

Wykrywanie i lokalizacja obiektów za pomocą skanowania przestrzeni niskorozdzielczą kamerą termowizyjną

Karol Erd, Krzysztof Sawicki

Instytut Optoelektroniki, Wojskowa Akademia Techniczna, ul. Gen. Sylwestra Kaliskiego 2, 01-476 Warszawa

Streszczenie: We współczesnych systemach monitoringu i detekcji często stosowane są technologie termowizyjne, które umożliwiają obserwację obiektów w warunkach słabego oświetlenia czy trudnych warunkach pogodowych. Szczególnym wyzwaniem pozostaje jednak skuteczne wykrywanie i lokalizacja obiektów przy użyciu niskorozdzielczych kamer termowizyjnych, które mimo ograniczeń sprzętowych oferują szereg zalet, takich jak niski koszt, niewielkie wymiary oraz energooszczędność. W artykule przedstawiono autorski algorytm przetwarzania obrazu termowizyjnego przeznaczony dla kamer o niskiej rozdzielczości. Przedstawiono także i przetestowano system skanujący, realizujący opracowany algorytm, oparty na kamerze termowizyjnej o rozdzielczości 32×24 piksele, zbudowany przy użyciu popularnych i łatwo dostępnych komponentów elektronicznych. Zwrócono uwagę na zagadnienia optymalizacji filtracji, binaryzacji oraz operacji morfologicznych, które pozwalają zredukować szумы, poprawić jakość detekcji oraz umożliwiają precyzyjne określenie położenia wykrytych obiektów. Wyniki przeprowadzonych eksperymentów pozwalają sądzić, że nawet przy zastosowaniu niskorozdzielczej matrycy możliwe jest uzyskanie użytecznych informacji o ruchu i lokalizacji obiektów, co otwiera nowe możliwości rozwoju kompaktowych i ekonomicznych systemów monitoringu w podczerwieni.

Słowa kluczowe: kamera termowizyjna, przetwarzanie obrazu, detekcja ruchu

1. Wprowadzenie

Współczesne systemy monitoringu środowiskowego, przemysłowego i bezpieczeństwa coraz częściej opierają się na technologiach pozwalających na autonomiczne funkcjonowanie, minimalizujące zaangażowanie człowieka oraz koszty wdrożenia [1, 2]. Jednym z kierunków rozwoju w tej dziedzinie jest integracja niskorozdzielczych kamer termowizyjnych z czujnikami typu „pozostaw i zapomnij” (ang. *deploy and forget*), które łączą prostotę implementacji, energooszczędność i zdolność do samoorganizacji w ramach sieci czujników [3], przy jednocześnie niskim koszcie jednostkowym. Kamery termowizyjne o rozdzielczości rzędu dziesiątek pikseli, mimo wielu ograniczeń sprzętowych, po odpowiednich przekształceniach obrazu oferują możliwość skutecznej detekcji i lokalizacji obiektów [4],

co może znaleźć zastosowanie w monitoringu ruchu, kontroli dostępu, czy wczesnym wykrywaniu zagrożeń.

Podstawową zaletą niskorozdzielczych kamer termowizyjnych jest ich niski koszt produkcji, kompaktowe rozmiary oraz niewielkie zapotrzebowanie na energię, co umożliwia integrację z przenośnymi urządzeniami IoT. Kluczowym wyzwaniem jest takie przetwarzanie obrazu termowizyjnego, które pozwala uniknąć zmniejszenia rozdzielczości przestrzennej. Dodatkowo umożliwia kompensację wysokich szumów, wynikających zarówno z niedoskonałości detektorów, jak i dynamicznych warunków otoczenia [4]. W tym artykule zaprojektowano algorytm przetwarzania obrazu termowizyjnego w kierunku detekcji i lokalizacji obiektów z użyciem kamer termowizyjnych o bardzo niskiej rozdzielczości. Zaprojektowano i wykonano także system integrujący komercyjnie dostępną, niskorozdzielczą kamerę z silnikiem krokowym oraz modułem przetwarzania i komunikacji sieciowej, w którym zaimplementowano omówiony algorytm. Cały system przetestowano w rzeczywistych warunkach pod względem skuteczności działania.

1.1. Omówienie problematyki

Stosowanie niskorozdzielczych kamer termowizyjnych, o rozdzielczości na przykład 32×24 piksele, wiąże się z fundamentalnymi ograniczeniami sprzętowymi, które znacząco wpływają na jakość danych wyjściowych. Główne wyzwania obejmują niską rozdzielczość przestrzenną, prowadzącą do utraty szczegółów strukturalnych obiektów, zwłaszcza przy większych

Autor korespondujący:

Krzysztof Sawicki, krzysztof.sawicki@wat.edu.pl

Artykuł recenzowany

nadesłany 12.05.2025 r., przyjęty do druku 16.01.2026 r.



Zezwala się na korzystanie z artykułu na warunkach licencji Creative Commons Uznanie autorstwa 4.0 Int.

odległościach od kamery [4]. Dodatkowo, kamery tego typu charakteryzują się wysokim poziomem szumów o charakterze zarówno termicznym (szum Johnsona-Nyquista), jak i systematycznym (niejednorodność matrycy, artefakty wynikające z niedoskonałości układu odczytu) [5, 6]. Ograniczona jest także dynamika sygnału, co utrudnia detekcję obiektów o niskim kontraście termicznym względem tła. W praktyce obserwuje się również obecność martwych pikseli, która w kamerach o niskiej rozdzielczości powoduje znaczne pogorszenie obrazu oraz niestabilność kalibracji w dynamicznych warunkach środowiskowych [7].

Wymienione czynniki sprawiają, że surowe dane z kamery termowizyjnej o niskiej rozdzielczości wymagają zaawansowanego przetwarzania, aby uzyskać użyteczne informacje o obecności i położeniu obiektów. Proponowany w dalszej części artykułu algorytm detekcji i lokalizacji został bezpośrednio zoptymalizowany pod kątem kompensacji powyższych ograniczeń.

1.2. Założenia systemu detekcji i lokalizacji ruchu

Projekt systemu detekcji i lokalizacji ruchu oparto na czterech fundamentalnych założeniach wynikających bezpośrednio z charakterystyki niskorozdzielczych kamer termowizyjnych. Po pierwsze, przyjęto strategię kompensacji ograniczeń sprzętowych przez wieloetapowe przetwarzanie, uwzględniające specyfikę szumów termicznych i systematycznych, charakterystycznych dla kamer termowizyjnych o niskiej rozdzielczości. Ponadto do projektowania toru przetwarzania obrazu przyjęto rozdzielczość kamery równą 32×24 px. Po drugie, przyjęto, że system zaprojektowano do pracy stacjonarnej w czasie quasi-rzeczywistym z niską częstotliwością próbkowania, rzędu pojedynczych Hz. Po trzecie, przyjęto minimalizację złożoności obliczeniowej przy zachowaniu kluczowych metryk skuteczności, rezygnując z zaawansowanych metod uczenia maszynowego na rzecz klasycznych technik przetwarzania obrazu. Po czwarte, założono, że system wyposażony jest w mechanizm skanowania przestrzennego umożliwiający obserwację otoczenia w szerokim kącie widzenia przez obrót kamery.

1.3. Algorytm przetwarzania obrazu

1.3.1. Korekcja niejednorodności matrycy

Detektory każdej matrycy nie dają jednakowej odpowiedzi na jednolite oświetlenie, ponieważ różnią się wartościami parametrów [5–8]. Zazwyczaj więc dokonuje się programowej kompensacji wprowadzanych przez to szumów [5]. Korekcja niejednorodności NUC (ang. *Non-uniformity correction*) eliminuje błędy pomiarowe spowodowane różnicami w wartościach parametrów detektorów [9]. Najczęściej stosuje się korekcje jedno- lub dwupunktową [6], z zastosowaniem jednego lub dwóch wzorców promieniowania. W pierwszym przypadku wystarczy jedynie jedno techniczne ciało czarne, powodujące jednolite oświetlenie matrycy. W drugim należy zastosować dwa takie wzorce o jednolitym rozkładzie promieniowania, odpowiadające różnym wartościom temperatury (lub jeden wzorec o przestrajalnej wartości temperatury). Korekcja dwupunktowa kompensuje wpływ niejednorodności detektorów w zakresie temperatury [6–8, 10].

Ze względu na typowy dla kamer o niskiej rozdzielczości silny dryf parametrów w funkcji temperatury obserwowanej sceny, a także na ich niski domyślny zakres dynamiczny, zdecydowano się na użycie w algorytmie korekcji dwupunktowej. Ponadto przyjęto, że korekcja niejednorodności zostanie przeprowadzona tylko raz, a uzyskane w jej wyniku macierze współczynników będą stosowane do wszystkich kolejnych rejestracji, ze względu na charakter pracy systemu. Korekcję dwupunktową należy przeprowadzić z użyciem technicznego ciała czarnego o wysokiej emisyjności. Kamery należy zbliżyć do powierzchni ciała czarnego tak, aby zajmowało cały widoczny obraz i zarejestrować po 10 ramek obrazu dla dwóch różnych wartości temperatury ciała:

przyjęto $20\text{ }^{\circ}\text{C}$ oraz $50\text{ }^{\circ}\text{C}$. Poszczególne serie uśredniono tak, aby otrzymać dwa wzorcowe obrazy dla każdej wartości temperatury. Następnie podstawiono je do wzorów (1) oraz (2) aby otrzymać macierz przesunięć $o(i, j)$ i wzmocnień $g(i, j)$:

$$g(i, j) = (V_H - V_L) / (V_H(i, j) - V_L(i, j)), \quad (1)$$

$$o(i, j) = V_H - g(i, j) \cdot V_H(i, j), \quad (2)$$

gdzie: $V(i, j)_H$, $V(i, j)_L$ – macierze sygnału z poszczególnych detektorów dla wyższej i niższej wartości referencyjnej temperatury, V_H , V_L – średni sygnał dla wyższej i niższej wartości referencyjnej temperatury.

Dzięki tak określonym macierzom korekcji można wyznaczyć macierz obrazu skorygowanego, korzystając ze wzoru (3):

$$V_{corr}(x, y) = g(x, y) \cdot V(x, y) + o(x, y), \quad (3)$$

gdzie: $V_{corr}(x, y)$ – macierz obrazu skorygowanego, $V(x, y)$ – macierz obrazu przed korekcją.

1.3.2. Wstępne przetwarzanie obrazu

Pomimo wykonania korekcji, obraz z kamery o niskiej rozdzielczości najczęściej nie nadaje się do analizy i należy go poddać wstępnemu przetwarzaniu [4]. Na początek należy poddać go cyfrowemu zwiększeniu rozdzielczości, w celu zapewnienia poprawnego działania przekształceń przestrzennych. Wśród najbardziej znanych metod zwiększenia rozdzielczości są: interpolacja najbliższego sąsiedztwa, interpolacja dwuliniowa, interpolacja sześcienna, resampling Lanczosa, czy neuronowe sieci głębokie.

Systemy detekcji oparte na kamerach o niskiej rozdzielczości mają najczęściej ograniczone zasoby obliczeniowe, co jest uwarunkowane użyciem komponentów o małych wymiarach i niskim poborze mocy. Dlatego też przyjęto, że najlepszą z metod będzie tu interpolacja sześcienna, ze względu na dobrą skuteczność przy akceptowalnym nakładzie obliczeniowym. Charakterystyczny dla interpolacji sześciennnej wygładzający wpływ, powoduje redukcję systematycznych szumów wynikających z niedoskonałości w układach odczytu.

Następnie obraz należy poddać normalizacji, wyrażonej wzorem (4) [11]:

$$I_{norm} = r \cdot \frac{I - \min(I)}{\max(I) - \min(I)}, \quad (4)$$

gdzie: I_{norm} – obraz po normalizacji, I – obraz wejściowy, r – zakres przyjmowanych wartości po normalizacji.

Zapewnia to stałość zakresu dynamiki przy późniejszym przetwarzaniu, a także zwiększa kontrast i ułatwia interpretację wizualną. Obraz po korekcji NUC został znormalizowany do zakresu 0–255 (reprezentacja 8-bitowa), co umożliwia standardowe przetwarzanie obrazu cyfrowego w dalszych krokach.

1.3.3. Filtracja obrazu

Podstawowa filtracja obrazów stanowi istotny etap przetwarzania sygnałów wizyjnych, mając na celu redukcję szumów, oraz przygotowanie obrazu do dalszej analizy [12]. Proces ten polega na modyfikacji wartości pikseli przez operacje matematyczne wykonywane w lokalnym otoczeniu każdego punktu obrazu. Głównym efektem filtracji jest tłumienie zakłóceń stochastycznych przy jednoczesnym zachowaniu istotnych cech strukturalnych, takich jak krawędzie [11–13]. Negatywnymi konsekwencjami mogą być jednak rozmycie szczegółów, zniekształcenia geometryczne obiektów oraz powstawanie artefaktów, szczególnie przy nieodpowiednim doborze parametrów

filtru. Głównym równaniem opisującym ogólną operację filtracji dla obrazu $I(i, j)$ jest:

$$I_{\text{filtrowany}}(i, j) = \sum_{m=-k}^k \sum_{n=-l}^l w(m, n) \cdot I(i + m, j + n), \quad (5)$$

gdzie $w(m, n)$ oznacza maskę filtru, a k, l określają wymiary otoczenia.

W kontekście obrazów termowizyjnych, gdzie szumy mają charakter zarówno addytywny (np. szum termiczny detektora), jak i multiplikatywny (nierównomierność odpowiedzi pikseli), szczególne znaczenie mają filtry nieliniowe [12]. Filtr medianowy realizuje operację statystyczną, zastępując wartość centralnego piksela medianą wartości z otoczenia określonego przez maskę filtracji. Dla obrazu $I(i, j)$ i kwadratowej maski o wymiarach $n \times n$, wynikową wartość każdego piksela oblicza się jako:

$$I_{\text{med}}(i, j) = \text{mediana}(I(x + i, y + j) \mid -n \leq i, j \leq n), \quad (6)$$

gdzie n definiuje rozmiar maski filtracji.

Mechanizm ten skutecznie eliminuje impulsowe zakłócenia (np. martwe piksele), zachowując przy tym ostre krawędzie obiektów. Wadą jest jednak ograniczona skuteczność wobec szumów o charakterze gaussowskim oraz wzrost zniekształceń przy dużych rozmiarach maski [11].

Filtr Gaussa jest kolejnym z podstawowych filtrów. Należy do liniowych metod filtracji przestrzennej, realizujących ważoną średnią ruchomą zgodnie z funkcją Gaussa. Wyraża się ją równaniem (7), implementowanym przez splot z maską o wartościach wyrażonych wzorem (8) [11].

$$G(i, j) = \frac{1}{2\pi\sigma^2} e^{-\frac{i^2 + j^2}{2\sigma^2}} \quad (7)$$

$$I_{\text{gauss}}(i, j) = \frac{1}{\sum_{m=-k}^k \sum_{n=-l}^l G(m, n)} \cdot I(i + m, j + n) \quad (8)$$

Wielkość σ określa stopień rozmycia filtru. W praktyce implementowany jest za pomocą dyskretnej aproksymacji jądra splotowego, gdzie wagi maleją wykładniczo wraz z odległością od centrum maski [11]. Filtr ten skutecznie redukuje szumy o rozkładzie normalnym, jednak powoduje rozmycie krawędzi proporcjonalnie do wartości σ . Zaleta w postaci separowalności maski znacząco redukuje złożoność obliczeniową w porównaniu z filtrami nieliniowymi.

W systemach termowizyjnych o niskiej rozdzielczości, gdzie występują szumy o zróżnicowanym charakterze, skutecznym podejściem jest zastosowanie kaskadowej filtracji obrazu. W opisywanym algorytmie połączono dwa komplementarne podejścia. W pierwszym kroku zastosowano filtr medianowy, który jako filtr nieliniowy, jest efektywny w eliminacji zakłóceń impulsowych (tzw. szumu typu „sól i pieprz”), wynikających z obecności wadliwych pikseli matrycy, przy jednoczesnym zachowaniu ostrości krawędzi obiektów [11]. Następnie obraz poddano filtracji z użyciem filtru Gaussa, który skutecznie kompensuje szumy o charakterze ciągłym, zbliżonym do rozkładu normalnego, takie jak szum termiczny detektorów [11, 12]. Optymalizacja parametrów – wymiarów maski n dla filtru medianowego oraz odchylenia standardowego σ dla filtru Gaussa – wymagała znalezienia kompromisu między stopniem redukcji szumu a zachowaniem kluczowych informacji o względnie małych obiektach, co ma znaczenie w aplikacjach monitoringu.

1.3.4. Binarizacja i przekształcenia morfologiczne

Do detekcji ruchu użyta została metoda różnic bezwzględnych. Dla kolejnych klatek I_{t-1} i I_t , różnicę tę wyznacza się według wzoru:

$$D(i, j) = |I_t(i, j) - I_{t-1}(i, j)|. \quad (9)$$

Pozwoliło to na wyodrębnienie obszarów, w których nastąpiła zmiana sygnału w czasie. Otrzymany obraz różnicy był następnie binaryzowany ze wzorem (10):

$$B(i, j) = \begin{cases} 255 & \text{gdy } D(i, j) \geq \mu_D + k\sigma_D \\ 0 & \text{w przeciwnym razie} \end{cases} \quad (10)$$

gdzie μ_D , σ_D oznaczają średnią i odchylenie standardowe różnicy w otoczeniu $\Omega(i, j)$, k oznacza czułość detekcji.

Binarizacja jest przekształceniem, którego działanie sprowadza się do zamiany obrazów wieloodcieniowych na obrazy binarne, które przyjmują tylko dwie wartości — zwykle 0 i 1 [11]. Zastosowanie binaryzacji jest bardzo szerokie i obejmuje przede wszystkim etapy wstępne w analizie obrazu, gdyż wiele kolejnych operacji (np. segmentacja, detekcja konturów, operacje morfologiczne) wymaga obrazów binarnych [11]. Binarizacja pozwala radykalnie zredukować ilość informacji zawartej w obrazie, co znacząco upraszcza dalsze przetwarzanie i umożliwia efektywne oddzielenie obiektów od tła.

Proces erozji i dyatacji stanowią podstawowe operacje morfologiczne używane w przetwarzaniu obrazów binarnych, gdzie pozwalają na modyfikację kształtu i struktury obiektów na obrazie [11]. Operacje te są realizowane przy użyciu elementu strukturalnego, który przesuwany jest po całym obrazie i determinuje sposób przekształcenia wartości pikseli.

Erozja jest operacją, która powoduje „kurczenie” obiektów w obrazie. W wyniku erozji piksel pozostaje aktywny (wartość 1 w obrazie binarnym) tylko wtedy, gdy wszystkie piksele w sąsiedztwie określonym przez element strukturalny są również aktywne [11]. Matematycznie erozję obrazu binarnego A przez element strukturalny B można zapisać jako:

$$A \ominus B = \{x \in E^N \mid B_x \subseteq A\} \quad (11)$$

gdzie x jest punktem, w którym znajduje się element B_x .

Dylatacja natomiast jest operacją odwrotną do erozji — powoduje „powiększanie” obiektów na obrazie. Piksel staje się aktywny, jeśli choć jeden piksel z sąsiedztwa określonego przez element strukturalny jest aktywny [11]. Definicja dyatacji obrazu binarnego A przez element strukturalny B wyraża się wzorem:

$$A \oplus B = \{x \in E^N \mid B_x \cap A \neq \emptyset\} \quad (12)$$

Obie operacje są często stosowane w parze, na przykład w procesie otwierania (najpierw erozja, potem dyatacja) lub zamykania (najpierw dyatacja, potem erozja), co pozwala na lepszą kontrolę nad kształtem i strukturą obiektów na obrazie [14]. W przypadku tego systemu zastosowano najpierw erozję, a później dyatację, co pozwoliło na wyodrębnienie prawidłowego kształtu wykrywanych obiektów od losowych przekroczeń progu na obrazie różnicowym. Parametry operacji morfologicznych takie jak rozmiar elementu strukturalnego oraz liczba iteracji, zostały dobrane eksperymentalnie pod kątem maksymalnej wykrywalności zlokalizowanego obiektu, przy braku wykryć fałszywie dodatnich.

Środek ciężkości (centroid) w przetwarzaniu obrazu to punkt geometryczny reprezentujący średnie położenie masy obiektu na obrazie [15]. W kontekście detekcji ruchu środek ciężkości służy do śledzenia ruchu obiektów przez określenie jego pozycji w kolejnych klatkach. Matematycznie środek ciężkości można zdefiniować jako:

$$c_x = \frac{\sum_x \sum_y x \cdot B(i, j)}{\sum_x \sum_y B(i, j)}, \quad c_y = \frac{\sum_x \sum_y y \cdot B(i, j)}{\sum_x \sum_y B(i, j)} \quad (13)$$

gdzie $B(i, j)$ jest binarnym obrazem, otrzymanym z obrazu różnic bezwzględnych. Suma wszystkich pikseli obrazu stanowi moment zerowego rzędu, opisujący środek ciężkości obiektu na obrazie.

Na binarnym obrazie różnicy wykrywane były kontury, spośród których wybrane były dwa największe (przyjęto, że odpowiadają one głównemu obiektowi w dwóch kolejnych pozycjach). Następnie obliczany był środek ciężkości obu konturów, które służyły jako punkt odniesienia do śledzenia ruchu obiektu w czasie.

1.3.5. Metoda oceny skuteczności algorytmu

Testowanie działania systemu odbywało się w następujący sposób. Urządzenie umieszczono w prostym korytarzu budynku. Wykonano cztery następujące po sobie skany o łącznym kącie obrotu urządzenia równym 180° tak, że kamera w położeniu początkowym była skierowana równolegle do korytarza.

W trakcie rejestracji, przed urządzeniem, wolnym tempem w odległości od 1 do 5 metrów przechodził człowiek tak, aby został zarejestrowany na przynajmniej kilku skanach.

Zarejestrowane obrazy poddano następnie omówionemu wcześniej przetwarzaniu, a rejestracje powtarzano iteracyjnie, w celu dobrania optymalnych parametrów wszystkich stosowanych przekształceń obrazu.

W celu ilościowego zbadania skuteczności detekcji obiektów wykonano 20 skanów podobnych do omówionych wyżej. Następnie dla każdej rejestracji dokonano oceny zgodnie ze standardową metodyką stosowaną w zadaniach detekcji obiektów. Wykorzystano do tego współczynnik IoU (ang. *intersection over union*, nazywany także indeksem Jaccarda) wyrażony wzorem (14):

$$\text{IoU} = \frac{|A \cap B|}{|A \cup B|}, \quad (14)$$

gdzie: A – referencyjny obszar obiektu, B – obszar obiektu wykrytego.

Prawdziwy obszar obiektu (ang. *ground truth*) porównano w ten sposób z obszarem wykrytym przez algorytm, stosując następujące kryteria:

- Wykrycie, dla którego współczynnik IoU był większy lub równy 0,5, klasyfikowano jako prawdziwie pozytywne (TP),
- Wykrycie, dla którego IoU było mniejsze od 0,5 lub które nie odpowiadało żadnemu obiektowi referencyjnemu, klasyfikowano jako fałszywie pozytywne (FP),
- Obiekt referencyjny, który nie został wykryty (żadne wykrycie nie osiągnęło progu $\text{IoU} = 0,5$), klasyfikowano jako fałszywie negatywny (FN).

Następnie obliczono parametr $F1$, będący średnią harmoniczną precyzji i czułości, który jest standardową metryką stosowaną do oceny wydajności w zadaniach detekcji i klasyfikacji [16]. Obliczono go zgodnie ze wzorem (15):

$$F1 = \frac{2 \cdot TP}{2 \cdot TP + FP + FN}. \quad (15)$$

Kolejnym etapem było określenie skuteczności lokalizacji obiektów przez system. W tym celu ponownie wykonano 20 rejestracji, przy czym tym razem w momencie każdego skanowania, człowiek znajdował się w jednej z ustalonych pozycji kątowych całkowitego pola widzenia systemu. Ustalono pozycje kątowe wynoszące $\{30^\circ, 60^\circ, 90^\circ, 120^\circ, 150^\circ\}$, jeżeli za 0° przyjmiemy środek pola widzenia kamery w pozycji początkowej. Dla każdej pozycji wykonano więc po cztery oddzielne rejestracje. Następnie wartość minimalną $\Delta\theta_{\min}$ oraz maksymalną $\Delta\theta_{\max}$ odchyłki wyznaczonej wartości kąta od wartości rzeczywistej, a także średni błąd bezwzględny kąta lokalizacji zgodnie ze wzorem (16) oraz błąd średniokwadratowy kąta lokalizacji zgodnie ze wzorem (17):

$$\text{MAAE} = \frac{\sum_{i=1}^n |\theta_{c_x, i} - \theta_{ob, i}|}{n}, \quad (16)$$

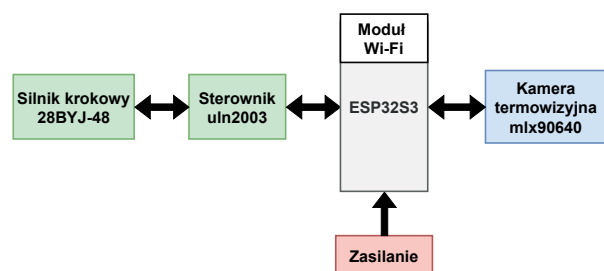
$$\text{RMSAE} = \sqrt{\frac{\sum_{i=1}^n (\theta_{c_x, i} - \theta_{ob, i})^2}{n}}, \quad (17)$$

gdzie: θ_{c_x} – położenie kątowe środka ciężkości wykrytego obiektu, θ_{ob} – położenie kątowe rzeczywistego obiektu, n – liczba rejestracji.

2. Projekt i realizacja systemu skanującego

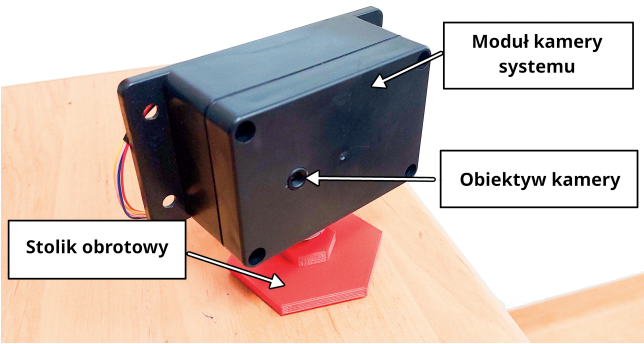
W celu określenia skuteczności opracowanej metody zaprojektowano i wykonano urządzenia spełniające opisane wcześniej założenia. System skanujący z niskorozdzielczą kamerą termowizyjną składał się z kamery mlx90640 o rozdzielczości 32×24 px i nominalnym polu widzenia $55^\circ \times 35^\circ$, płytki PCB z procesorem ESP32-S3 oraz silnika krokowego 28BYJ-48 wraz ze sterownikiem uln2003. Schemat blokowy systemu został przedstawiony na rysunku 1. Wszystkie elementy systemu zostały umieszczone w specjalnie wykonanej obudowie i przytwierdzone do silnika krokowego, którego oś została zamontowana na stoliku obrotowym. W ten sposób system mógł być w całości obracany o dowolnie duży kąt, a wbudowany w procesor moduł Wi-Fi zapewnił komunikację bezprzewodową, eliminując potrzebę użycia specjalistycznych złącz obrotowych. Wykonany model systemu został przedstawiony na rysunku 2.

Zasada działania systemu polegała na naprzemiennym obracaniu platformy i rejestracji obrazu z kamery termowizyjnej. Krok obrotu silnika został dobrany eksperymentalnie. Chociaż nominalne horyzontalne pole widzenia (FOV) kamery MLX90640, zgodnie z dokumentacją producenta, wynosi 55° [21], w praktyce jego użyteczna szerokość, pozwalająca na uzyskanie obrazów o dobrej jakości aż po krawędzie, okazała się bliższa 45° . W związku z tym, w celu zapewnienia ciągłości panoramy przy



Rys. 1. Schemat blokowy systemu skanującego z niskorozdzielczą kamerą termowizyjną mlx90640

Fig. 1. Block diagram of scanning system with low-resolution thermal camera mlx90640



Rys. 2. Realizacja praktyczna systemu skanującego zamontowanego na platformie obrotowej
Fig. 2. Practical implementation of scanning system mounted on rotating platform

jednoczesnej minimalizacji nakładu obliczeniowego, przyjęto kąt obrotu równy 45°. Taka strategia pozwoliła na sekwencyjne składanie kolejnych klatek w celu utworzenia panoramicznego obrazu termowizyjnego bez konieczności stosowania algorytmów korelacyjnych do obsługi obszarów wspólnych. Mimo że efektywne pole widzenia może zależeć od odległości do obserwowanego obiektu, dla przyjętego w eksperymencie zakresu od 1 do 5 metrów wartość 45° zapewniała poprawne wyniki.

3. Analiza zarejestrowanych obrazów

Przykładowy panoramiczny termogram przed korekcją niejednorodności przedstawiono na rysunku 3. Poziom szumu jest wysoki, ale sylwetka człowieka jest widoczna na jego tle. Szum wynikający z niejednorodności ma widoczny powtarzalny charakter, wynikający z metody składania obrazów.



Rys. 3. Przykładowy obraz termowizyjny przed korekcją niejednorodności (NUC) z widoczną sylwetką człowieka na tle wysokiego poziomu szumu
Fig. 3. Sample thermal image before non-uniformity correction (NUC) showing human silhouette against high noise background

Na rysunku 4 przedstawiono przykładowy obraz po korekcji niejednorodności. Widać na nim znaczną poprawę kontrastu, a sylwetka człowieka jest teraz znacznie bardziej widoczna. Obraz zawiera artefakty w kształcie szachownicy, co jest szczególnie widoczne w miejscu występowania sylwetki.



Rys. 4. Obraz termowizyjny po korekcji niejednorodności
Fig. 4. Thermal image after non-uniformity correction

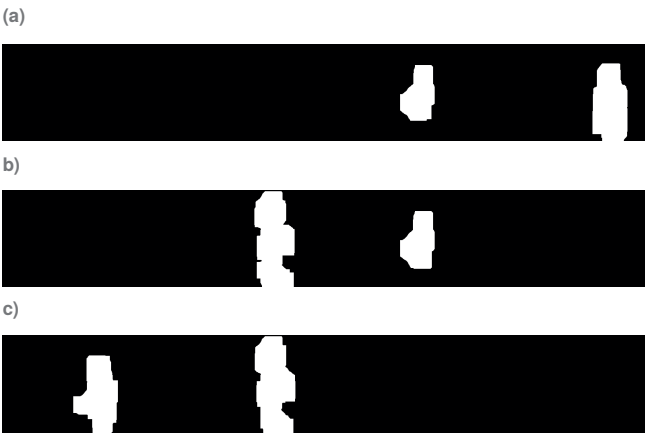
Na rysunku 5 zaprezentowano całą serię obrazów poddanych wstępnemu przetwarzaniu, czyli normalizacji, zwiększeniu rozdzielczości oraz filtracji. Sylwetki na obrazach wciąż zawierają artefakty, choć są one nieznacznie rozmyte względem surowego obrazu po korekcji.

Na rysunku 6 przedstawiono zbinaryzowane obrazy różnic bezwzględnych między każdą parą kolejnych obrazów. Obszary aktywne na obrazie binarnym (piksele o wartości logicznej 1) w znacznej mierze odpowiadają rzeczywistemu położeniu sylwetki obserwowanej osoby. Należy jednak zauważyć, że wykryte

regiony są lokalnie mniejsze od faktycznych wymiarów sylwetki, co wynika z zastosowania progowania adaptacyjnego oraz wpływu operacji morfologicznych.



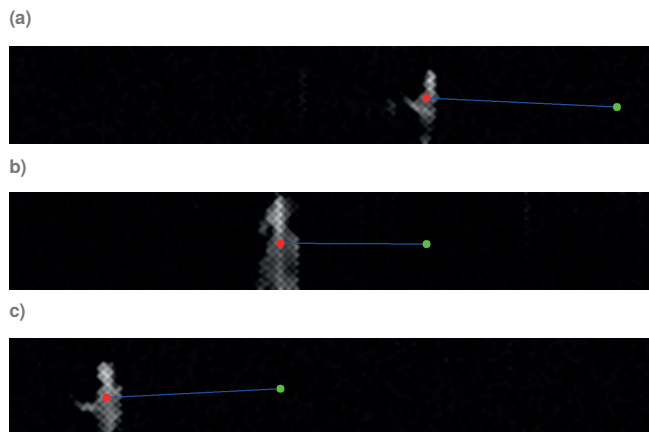
Rys. 5. Seria czterech kolejnych panoramicznych obrazów termowizyjnych po wstępnym przetwarzaniu (normalizacja, zwiększenie rozdzielczości, filtracja) w kolejności rejestracji: (a) pierwszy obraz, (b) drugi obraz, (c) trzeci obraz, (d) czwarty obraz
Fig. 5. Series of four consecutive thermal panoramic images after preprocessing (normalization, resolution enhancement, filtering) in recording order: (a) first image, (b) second image, (c) third image, (d) fourth image



Rys. 6. Zbinaryzowane obrazy różnic bezwzględnych po operacjach morfologicznych między kolejnymi parami obrazów: (a) różnica między obrazami 1-2, (b) różnica między obrazami 2-3, (c) różnica między obrazami 3-4
Fig. 6. Binarized absolute difference images after morphological operations between consecutive image pairs: (a) difference between images 1-2, (b) difference between images 2-3, (c) difference between images 3-4

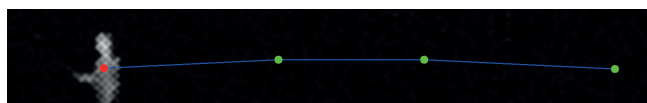
Na rysunku 7 zaznaczono ruch wykrytego na podstawie różnic między obrazami obiektu. Środki ciężkości wyznaczone w oparciu o zbinaryzowane obrazy odzwierciedlają rzeczywiste punkty lokalizacji obiektu na obrazie. Z kolei na rysunku 8 pokazano całkowitą trajektorię ruchu obiektu na przestrzeni całego pomiaru.

W tabeli 1 zestawiono klasyfikację wykryć oraz wyznaczony parametr $F1$, określający skuteczność systemu pod względem zdolności do detekcji obiektu. W większości przypadków skutecznie wykryto obiekt na podstawie przyjętego kryterium. W pozostałych przypadkach częściej wystąpił brak detekcji, niż detekcja fałszywa.



Rys. 7. Trajektorie ruchu wykrytego obiektu (niebieski) wyznaczone z kolejnych par obrazów z zaznaczonymi pozycjami: aktualną (czerwony) i poprzednią (zielony): (a) trajektoria 1-2, (b) trajektoria 2-3, (c) trajektoria 3-4

Fig. 7. Motion trajectories of detected object (blue) determined from consecutive image pairs with marked positions: current (red) and previous (green): (a) trajectory 1-2, (b) trajectory 2-3, (c) trajectory 3-4



Rys. 8. Całkowita trajektoria ruchu obiektu (niebieski) wykryta w całej zarejestrowanej serii obrazów z zaznaczonymi punktami detekcji (zielony) i ostatnią pozycją obiektu (czerwony)

Fig. 8. Complete motion trajectory of object (blue) detected throughout entire recorded image series with marked detection points (green) and final object position (red)

Tabela 1. Zestawienie parametrów skuteczności detekcji obiektów w systemie skanującym z niskorozdzielczą kamerą termowizyjną (TP – prawdziwie pozytywne, FP – fałszywie pozytywne, FN – fałszywie negatywne, F1 – parametr F1)

Table 1. Summary of object detection performance parameters in scanning system with low-resolution thermal camera (TP – true positive, FP – false positive, FN – false negative, F1 – F1 score)

Parametr	TP	FP	FN	F1
Wartość	13	5	2	0,79

Tabela 2. Zestawienie parametrów lokalizacji kątowej obiektów (MAAE – średni błąd bezwzględny kąta, RMSAE – błąd średniokwadratowy kąta, $\Delta\theta_{\min}$ – minimalny błąd kątowy, $\Delta\theta_{\max}$ – maksymalny błąd kątowy)

Table 2. Summary of angular localization parameters (MAAE – mean absolute angular error, RMSAE – root mean square angular error, $\Delta\theta_{\min}$ – minimum angular error, $\Delta\theta_{\max}$ – maximum angular error)

Parametr	MAAE	RMSAE	$\Delta\theta_{\min}$	$\Delta\theta_{\max}$
Wartość [°]	4,72	6,03	1,17	18,70

W tabeli 2 zamieszczono wyniki badania skuteczności lokalizacji obiektu przez system. Średni błąd lokalizacji kątowej ma wartość poniżej 5°. W niektórych pomiarach wystąpiły odchyłki od wartości oczekiwanej kąta powyżej 10°, co przekłada się na większą wartość błędu średniokwadratowego, która wynosi 6,03°.

4. Wnioski

W ramach niniejszej pracy opracowano i zweryfikowano eksperymentalnie algorytm detekcji i lokalizacji obiektów dedykowany niskorozdzielczym kamerom termowizyjnym. Głównym osiągnięciem jest demonstracja możliwości skutecznej detek-

cji ruchu przy użyciu matrycy o rozdzielczości zaledwie 32×24 piksele, co stanowi znaczące ograniczenie w porównaniu z konwencjonalnymi kamerami termowizyjnymi.

Uzyskany w badaniach parametr $F1$ na poziomie 0,79 należy oceniać w kontekście fundamentalnych ograniczeń sprzętowych, w szczególności ekstremalnie niskiej rozdzielczości matrycy (32×24 piksele), oraz przyjętego założenia o stosowaniu klasycznych, obliczeniowo tanich algorytmów przetwarzania obrazu [17]. W porównaniu do zaawansowanych systemów detekcji opartych na sieciach neuronowych (np. YOLO) i kamerach wysokiej rozdzielczości, które osiągają skuteczność $F1$ na poziomie 97 % [18], uzyskany wynik jest znacząco niższy. Różnica ta wynika jednak bezpośrednio z odmiennych założeń projektowych. W niniejszej pracy priorytetem była minimalizacja kosztów i złożoności obliczeniowej, a nie maksymalizacja metryk skuteczności.

Bardziej wartościowe jest porównanie z systemami o zbliżonej charakterystyce. Przykładowo, w systemie TADAR, wykorzystującym również niskorozdzielczą matrycę termowizyjną, osiągnięto $F1$ na poziomie 88,8 % dla detekcji wielu użytkowników [19]. Z kolei wdrożenie komercyjnych kamer termowizyjnych do detekcji pieszych w warunkach miejskich pozwoliło na uzyskanie wysokich wartości $F1$ oraz dokładności sięgającej 90 % [22]. W tym kontekście wynik 0,79 uzyskany dla modelu systemu skanującego, obciążonego dodatkowymi artefaktami wynikającymi z ruchu kamery, można uznać za obiecujący. Potwierdza on, że zaproponowana metoda stanowi krok w kierunku opracowania efektywnych i ekonomicznych systemów monitoringu, stanowiąc kompromis między skutecznością a kosztem wdrożenia.

Średni błąd lokalizacji kątowej obiektów wyniósł 4,72°, przy błędzie średniokwadratowym 6,03°, co potwierdza przydatność metody do aplikacji monitoringu o umiarkowanych wymaganiach dotyczących ich precyzji.

Istotnym wkładem metodologicznym jest zastosowanie dwupunktowej korekcji niejednorodności w połączeniu z sekwencyjną filtracją medianową i gaussowską, co pozwoliło na skuteczną kompensację szumów charakterystycznych dla niskorozdzielczych detektorów termowizyjnych. Szczególnie istotne okazało się dostosowanie parametrów operacji morfologicznych do specyfiki artefaktów wynikających z architektury odczytu „wzorem szachownicowym” sensora mlx90640 [21].

Praktyczna realizacja systemu skanującego potwierdziła wykonalność implementacji opracowanego algorytmu na platformie o ograniczonych zasobach obliczeniowych (ESP32-s3). System umożliwił utworzenie panoramicznego obrazu termowizyjnego o kącie widzenia 180° przy zachowaniu funkcjonalności detekcji i śledzenia obiektów w czasie quasi-rzeczywistym.

Ograniczenia metody obejmują wrażliwość na parametry filtracji oraz występowanie artefaktów ruchowych w scenariuszach dynamicznych. Architektura odczytu sensora mlx90640 wprowadza charakterystyczny szum geometryczny, który wymaga specjalizowanego przetwarzania i może ograniczać skuteczność detekcji szybko poruszających się obiektów.

Na skuteczność detekcji i lokalizacji obiektu wpływa szereg parametrów samego przetwarzania obrazu. Wśród najbardziej kluczowych parametrów są: wymiary masek filtrów, wymiar maski operacji morfologicznych oraz liczba ich iteracji. Wielkość maski filtru medianowego musiał być znacznie większa od filtru Gaussa, w celu kompensacji obszarów zaszumionych powiększonych na skutek zwiększenia rozdzielczości. Samo zwiększenie rozdzielczości jest kluczowe, ponieważ umożliwia prawidłowe działanie filtru Gaussa.

Parametry erozji i dylatacji muszą powodować wyodrębnienie obszarów zajmowanych przez obiekt na obrazie. Z uwagi na charakterystyczny dla niskorozdzielczej kamery tego typu szum w kształcie szachownicy, liczba iteracji dylatacji musiała być

większa niż erozji, W innym przypadku sylwetka rejestrowanej osoby po progowaniu i binaryzacji była błędnie lokalizowana.

Wyniki pracy wskazują, że niskorozdzielcze kamery termowizyjne, przy odpowiednim przetwarzaniu obrazu, mogą stanowić alternatywę dla konwencjonalnych systemów monitoringu w aplikacjach o umiarkowanych wymaganiach dokładnościowych, głównie ze względu na niski koszt i łatwość implementacji. Otwiera to nowe możliwości dla rozwoju kompaktowych systemów IoT w podczerwieni.

Podziękowania

Praca została dofinansowana przez Wojskową Akademię Techniczną w ramach projektu nr UGB 531-000080-W600-2.

Bibliografia

- Portocarrero J.M.T., Delicato F.C., Pires P.F., Gámez N., Fuentes L., Ludovino D., Ferreira P., *Autonomic Wireless Sensor Networks: A Systematic Literature Review*, "Journal of Sensors", Vol. 2014, DOI: 10.1155/2014/782789.
- Diaz S., Mendez D., Kraemer R., *A Review on Self-Healing and Self-Organizing Techniques for Wireless Sensor Networks*, "Journal of Circuits, Systems and Computers", Vol. 28, No. 5, 2019, DOI: 10.1142/S0218126619300058.
- Bieszczad G., Sosnowski T., Sawicki K., Gogler S., Ligienza A., Mścichowski M., *Sieć miniaturowych czujników termowizyjnych do wykrywania i śledzenia obiektów*, „Pomiary Automatyka Robotyka”, T. 25, Nr 242, 2021, 57–66, DOI: 10.14313/PAR_242/57.
- Chiu S.-Y., Tseng Y.-C., Chen J.-J., *Low-Resolution Thermal Sensor-Guided Image Synthesis*, IEEE/CVF Winter Conference on Applications of Computer Vision Workshops (WACVW), 2023, 60–69, DOI: 10.1109/WACVW58289.2023.00011.
- Bielecki Z., Rogalski A., *Detekcja sygnałów optycznych*, Wydawnictwo naukowe PWN, 2021.
- Więcek B., De Mey G., *Termowizja w Podczerwieni. Podstawy i Zastosowania*, Wydawnictwo Naukowe PWN, 2017.
- Minkina W., *Pomiary termowizyjne: przyrządy i metody*, Wydawnictwa Politechniki Częstochowskiej, 2004, ISBN: 978-83-7193-237-3.
- Howell J., Mengüç M.P., Siegel R., Daun K., *Thermal Radiation Heat Transfer*, CRC Press Taylor & Francis Group, 2020, DOI: 10.1201/9780429327308.
- Krupiński M., Bieszczad G., Sosnowski T., Madura H., Gogler S., *Nonuniformity correction in microbolometer array with temperature influence compensation*, "Metrology and Measurement Systems", Vol. 21, No. 4, 2014, DOI: 10.2478/mms-2014-0050.
- Sosnowski T., Kastek M., Sawicki K., Ligienza A., Gogler S., Więcek B., *High-Accuracy Calibration Method of a Thermal Camera Using Two Reference Blackbodies*, "Sensors", Vol. 24, No. 17, 2024, DOI: 10.3390/s24175831.
- Smith S.W., *The scientist and engineer's guide to digital signal processing*, USA: California Technical Publishing, 1997, ISBN: 978-09-6601-763-3.
- Kim H.J., Shrestha A., Sapkota E., Pokharel A., Pandey S., Kim C.S., Shrestha R., *A Study on the Effectiveness of Spatial Filters on Thermal Image Pre-Processing and Correlation Technique for Quantifying Defect Size*, "Sensors", Vol. 22, No. 22, 2022, DOI: 10.3390/s22228965.
- Lyons R., *Understanding Digital Signal Processing*, Pearson Education International, 2011, ISBN: 978-01-3211-937-5.
- Bergenroth H., *Use of Thermal Imagery for Robust Moving Object Detection Hannah*, 2021, student thesis, [https://liu.diva-portal.org].
- Gahramanova A., *Locating Centers of Mass with Image Processing*, "Undergraduate Journal of Mathematical Modeling: One + Two", Vol. 10, No. 1, 2019, DOI: 10.5038/2326-3652.10.1.4906.
- van Rijsbergen C.J., *Information Retrieval*, 2nd. Butterworth-Heinemann, 1979.
- DeVries L., Underwood M., Holtta V., Lacher A.J., Lary D.J., *Assessing thermal imagery integration into object detection methods for improved situational awareness on the flight deck*, "Scientific Reports", Vol. 13, No. 1, 2023, DOI: 10.1038/s41598-023-34791-8.
- Gupta R., Jain S., Kumar M., *An Enhanced Algorithm for Small Object Detection based on Thermal Imaging Using YOLOv8-EPB*, "Journal of Computer Science", Vol. 21 No. 6, 2025, 1391–1403, DOI: 10.3844/jcssp.2025.1391.1403.
- Zhang X., Wu C., *TADAR: Thermal Array-based Detection and Ranging for Privacy-Preserving Human Sensing*, Proceedings of the Twenty-fifth International Symposium on Theory, Algorithmic Foundations, and Protocol Design for Mobile Networks and Mobile Computing (MobiHoc'24), ACM, 2024, 11–20, DOI: 10.1145/3641512.3686357.
- Chidurala V., Li X., *Detection of moving objects using thermal imaging sensors for occupancy estimation*, "Internet of Things", Vol. 17, 2022, DOI: 10.1016/j.iot.2021.100487.

Inne źródła

- Melexis, *MLX90640 32 × 24 IR array*, Datasheet, 2019, [www.melexis.com/-/media/files/documents/datasheets/mlx90640-datasheet-melexis.pdf].
- U.S. Department of Transportation, *Connected Vehicle Pilot Deployment Program-Phase 4, New York City Department of Transportation-Pedestrian Detection Analysis Report*, Intelligent Transportation Systems Joint Program Office, FHWAJPO-24-100, 2024.

Detection and Localization of Objects Using Area Scanning with a Low-Resolution Thermal Camera

Abstract: Modern monitoring and detection systems frequently employ thermographic technologies, which enable observation of objects under low-light conditions or adverse weather. However, the effective detection and localization of objects using low-resolution thermal cameras remains a significant challenge. Despite their hardware limitations, such cameras offer numerous advantages, including low cost, compact size, and energy efficiency. This paper presents a custom developed algorithm for thermal image processing, designed for low-resolution cameras. It also describes and evaluates a scanning system that implements the developed algorithm. The system is based on a thermal imaging camera with a 32×24 pixel resolution and was constructed using commercially available electronic components. Particular attention is given to the optimization of filtering, binarization, and morphological operations, which reduce noise, enhance detection quality, and enable precise determination of the location of detected objects. The results of conducted experiments demonstrate that even with a low-resolution sensor, it is possible to obtain useful information regarding the movement and localization of objects. This opens new possibilities for the development of compact and cost-effective infrared monitoring systems.

Keywords: thermal camera, image processing, motion detection

mgr inż. Karol Erd
karol.erd@wat.edu.pl
ORCID: 0009-0007-0532-6578

Ukończył w 2023 r. Wojskową Akademię Techniczną w Instytucie Optoelektroniki. Zajmuje się zagadnieniami związanymi z przetwarzaniem obrazu, ich zastosowaniem w technice podczerwieni, a także zdalnym pomiarem temperatury. Obecnie kształci się w szkole doktorskiej WAT i realizuje pracę doktorską w dziedzinie techniki terahercowej.



dr inż. Krzysztof Sawicki
krzysztof.sawicki@wat.edu.pl
ORCID: 0000-0002-1368-3854

Ukończył Wydział Elektroniki Wojskowej Akademii Technicznej w 2009 r. Rozprawę doktorską z tematyki steganografii w sieciach bezprzewodowych obronił w WAT w 2019 r. Jego zainteresowania skupiają się na sieciach bezprzewodowych, bezpieczeństwie teleinformatycznym, steganografii i systemach wbudowanych oraz ich zastosowaniu w technice podczerwieni i termowizji.

