności. Duży wpływ na badanie miały możliwości kamery (szczególnie światłoczułość i rozdzielczość).

Metody filtracji przetestowano także pod kątem szybkości przetwarzania obrazów. Wykorzystano do tego aplikację testującą. Algorytmy sprawdzono dla różnych rozdzielczości obrazów wejściowych. Testowano całość algorytmu do wykrywania ruchu (nie tylko metody filtracji). Uzyskane czasy to wartości średnie czasów wykonania algorytmu z 1000 (dla rozdzielczości 160 x 120 pikseli i 320 x 240 pikseli) i 100 (dla pozostałych rozdzielczości) przebiegów.

Tabela 4. Średnie czasy wykonania przebiegu algorytmu w zależności od metody filtracji

Rozdzielczość [piksele]	Średni czas wykonania algorytmu [ms]		
	Pikselizacja	Filtr medianowy	Filtr otwierający
160 x 120	3	11	10
320 x 240	9	39	38
640 x 480	32	146	152
1280 x 960	125	559	603

Na podstawie uzyskanych wyników obliczono ile przebiegów na sekundę jest w stanie wykonać algorytm w zależności od metody filtracji. Jest to istotne, ponieważ kamera może przysyłać 30 klatek/s. Wartości w poniższej tabeli wskazują czy algorytm jest w stanie przetworzyć wszystkie przesyłane klatki (o danej rozdzielczości) w czasie rzeczywistym, czy też istnieje potrzeba ograniczenia się do pobierania klatek tylko co jakiś czas.

Tabela 5. Ilość klatek na sekundę, jaką jest w stanie przetworzyć algorytm w zależności od metody filtracji

Rozdzielczość [piksele]	Ilość klatek na sekundę, którą jest w stanie przetworzyć algorytm		
	Pikselizacja	Filtr medianowy	Filtr otwierający
160 x 120	333	90	100
320 x 240	111	25	26
640 x 480	31	6	6
1280 x 960	8	1	1

Z powyższych wartości wynika, że pikselizacja jest najszybszą z testowanych metod i na wykorzystanym sprzęcie jako jedyna była w stanie zapewnić przetwarzanie wszystkich klatek wysłanych przez kamerę dla rozdzielczości 640 x 480. Co ciekawe filtr medianowy i otwierający przetwarzają klatki w bardzo podobnym tempie (medianowy jest nieco szybszy). Z badań wynika też, że wraz ze zwiększeniem rozdzielczości analizowanych obrazów dwukrotnie, w pionie i poziomie, czas przetwarzania zwiększa się dla każdego algorytmu w przybliżeniu czterokrotnie, co jest dość oczywiste.

5. Wnioski

Wybór metody filtracji do usuwania zakłóceń z obrazu jest zależny od preferencji użytkownika. Jeżeli stosujemy kamerę o rozdzielczości co najmniej 640 x 480 pikseli, najlepszym rozwiązaniem wydaje się być pikselizacja, z racji stosunku skuteczności do

czasu przetwarzania klatki. Kamery internetowe są dosyć tanie i bardziej opłaca się zastosować lepszą kamerę i szybszy, nieco mniej dokładny algorytm, niż słabą kamerę i złożony algorytm. Jeżeli zależy nam na dużej czułości algorytmu, najlepszym rozwiązaniem jest stosowanie filtra otwierającego. Jeżeli z kolei chcemy operować na dokładnym obrazie różnicowym (np. do rozpoznawania kształtów, czy też innej wyspecjalizowanej analizy) najlepszy jest filtr statystyczny medianowy. W systemach monitoringu bywa tak, że „lepsze jest wrogiem dobrego”, ponieważ duże zwiększenie czułości systemu na ruch, może w pewnym momencie skutkować fałszywymi alarmami. Należy o tym pamiętać zwłaszcza stosując taki monitoring na zewnątrz. Testując napisaną przeze mnie aplikację, kierując kamerę na zewnątrz, doszedłem do wniosku, że do zwykłego wykrycia ruchu najlepszą metodą usuwania zakłóceń jest pikselizacja. Generowała mniejszą ilość fałszywych alarmów niż pozostałe i wystarczająco skutecznie pozwalała wykryć ruch obiektów. Z moich badań wynika również, że za pomocą zwykłej kamery internetowej można uzyskać satysfakcjonujący, dla niektórych zastosowań poziom bezpieczeństwa. Serdecznie dziękuję Pani inż. Lidii Dmytrzak, Panu dr Waleremu Susłowowi, oraz Panu prof. dr hab. inż. Zbigniewowi Suszyńskiemu za wszelką udzieloną mi pomoc w przeprowadzeniu badań i pisaniu niniejszego artykułu.

Literatura

1. Kruegle H.: CCTV Surveillance Second Edition: Analog and Digital Video Practices and Technology, Elsevier, 2007.
2. Kirillov A.: Motion Detection Algorithms,
www: <http://www.codeproject.com/Articles/10248/Motion-Detection-Algorithms>, 2007.

Streszczenie

Artykuł opisuje wybrane metody usuwania zakłóceń z obrazów z kamer monitorujących. Przedstawiono w nim przykładowy algorytm do wykrywania ruchu na podstawie klatek przechwyconych z kamery. Metody filtracji (pikselizacja, statystyczny filtr medianowy, filtr otwierający) zostały porównane pod względem czułości na ruch, oraz średniego czasu przetwarzania obrazów. Pojęcie „zakłócenia” w niniejszym artykule rozumiane jest jako informacja zbędna z punktu widzenia systemu monitoringu. Dotyczy więc nie tylko szumów wynikających z niedoskonałości kamery, ale również zjawisk takich jak ruchy liści, refleksy i tym podobne.

Abstract

The article describes selected methods for removing noise from images from surveillance cameras. It presents an example of an algorithm for motion detection based on frames captured from the camera. Filtration methods (pixelization, statistical median filter, the

filter opening) were compared in terms of sensitivity to movement, and the average time for processing images. The concept of "interference" in this paper shall be construed as superfluous information from the point of view of the monitoring system. This applies not only noise resulting from imperfections of the camera but also the movement phenomena, such as leaves, reflections, and the like.