

Porównanie dwóch klasyfikatorów opartych na metodzie DTW i połączonych metodach PCA-LDA do rozpoznawania obiektów morskich na podstawie obrazów FLIR

Streszczenie. Praca dotyczy rozpoznawania obiektów morskich na podstawie ich obrazów wykonanych przez sensory podczerwieni (FLIR – forward looking infra-red). W pracy zaproponowano wykorzystanie połączonych transformacji analizy głównych składowych PCA (Principal Component Analysis) i liniowej analizy dyskryminacyjnej LDA (Linear Discriminant Analysis) oraz transformaty dopasowania szeregów czasowych DTW (Dynamic Time Warping) wraz z metodą wyznaczania odległości dwóch szeregów czasowych za pomocą algorytmu programowania dynamicznego.

Abstract. This paper presents methods of recognition of maritime objects based on FLIR (forward looking infra-red) sensor images. Two methods of recognition are proposed: a method of combined transforms of the Principal Component Analysis PCA with the Linear Discriminant Analysis LDA and the DTW (Dynamic Time Warping) transform with the method of determining the distance of two time series by means of a dynamic programming algorithm. (Comparison of two classifiers based on the method of comparing time series DTW and the combined PCA-LDA methods to recognize maritime objects upon FLIR images)

Słowa kluczowe: rozpoznawanie obrazowe w podczerwieni, klasyfikator PCA-LDA, klasyfikator DTW.

Keywords: FLIR images recognition, PCA-LDA classifier, DTW classifier.

Wstęp

Pasywne sensory podczerwieni FLIR (forward looking infra-red) wykorzystywane do rozpoznawania obiektów morskich rozmieszczane są na powietrznych platformach rozpoznania i okrętach. Używane są przede wszystkim do rozpoznania krótkiego i średniego zasięgu i identyfikacji obiektów nawodnych (statków i okrętów). Sensor FLIR dostarcza cyfrowe obrazy monochromatyczne, w których każdy piksel scharakteryzowany jest jedną liczbą – luminancją piksela, mającą interpretację temperatury punktu.

Zadanie rozpoznawania obiektów morskich na podstawie obrazów FLIR sprowadza się do odpowiedzi na pytanie, czy obiekt zawarty w obrazie jest rzeczywiście obiektem morskim (w szczególności możemy zawęzić zbiór interesujących nas obiektów do pewnego podzbioru), a w przypadku uzyskania odpowiedzi pozytywnej oczekujemy odpowiedzi na kolejne pytanie, czy rozpoznawany obiekt należy do którejś z klas obiektów wcześniej sklasyfikowanych, a jeśli tak, to do której. Obraz wejściowy poddawany wyżej opisanej procedurze jest najczęściej zniekształcony na skutek zaszumienia, a także wykonania go z różnej odległości i pod różnym kątem obserwacji. Na jakość obrazu mają również wpływ charakterystyki kamery FLIR. To wszystko powoduje, iż proces rozpoznawania jest procesem wieloetapowym, a wyniki rozpoznania noszą cechy przypadkowości. Niemniej jednak obrazy tego samego obiektu wykonywane w różnych warunkach mają cechy wspólne, które zwane są cechami dystynktywnymi. Ich wartości decydują o odpowiedziach na wcześniej sformułowane pytania, w szczególności o wyznaczeniu wartości odległości rozpoznawanego obiektu od obiektów wzorcowych przechowywanych w bazie wzorców. Cechy dystynktywne obiektów tworzą przestrzeń wartości cech dystynktywnych. Postać przestrzeni cech dystynktywnych zależy od przyjętej metody rozpoznawania i przyjętego klasyfikatora. Należy podkreślić, że w literaturze przedmiotu można znaleźć definicje kilku klasyfikatorów, które mogą być użyte do rozpoznawania obiektów za pośrednictwem sensorów FLIR [10,12,13,14,24,25,26].

W pracy wykorzystano dwie metody klasyfikacji sylwetek obiektów morskich: metodę liniowej analizy dyskryminacyjnej Fishera (LDA – Linear Discrimination Analysis) [19,20,21,22,28] połączoną z metodą analizy

głównych składowych (PCA – Principal Component Analysis) [3,5,6,9,28] zwaną dalej klasyfikatorem PCA-LDA oraz metodę porównywania szeregów czasowych DTW (Data Time Warping), zwaną w pracy klasyfikatorem DTW [15, 16, 17].

Zastosowanie transformacji PCA pozwala na zmniejszenie wymiarowości przestrzeni cech dystynktywnych analizowanych obiektów. Metoda PCA pozwala na uszeregowanie cech obrazów (wartości jasności pikseli) w ciąg nierosnący od pikseli o największej rozróżnialności (największej wariancji cechy jasności w zbiorze obrazów wzorcowych) począwszy do pikseli o najmniejszej rozróżnialności (najmniejszej wariancji cechy jasności w zbiorze obrazów wzorcowych). W naturalny sposób można z obrazu wyeliminować piksele, dla których wariancja jasności w zbiorze obrazów wzorcowych wynosi zero. O eliminacji kolejnych pikseli o wariancji jasności większej niż zero należy zdecydować arbitralnie. Proces eliminacji pikseli jest odpowiednikiem stratnej kompresji obrazów i prowadzi do utraty informacji. Zbiór pikseli, które pozostały w obrazie, nazywa się, w przypadku rozpoznawania sylwetek, sylwetką własną. Zmiana bazy obrazów wzorcowych prowadzi najczęściej do zmiany sylwetek własnych. Metoda PCA bazuje na liniowym przekształceniu Loewa-Karhunen przestrzeni $N=m \times n$ wymiarowej w przestrzeń N' wymiarową ($N' < m \times n$).

Pierwszą metodą rozpoznawania zastosowaną w pracy jest klasyfikacja metodą liniowej analizy dyskryminacyjnej Fishera połączoną ze wstępnym użyciem metody PCA. Metoda Fishera została opracowana przez R. Fishera w roku 1936 do klasyfikacji taksonomicznej. Podstawową trudnością w zastosowaniu metody Fishera do rozpoznawania obrazów jest występująca najczęściej osobliwość wewnątrzklasowej macierzy rozproszenia. Wynika to z faktu, iż liczba obrazów w zbiorze treningowym jest zdecydowanie mniejsza niż liczba pikseli w każdym obrazie. W roku 1997 Belhumeur i in. [19] zaproponowali rozwiązanie problemu małej liczności zbioru treningowego poprzez zmniejszenie wymiarowości każdego obrazu. Dokonali tego za pomocą projekcji zbioru obrazów w przestrzeń o dużo mniejszej wymiarowości. Metodę tę nazwali metodą obrazów Fishera (Fisherfaces). Istota rozwiązania sprowadza się do wstępnego zastosowania transformacji PCA w celu zmniejszenia wymiarowości przestrzeni cech a

następnie użycie standardowej analizy dyskryminacyjnej Fishera. W pracach [20-22] wykazano, że takie rozwiązanie ma pewne wady i zaproponowano użycie tak zwanych bezpośrednich metod liniowej analizy dyskryminacyjnej. W niniejszej pracy wykorzystano jednakże metodę obrazów Fishera (metodę połączonych transformacji PCA i LDA) i porównano jej skuteczność z metodą PCA (obrazów własnych).

Stosując wybraną miarę odległości w przestrzeni sylwetek Fishera możemy określić, który z obiektów wzorcowych położony jest najbliżej w stosunku do analizowanego obiektu. Należy zauważyć, że wyżej omówiona procedura może być zastosowana do sytuacji, gdy obrazy zebrane w bazie wzorców są pogrupowane w klasy reprezentujące typy obiektów, a zadanie rozpoznania polega na rozpoznaniu typu obiektu. Jako punkty reprezentujące klasy w przestrzeni wzorców przyjmuje się punkty średnie w klasie [2].

Drugą metodą rozpoznawania zastosowaną w pracy jest klasyfikacja metodą porównania szeregów czasowych DTW. Jako cechy dystynktywne sylwetek przyjęto histogram pionowego rzutu jasności pikseli na oś poziomą. Histogram ten w dalszej kolejności został wykorzystany do wyznaczania odległości porównywanego obrazu FLIR z obrazami zawartymi w bazie wzorców. Podstawą tego procesu jest algorytm porównywania szeregów czasowych DTW (Dynamic Time Warping) oparty na metodzie programowania dynamicznego. Algorytm ten określa globalną odległość dwóch szeregów czasowych. Szczególną cechą tego algorytmu jest możliwość dokonania tej oceny, gdy przebieg czasu jednego z szeregów został zdeformowany (rozciągnięty lub skompresowany). Ta ostatnia cecha jest szczególnie przydatna w analizie obrazów obiektów, które zostały uzyskane pod kątem obserwacji różnym od zera. Punkty startowe metody DTW są związane z punktami, w których pojawia się niezerowa wartość histogramu, natomiast punkty końcowe związane są z punktami, w których znika niezerowa wartość histogramu. Odległości między punktami końcowymi dla obu analizowanych histogramów nie muszą być takie same. Jedynym problemem do rozwiązania pozostaje całkowita niezgodność kierunku ustawienia obu obiektów. Można go rozwiązać poprzez porównanie analizowanej sylwetki z sylwetką wzorcową i jej lustrzanym odbiciem (symetria względem prostej).

Klasyfikator DTW jest jednowymiarowym klasyfikatorem odległościowym z odległością wyznaczaną przez algorytm DTW.

Wszystkie klasyfikatory wymagają, aby obrazy FLIR poddawane procesowi klasyfikacji zostały wcześniej poddane procesowi wstępnego przetwarzania, którego celem jest wyeliminowanie zbędnych informacji dotyczących tła obiektu oraz zakłóceń, a także normalizacja sylwetki obiektu. Proces wstępnego przetwarzania obrazu może obejmować segmentację, normalizację jasności, przeskalowanie sylwetki, centrowanie sylwetki oraz poziomowanie sylwetki.

W końcowej części pracy zostały przedstawione wyniki badań klasyfikatorów PCA-LDA i DTW.

Model procesu rozpoznania obrazowego obiektów

Proces klasyfikacji dzieli się zwykle na cztery podstawowe etapy:

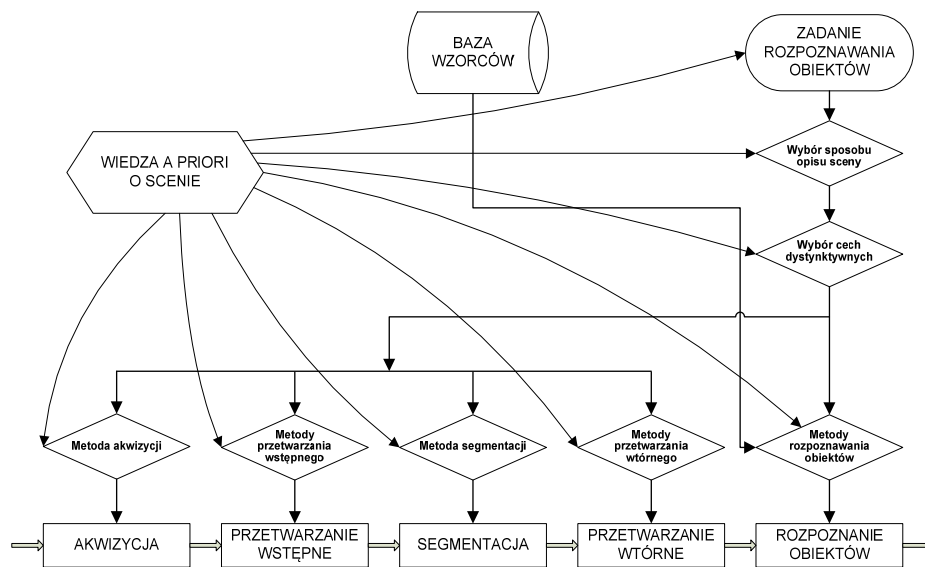
- wstępnego przetwarzania obrazu,
- segmentacji,
- ekstrakcji cech (wyznaczania wartości cech dystynktywnych),
- klasyfikacji właściwej.

Na rysunku 1 został przedstawiony dokładny model procesu klasyfikacji obiektów morskich na podstawie obrazów FLIR.

W ramach czynności wykonywanych podczas przetwarzania wstępnego obrazów FLIR dokonywane są między innymi następujące czynności: przekształcenie obrazu do postaci obrazu achromatycznego w skali szarości, usunięcie elementów stałych obrazów FLIR (skala temperatury, dane opisujące stan otoczenia, logo producenta, znak środka obrazu), przeskalowanie obrazu.

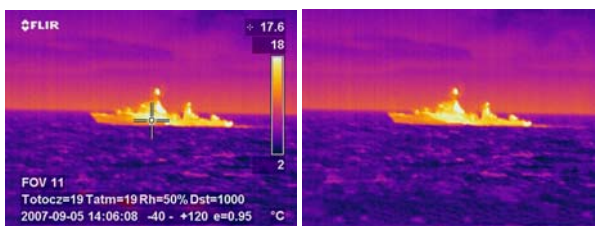
W ramach przetwarzania wtórnego realizowane są operacje poprawiające kształt sylwetki obiektu morskiego (m.in. operacje morfologiczne), poziomowanie sylwetki obiektu oraz ewentualnie jej powtórne przeskalowanie.

Obrazy rozpoznawane w pracy zostały poddane przetwarzaniu wstępnemu i segmentacji, którego efektem była sylwetka obiektu morskiego w skali szarości. Proces segmentacji był wspomagany komputerowo za pomocą modułu segmentacji metodą Otsu. Niestety w większości przypadków wynik działania algorytmu z automatycznie określonym progiem był niezadawalający, pozostawiono więc możliwość ręcznej korekty wartości progów.

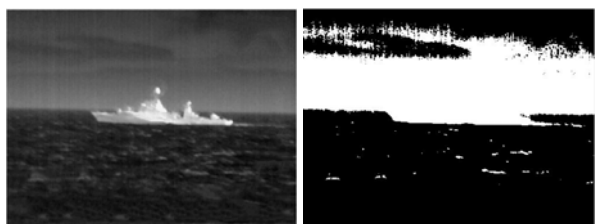


Rys.1. Model procesu rozpoznania obrazowego obiektów

Na rysunkach 2 – 6 zaprezentowano wyniki realizacji kolejnych etapów przetwarzania obrazu pierwotnego z kamery FLIR.



Rys.2. Pierwotny obraz z kamery FLIR oraz z usuniętymi zbytecznymi artefaktami



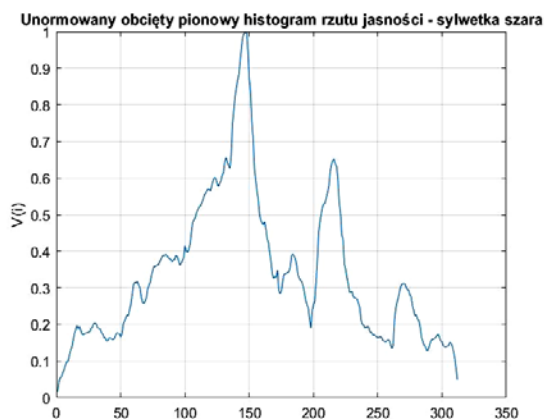
Rys.3. Obraz w skali szarości oraz obraz po wykonaniu progowania binarnego z progiem wyznaczonym przez algorytm Otsu



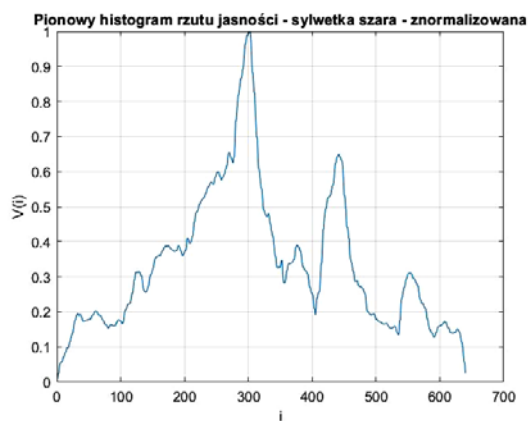
Rys.4. Obraz po wykonaniu progowania binarnego z progiem określonym ręcznie oraz obraz po wykonaniu operacji usunięcia artefaktów typu „pieprz i sól” oraz zamknięcia



Rys.5. Iloczyn obrazu binarnego i obrazu w skali szarości oraz obraz po wykonaniu operacji normalizacji wysokości sylwetki i centrowania



Rys. 6. Unormowany obcięty pionowy histogram rzutu jasności sylwetki w skali szarości



Rys. 7. Unormowany histogram pionowego rzutu jasności sylwetki w skali szarości przeskalowany w poziomie do 640 pikseli

Na rysunku 7 przedstawiono ostateczny wektor 1×640 zawierający wartości cech dystyngtywnych obiektu morskiego podawany na wejście klasyfikatora DTW.

Określenie sylwetek obiektów morskich na podstawie obrazów FLIR w klasyfikatorze PCA

Niech obraz sylwetki obiektu morskiego $I(x,y)$ będzie dwuwymiarową tablicą $m \times n$ zawierającą liczby całkowite 8-bitowe określające wartość jasności każdego piksela (obraz achromatyczny). Tablicę tę przekształca się najczęściej w wektor o wymiarze $(m \cdot n) \times 1$ [1,7,8], stąd dla rozmiaru obrazu 480×640 otrzymamy wektor o długości 307.200 lub ekwiwalentnie punkt w przestrzeni 307.200 wymiarowej. Zbiór obrazów obiektów morskich, które będziemy traktować jako obrazy wzorcowe stanowić będzie zbiór punktów tej przestrzeni. Punkty te nie są rozłożone w przestrzeni przypadkowo, ponieważ posiadają pewne podobieństwo, stąd mogą być opisane w przestrzeni o mniejszej liczbie wymiarów. Główną ideą metody PCA lub przekształcenia Loeve-Karhunen jest znalezienie takiego zbioru wektorów ortonormalnych, które najlepiej opisują taką podprzestrzeń.

W pierwszej kolejności metoda PCA zapewnia możliwość zbudowania nowej przestrzeni, w której dowolny rozpoznawany obraz sylwetki będzie liniową kombinacją wektorów określanych jako sylwetki własne bazy sylwetek wzorcowych. Wektory te są wektorami własnymi macierzy kowariancji macierzy zbudowanej z kolumn – obrazów bazy obrazów wzorcowych. Poniżej przedstawiono sposób wyznaczania sylwetek własnych obrazów zawartych w zbiorze treningowym zwanym w pracy bazą obrazów wzorcowych.

Każdy wejściowy wektor x_i jest skonstruowany poprzez konkatencję kolejnych kolumn obrazu w skali szarości zawierającego obiekt morski.

Niech X będzie zbiorem obrazów wzorcowych x_1, x_2, \dots, x_M . Każdy obraz jest wektorem o długości $N=m \cdot n$, stąd

$$(1) \quad X = \{x_i \in \{\mathbf{N} \cup \{0\}\}^N, i = 1, \dots, M\}.$$

W [7,8] sformułowano wymagania dotyczące obrazów w bazie wzorców – obrazy powinny być takich samych rozmiarów, a środek sylwetki powinien być w środku obrazu.

Wektor średni wyznaczany jest zgodnie ze wzorem

$$(2) \quad \bar{x} = \frac{1}{M} \sum_{i=1}^M x_i.$$

Każdy obraz r_i i różni się od obrazu średniego o wektor

$$(3) \quad r_i = x_i - \bar{x}.$$

Zbiór tych wektorów jest przedmiotem analizy głównych składowych PCA [28]. Wyznaczane są wektory własne u_i i wartości własne λ i macierzy kowariancji

$$(4) \quad C = \frac{1}{M} \sum_{i=1}^M r_i r_i^T = AA^T,$$

gdzie macierz $A = [r_1 r_2 \dots r_L]$.

Wektory własne oraz wartości własne spełniają następujące równania

$$(5) \quad Cu_i = \lambda_i u_i \text{ oraz } AA^T u_i = \lambda_i u_i.$$

Macierz kowariancji C ma wymiar $N \times N$ i jest symetryczną macierzą liczb rzeczywistych i określa $N=m \cdot n$ wektorów własnych i wartości własnych. Ze względu na rozmiar macierzy C w praktycznych zagadnieniach problem obliczeniowy staje się bardzo złożony, więc należy znaleźć inną metodę wyznaczania wektorów własnych.

Zamiast posługiwać się macierzą kowariancji $C = AA^T$ rozważmy macierz $A^T A$ i jej wektory własne v_i ($i=1, \dots, M$) [1,7,8]. Z definicji możemy napisać

$$(6) \quad A^T A v_i = \mu_i v_i \quad i=1, \dots, M.$$

Mnożąc obie strony przez macierz A otrzymujemy

$$(7) \quad AA^T A v_i = \mu_i A v_i \quad (i=1, \dots, M),$$

co możemy zapisać jako

$$(8) \quad C A v_i = \mu_i A v_i \quad i=1, \dots, M.$$

Porównując (5) i (8) można napisać, że

$$(9) \quad \lambda_i = \mu_i$$

$$(10) \quad u_i = A v_i \quad i=1, \dots, M.$$

Z (10) wynika, że

$$(11) \quad AA^T A v_i = \mu_i A v_i \quad i=1, \dots, M,$$

$$(12) \quad u_i^j = \sum_{k=1}^M r_k^j v_i^k = \sum_{k=1}^M (x_k^j - \bar{x}^j) v_i^k, \quad i=1, \dots, M, \quad j=1, \dots, N.$$

Wektory $u_i, i=1, \dots, M$, są wektorami ortogonalnymi, stąd należy je znormalizować, aby stały się wersorami osi w nowym układzie współrzędnych:

$$(13) \quad u_i := \frac{u_i}{\|u_i\|}.$$

W przypadku gdy liczność bazy obrazów jest duża, w procesie rozpoznawania można wykorzystać część wektorów własnych odpowiadających największym wartościom własnym. Przyjmujemy, że w dalszej części pracy, że M' oznacza liczbę wykorzystanych wektorów własnych, a tym samym wymiarowość przestrzeni sylwetek.

Każdą znormalizowaną sylwetkę z bazy wzorców (zbioru treningowego) można przedstawić jako liniową kombinację M' sylwetek własnych w przestrzeni M' -wymiarowej:

$$(14) \quad r_i = x_i - \bar{x} = \sum_{j=1}^{M'} y_j^i u_j \quad i=1, \dots, M'.$$

Stąd można napisać w postaci iloczynu skalarnego

$$(15) \quad u_j^T r_i = u_j^T \sum_{j=1}^{M'} y_j^i u_j = u_j^T y_1^i u_1 + u_j^T y_2^i u_2 + \dots + u_j^T y_j^i u_j + \dots + u_j^T y_{M'}^i u_{M'} = y_j^i$$

ponieważ

$$(16) \quad u_j^T u_i = \begin{cases} 1 & \text{gdy } i = j \\ 0 & \text{gdy } i \neq j \end{cases}.$$

Z (14) wynika, że

$$(17) \quad y_j^i = u_j^T r_i = u_j^T (x_i - \bar{x}), \quad i=1, \dots, M, \quad j=1, \dots, M'.$$

Każda znormalizowana sylwetka z bazy wzorców jest reprezentowana przez wektor:

$$(18) \quad y^{iT} = [y_1^i, y_2^i, \dots, y_{M'}^i], \quad i=1, 2, \dots, M,$$

który można uzyskać poprzez liniową transformację znormalizowanej sylwetki r_i :

$$(19) \quad y^i = U_E^T r_i = U_E^T (x_i - \bar{x}), \quad i=1, 2, \dots, M.$$

Charakterystyka klasyfikatora LDA

Mimo że teoretycznie możliwe byłoby bezpośrednie zastosowanie liniowej analizy dyskryminacyjnej Fishera LDA (Fisher's Linear Discriminant Analysis) do rozpoznawania sylwetki obiektu morskiego jako wektora $N=m \cdot n$, to jednak w praktyce okazuje się to zwykle niemożliwe ze względu to, że wewnątrzklasowa macierz rozproszenia jest macierzą osobiwą, co uniemożliwia jej odwrócenie. Spowodowane to jest tzw. problemem małej liczności zbioru treningowego (ang. *small sample size problem* – SSS *problem*) w porównaniu z rozmiarem wektora [20-22,28]. Tradycyjne rozwiązanie problemu SSS [20] wymaga włączenia metody PCA jako wstępnego etapu przetwarzania w metodzie LDA, mającego na celu zmniejszenie wymiarowości problemu, aby usunąć zerową podprzestrzeń w macierzy rozproszenia wewnątrzklasowego zbioru treningowego. Po tej operacji transformacja LDA jest realizowana w podprzestrzeni PCA o mniejszej liczbie wymiarów [19]. Jak wykazano w [21,22] zmniejszenie wymiarowości przestrzeni może prowadzić do utraty informacji dotyczących separacji w zbiorze treningowym. Dlatego też zaproponowano tam tzw. metody bezpośrednie D-LDA, w których obrazy przetwarzane są bezpośrednio w przestrzeni pierwotnej.

W niniejszej pracy zastosowano algorytm obrazów własnych jako wstępny etap przetwarzania informacji obrazowej w celu zmniejszenia wymiarowości problemu. Innymi słowy wyjściowe wektory metody obrazów własnych są wektorami wejściowymi metody LDA. Takie podejście do problemu rozpoznawania obrazów zwane jest metodą obrazów Fishera [19]. Zakładamy, że przedmiotem transformacji Fishera są obrazy własne uzyskane za pomocą przekształcenia PCA i określone za pomocą wzoru (18).

W pierwszej kolejności należy określić średnią zbioru treningowego (bazy wzorców) \bar{y} i średnie \bar{y}_i każdej z L klas obrazów Y_i .

$$(20) \quad \bar{y} = \frac{1}{L} \sum_{y \in Y} y$$

$$(21) \quad \bar{y}_i = \frac{1}{L_i} \sum_{y \in Y_i} y$$

W dalszej kolejności wyznaczamy macierz rozprożeń międzyklasowych

$$(22) \quad S_B = \sum_{i=1}^L L_i (\bar{y}_i - \bar{y})(\bar{y}_i - \bar{y})^T$$

oraz macierz rozprożeń wewnątrzklasowych

$$(23) \quad S_W = \sum_{i=1}^L \sum_{x_i \in X_i} (y_i - \bar{y}_i)(y_i - \bar{y}_i)^T$$

gdzie L jest liczbą klas a L_i jest licznością l -tej klasy Y_l .

Jeżeli macierz S_W jest nieosobliwa, to optymalna transformacja (projekcja) U_F jest wyznaczana jako macierz ortonormalnych kolumn, która maksymalizuje stosunek wyznacznika macierzy rozproszenia międzyklasowego rzutowanego zbioru treningowego do wyznacznika macierzy rozproszenia wewnątrzklasowego rzutowanego zbioru treningowego, to znaczy

$$(24) \quad U_F = \arg \max_U \frac{U^T S_B U}{U^T S_W U} = [u_1 \ u_2 \ \dots \ u_K],$$

przy czym zbiór $\{u_i : i=1, \dots, K\}$ jest zbiorem wektorów własnych odpowiadających K największym wartościom własnym $\{\lambda_i : i=1, \dots, K\}$ wyznaczonym z rozwiązania równania

$$(25) \quad S_B u_i = \lambda_i S_W u_i, \quad i=1, \dots, K.$$

Należy zauważyć, że istnieje co najwyżej $L-1$ niezerowych wartości własnych [19].

Zakładając, że w etapie wstępnym przetwarzania zastosowano metodę PCA, możemy przyjąć, iż wektory wejściowe charakteryzują się średnią $\bar{y} = \theta$. Ostatecznie transformację opisaną metodą obrazów Fishera możemy przedstawić w następującej postaci

$$(26) \quad z^i = U_F^T r_i = U_F^T (y_i - \bar{y}) = U_F^T y_i, \quad i=1, 2, \dots, K$$

Wykorzystanie sylwetek własnych obiektów morskich w procesie klasyfikacji ich obrazów za pomocą klasyfikatora PCA-LDA

Proces rozpoznawania nowego obrazu jest realizowany w czterech etapach:

1. Wyznaczenie nowego obrazu znormalizowanego:

$$(27) \quad r_n = x_n - \bar{x}$$

2. Projekcja obrazu znormalizowanego r_x na przestrzeń własną bazy wzorców sylwetek PCA – wyznaczenie składowych wektora \hat{r}_n w nowym układzie współrzędnych

$$(28) \quad \hat{r}_n = \sum_{j=1}^{M'} y_j^n u_j$$

przy czym składowe wektora wag y_j^n określone są za pomocą iloczynów skalarnych:

$$(29) \quad y_j^n = u_j^T \hat{r}_n$$

$$(30) \quad y^n = [y_1^n, y_2^n, \dots, y_{M'}^n].$$

3. Projekcja obrazu własnego y^n w przestrzeń obrazów Fishera LDA:

$$(31) \quad z^n = U_F^T y^n.$$

4. Znajdujemy

$$(32) \quad i^* = \arg \min_{i=1, \dots, L} \|z^n - \bar{z}_i\|.$$

i stwierdzamy, że obraz x_n został rozpoznany jako obraz klasy i^* w bazie wzorców, przy czym \bar{z}_i jest średnią każdej z L klas obrazów Z_i

Jako odległość w przestrzeni własnej bazy wzorców można przyjąć odległość euklidesową:

$$(33) \quad \|z^n - \bar{z}_i\|_E = \sum_{j=1}^K (z_j^n - \bar{z}_j^i)^2.$$

Natomiast w [7,28] stwierdzono, iż należy rozważyć zastosowanie odległości Mahalanobisa:

$$(34) \quad \|z^n - \bar{z}_i\|_M = \sum_{j=1}^K \frac{1}{\lambda_j} (z_j^n - \bar{z}_j^i)^2,$$

przy czym może być to niemożliwe do zastosowania w przypadku, gdy macierz kowariancji C_F jest macierzą osobliwą. Wartości λ_j są wartościami własnymi tej macierzy.

Charakterystyka klasyfikatora DTW

W niniejszej pracy przyjęto, że cechy dystynktywne obiektu stanowią wartości histogramu pionowego rzutu jasności sylwetki obiektu.

Histogram pionowego rzutu jasności powstaje w wyniku sumowania poziomu jasności pikseli w poszczególnych kolumnach i wyraża się następującym wzorem:

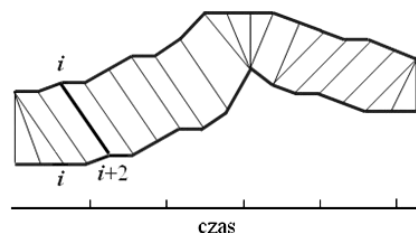
$$(35) \quad LV[j] = \sum_{i=1}^h L(i, j) \text{ dla } j=1, \dots, w,$$

gdzie: $L(i, j)$ – jasność punktu o współrzędnych (i, j) obrazu achromatycznego lub binarnego, LV – pionowy rzut jasności kolumn (wektor),

i – numer wiersza, j – numer kolumny, w – szerokość obszaru (obrazu), h – wysokość obszaru (obrazu).

Na rys. 5 i 6 przedstawiono sylwetkę obrazu obiektu morskiego oraz jego pionowy histogram rzutu jasności.

Do oceny odległości między histogramami wykorzystano metodę DTW porównywania szeregów czasowych (Dynamic Time Warping – Dynamiczne Dopasowanie Czasu), która określa klasę algorytmów służących do porównywania szeregów (ciągów) wartości. Zaletą metody DTW jest możliwość określenia podobieństwa dwóch szeregów czasowych przy założeniu, że oba mogą opisywać różne przedziały czasu oraz zmienna czasu może być lokalnie przedziałowo skompresowana lub rozciągnięta. Odległość między dwoma szeregami czasowymi jest określana po ewentualnym skompresowaniu bądź rozciągnięciu czasu dla kojarzonych punktów szeregów czasowych. zilustrowane to zostało na rysunku 8.



Rys.8. Dopasowanie dwóch szeregów czasowych – nieliniowe (elastyczne) dopasowanie określa miarę podobieństwa

Celem metody DTW jest porównanie dwóch ciągów (najczęściej zależnych od chwil czasowych, a więc szeregów czasowych)

$$(36) \quad LV_1 = \{x_1, x_2, \dots, x_N\}$$

o długości $N \in \mathcal{N}$ oraz

$$(37) \quad LV_2 = \{y_1, y_2, \dots, y_M\}$$

o długości $M \in \mathcal{N}$.

Te szeregi mogą być w ogólności ciągami wartości pewnych cech zmierzonych w jednakowych odstępach czasu. Oznaczmy przestrzeń cech przez F . Wówczas możemy zapisać, że $x_i, y_j \in F$ dla $i=1, \dots, N$ oraz $j=1, \dots, M$.

Aby porównać dwie wartości x_i oraz y_j , należy dysponować lokalną miarą odległości zwaną lokalną miarą kosztu, która może być zdefiniowana w sposób następujący:

$$(38) \quad c : F \times F \rightarrow \mathbb{R}_{\geq 0}.$$

Miara ta określa brak dopasowania (brak podobieństwa) obu punktów. Można również wprowadzić funkcję operującą na zbiorach indeksów i oraz j określoną w sposób następujący:

$$(39) \quad d(i, j) = c(x_i, y_j).$$

Funkcja d określa macierz $D_{N \times M}$ odległości skrośnych między wszystkimi kombinacjami par elementów ciągów X oraz Y :

$$(40) \quad D[i, j] = d(i, j).$$

Funkcja odległości może być na przykład określona jako

$$(41) \quad d(i, j) = |x_i - y_j|$$

lub

$$(42) \quad d(i, j) = (x_i - y_j)^2.$$

Funkcja DTW pozwala określić odległość $DTW(LV_1, LV_2)$ dwóch szeregów czasowych na podstawie macierzy D .

Jako klasyfikator DTW w procesie wyboru wzorców z bazy danych wzorców został przyjęty klasyfikator minimalnoodległościowy, przy czym odległość dwóch obrazów określa odległość dwóch połączonych histogramów tych obrazów wyznaczona metodą DTW.

Przyjmijmy następujące oznaczenia:

\overline{LV}_j – średni histogram pionowego rzutu jasności j -tego typu obrazu w bazie wzorców,

LV_r – histogram pionowego rzutu jasności rozpoznawanego obrazu,

$DTW(LV_r, \overline{LV}_j)$ – odległość między histogramem pionowego rzutu jasności rozpoznawanego obrazu oraz średnim histogramem pionowego rzutu jasności j -tego typu obrazu w bazie wzorców.

Zadanie rozpoznania obiektu jest sprowadzone do wyznaczenia takiego numeru typu j^* w bazie wzorców, aby

$$(43) \quad j^* = \arg \min_{j \in NRTW} DTW(LV_r, \overline{LV}_j),$$

gdzie $NRTW$ jest zbiorem numerów typów obrazów wzorców.

Wyniki rozpoznawania obiektów morskich za pomocą klasyfikatora PCA-LDA

W bazie wzorców znajdowało się 110 obrazów FLIR siedmiu typów obiektów morskich. Dla wszystkich obrazów wykonano czynności wstępnego przetwarzania obrazów,

segmentacji, przetwarzania wtórnego oraz wyznaczono histogram pionowego rzutu jasności.

Badanie polegało na wyborze kolejnych różnych obrazów z bazy danych, a następnie wykonaniu procedury rozpoznawania poprzez wyznaczenie odległości LDA między obrazem sylwetki rozpoznawanego obiektu a obrazami pozostałych sylwetek w bazie traktowanych jako obrazy wzorcowe. Procedura wyznaczania odległości zawierała transformację PCA obrazów pierwotnych a następnie transformację otrzymanych wyników za pomocą przekształcenia Fishera LDA. W takiej przestrzeni wyników została określona odległość między transformowanymi obrazami. Wybierany jest obraz wzorcowy o najmniejszej odległości LDA. Jeżeli typ wybranego obrazu jest zgodny z typem rozpoznawanego obrazu, to uznajemy, że obraz został rozpoznany poprawnie, w przeciwnym przypadku obraz został rozpoznany błędnie. Wyniki zostały przedstawione w poniższej tabeli.

Tabela 1. Wyniki rozpoznawania obrazów obiektów morskich za pomocą klasyfikatora PCA-LDA

Nr typu obiektu morskiego	Liczba testowanych obrazów	Liczba poprawnie rozpoznanych obiektów morskich	Liczba błędnie rozpoznanych obiektów morskich*	Wartość średnia poprawnie rozpoznanych obiektów morskich
1	14	4	10	29%
2	28	26	2	93%
3	16	11	5	69%
4	18	5	13	28%
5	10	7	3	70%
6	15	15	0	100%
7	9	6	3	67%
Łącznie	110	76	34	69%

*W algorytmie rozpoznawania nie przyjęto żadnej odległości dyskryminacyjnej, stąd zawsze rozpoznawano pewien typ obiektu.

Tabela 2. Wyniki rozpoznawania obrazów obiektów morskich za pomocą klasyfikatora DTW

Nr typu obiektu morskiego	Liczba testowanych obrazów	Liczba poprawnie rozpoznanych obiektów morskich	Liczba błędnie rozpoznanych obiektów morskich*	Wartość średnia poprawnie rozpoznanych obiektów morskich
1	14	14	0	100%
2	28	22	16	57%
3	16	16	0	100%
4	18	12	6	67%
5	10	10	0	100%
6	15	3	12	20%
7	9	3	6	33%
Łącznie	110	70	40	64%

*W algorytmie rozpoznawania nie przyjęto żadnej odległości dyskryminacyjnej, stąd zawsze rozpoznawano pewien typ obiektu.

Wyniki rozpoznawania obiektów morskich dla klasyfikatora DTW

W bazie wzorców znajdowały się identyczne obrazy FLIR jak podczas badań klasyfikatora PCA-LDA.

Badanie polegało na wyborze kolejnych różnych obrazów z bazy danych, a następnie wykonaniu procedury rozpoznawania poprzez wyznaczenie odległości między histogramem rozpoznawanego obiektu a histogramami pozostałych obrazów w bazie traktowanych jako obrazy wzorcowe. Wybierany jest obraz wzorcowy o najmniejszej odległości DTW histogramów. Jeżeli typ wybranego obrazu jest zgodny z typem rozpoznawanego obrazu, to uznajemy, że obraz został rozpoznany poprawnie, w przeciwnym przypadku obraz został rozpoznany błędnie. Wyniki zostały przedstawione w poniższej tabeli.

Podsumowanie

W tabeli 3 przedstawiono porównanie skuteczności rozpoznawania obiektów morskich na podstawie zdjęć FLIR za pomocą klasyfikatorów PCA-LDA i DTW. Porównując te dwa klasyfikatory należy zauważyć zbliżoną ich średnią skuteczność (69% i 64%), przy czym wyniki dla poszczególnych typów klasyfikatorów znacznie się różnią dla poszczególnych typów obiektów. Może to wynikać z różnych sposobów określania wartości cech dystynktywnych, w istocie różnych algorytmów kompresji informacji o jasności poszczególnych punktów sylwetek obiektów morskich.

Średnie wyniki nie odbiegają zasadniczo od wyników prezentowanych w badaniach z zastosowaniem innych metod segmentacji oraz innych klasyfikatorów [13].

Tabela 3. Porównanie skuteczności rozpoznawania obrazów obiektów morskich za pomocą klasyfikatorów PCA-LDA i DTW

Nr typu obiektu morskiego	Liczba testowanych obrazów	Skuteczność klasyfikatora PCA-LDA	Skuteczność klasyfikatora DTW
1	14	29%	100%
2	28	93%	57%
3	16	69%	100%
4	18	28%	67%
5	10	70%	100%
6	15	100%	20%
7	9	67%	33%
Łącznie	110	69%	64%

W przypadku klasyfikatora DTW zastanawiająco słabe wyniki uzyskano dla szóstego typu obiektu. Zastosowanie połączonego klasyfikatora PCA-LDA pozwoliło osiągnąć lepsze rezultaty niż klasyfikator DTW. Analizując powyższe wyniki należy również mieć na uwadze stosunkowo małą liczbę zbioru treningowego (bazy wzorców) i występujące w niej zdjęcia o złej jakości. Zdjęcia złej jakości, w których temperatura punktów tła jest zbliżona do temperatury sylwetki powoduje istotną trudność uzyskania zadawalającego obrazu sylwetki binarnej obiektu. Wydaje się, że może mieć to istotny wpływ na skuteczność rozpoznawania sylwetek.

Dalsze kierunki prac w dziedzinie rozpoznawania obiektów morskich na podstawie obrazów FLIR mogą uwzględniać zastosowanie:

- doskonalszych metod wstępnego przetwarzania obrazów, przede wszystkim segmentacji, np. zastosowanie metod rekurencyjnych uwzględniających unimodalność histogramów jasności i automatyzację procesu segmentacji,
- metod poprawy jakości obrazu segmentu poprzez zastosowanie operacji morfologicznych,
- innych klasyfikatorów liniowych, np. D-LDA (Direct Linear Discriminant Analysis), ICA (Independent Component Analysis) i SVM (Support Vector Machine),
- innych metod oceny odległości histogramów projekcji poprzez zastosowanie algorytmów GTW (General Time Warping) i GCTW (Generalized Canonical Time Warping),
- głębokich sieci neuronowych [26,27].

Zaprezentowane wcześniej wyniki pozwalają wyciągnąć wniosek, że prace dotyczące rozpoznawania obiektów morskich na podstawie obrazów FLIR powinny być kontynuowane i powinny prowadzić do opracowania systemu rozpoznania obrazowego współpracującego z kamerami FLIR w trybie on-line. System taki powinien posiadać następujące funkcjonalności:

- akwizycja obrazów z kamery FLIR w trybie on-line,
- wstępne przetwarzanie obrazów,

- określanie sylwetek obiektów morskich,
- tworzenie i edycja bazy danych obrazów i ich cech dystynktywnych,
- rozpoznawanie obiektów morskich uwzględniające różne typy klasyfikatorów oraz metody ich fuzji.

Autor: dr inż. Tadeusz Pietkiewicz, Wojskowa Akademia Techniczna, Wydział Elektroniki, Instytut Radioelektroniki, ul. Gen. S. Kaliskiego 2, 00-908 Warszawa, E-mail: tadeusz.pietkiewicz@wat.edu.pl.

LITERATURA

- [1] Atalay I., Face Recognition Using Eigenface, *Istanbul Technical University, Turkey*, 1996. URL http://www.ilkeratalay.com/download/eigenfaces_msc_thesis.pdf (dostęp 20.10.2014 r.)
- [2] Gouaillieri V., Gagnon L., Ship Silhouette Recognition Using Principal Components Analysis, *SPIE Proc. #3164, Conference "Applications of Digital Image Processing XX"*, San Diego, 1997
- [3] Jafri R., H. Arabina R., A Survey of Face Recognition Techniques, *Journal of Information Processing Systems*, Vol.5, No.2, June 2009
- [4] Smith L. I., A tutorial on principal components analysis, February 2002. URL http://www.cs.otago.ac.nz/cosc453/student_tutorials/principal_components.pdf (dostęp 20.10.2014 r.)
- [5] Tat-Jun Chin, Suter D., A Study of the Eigenface Approach for Face Recognition, *MECSE-6-2004, Monash University*, 2004
- [6] Turk M., Pentland A., Eigenfaces for Recognition, *Journal of Cognitive Neuroscience*, vol. 3, no. 1, 1991, 71-86. URL <http://www.cs.ucsb.edu/~mturk/Papers/jcn.pdf> (dostęp 20.10.2014 r.)
- [7] Turk M., Pentland A., Face Recognition Using Eigenface. *Proc. IEEE Conference on Computer Vision and Pattern Recognition. IEEE*, June 1991, 586-591. URL <http://www.cs.wisc.edu/~dye/cs540/handouts/mturk-CVPR91.pdf> (dostęp 20.10.2014 r.)
- [8] Zhao W., Chellappa R., Phillips P.J., Rosenfeld A., Face Recognition: A Literature Survey, *ACM Computing Surveys*, Vol. 35, No. 4, December 2003, 399-458
- [9] Gonzalez R.C., Woods R.E., Digital Image Processing, Ed. 3, *Pearson Prentice Hall*, 2008
- [10] Jouan A., Valin P., Bossé É., Concepts of data/information fusion for naval C2 and airborne ISR platforms, *Defence R&D Canada - Valcartier, Technical Report DRDC Valcartier*, TR 2004-284, November 2006
- [11] Malina W., Smiatacz M., Cyfrowe przetwarzanie obrazów, *Akademicka Oficyna Wydawnicza EXIT*, Warszawa 2008
- [12] Park Y., Sklansky J., Automated Design of Linear Tree Classifiers, *Pattern Recognition*, Vol. 23, No. 12, 1990, 1393-1412
- [13] Valin P., Methods for the Fusion of Multiple FLIR Classifiers, [w] *Proceedings of the Workshop on Signal Processing, Communication, Chaos and Systems: a Tribute to Rabinder N. Madan*, June 20, 2002, Newport, RI, 2002, 117-122
- [14] Valin P., Bossé É., Jouan A., Airborne application of information fusion algorithms to classification, *Defence R&D Canada - Valcartier, Technical Report DRDC Valcartier*, TR 2004-282, May 2006
- [15] Müller M., *Information Retrieval for Music and Motion*, Springer-Verlag Berlin Heidelberg, 2007, 69-84
- [16] Teisseyre P., Porównanie szeregów czasowych z wykorzystaniem algorytmu DTW. *Instytut Podstaw Informatyki, Polska Akademia Nauk*
- [17] Giorgino T., Computing and Visualizing Dynamic Time Warping Alignments in R: The DTW Package. *Journal of Statistical Software*, vol. 31, iss. 7, August 2009
- [18] Smiatacz M., Eigenfaces, Fisherfaces, Laplacianfaces, Marginfaces - how to face the face verification task. *Proceedings of the 8th International Conference on Computer Recognition Systems CORES 2013. Advances in Intelligent Systems and Computing*, vol 226. Springer, Heidelberg, 187-196. doi: 10.1007/978-3-319-00969-8_18
- [19] Belhumeur P.N.; Hespanha J.P., Kriegman D.J., Eigenfaces vs. Fisherfaces: recognition using class specific linear

- projection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, iss. 7, Jul 1997, 711 – 720
- [20] Juwei Lu, Plataniotis K.N., Venetsanopoulos A.N., Face recognition using LDA-based algorithms. *IEEE Transactions on Neural Networks*, volume: 14, Issue: 1, Jan 2003, 195 – 200
- [21] Chen L.-F., Mark Liao H.-Y., Ko M.T., Lin J.-C., and Yu G.-J., A new LDA-based face recognition system which can solve the small sample size problem. *Pattern Recognition*, vol. 33, 2000, 1713–1726
- [22] Yu H., Yang J., A direct LDA algorithm for high-dimensional data with application to face recognition. *Pattern Recognition*, vol. 34, 2001, 2067–2070
- [23] Pietkiewicz T., A method of recognition of maritime objects based on FLIR (forward looking infra-red) sensor images using dynamic time warping. *Proc. SPIE 10434, Electro-Optical Remote Sensing XI, 1043409* (5 October 2017); doi: 10.1117/12.2278419
- [24] Pietkiewicz T., *Application of fusion of two classifiers based on principal component analysis method and time series comparison to recognize maritime objects upon FLIR images*. Proc. SPIE 11055, XII Conference on Reconnaissance and Electronic Warfare Systems, 110550Z (27 March 2019); doi: 10.1117/12.2524975.
- [25] Pietkiewicz T., Sikorska-Łukasiewicz K., Comparison of two classifiers based on neural networks and the DTW method of comparing time series to recognize maritime objects upon FLIR images. *Proc. SPIE 11055, XII Conference on Reconnaissance and Electronic Warfare Systems, 110550V* (27 March 2019); doi: 10.1117/12.2524918.
- [26] Pietkiewicz T., Matuszewski J., Recognition of maritime objects based on FLIR images using the method of eigenimages. *IEEE Xplore Digital Library, 2018 14th International Conference on Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering (TCSET)*, Lviv-Slavske, Ukraine, 20-24 Feb. 2018, doi: 10.1109/TCSET.2018.8336345.
- [27] Matuszewski J., Radar signal identification using a neural network and pattern recognition methods. *IEEE Xplore Digital Library, 2018 14th International Conference on Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering (TCSET)*, Lviv-Slavske, Ukraine, 29-24 Feb. 2018, 79-83, doi: 10.1109/TCSET.2018.8336160.
- [28] Osowski S., *Metody i narzędzia eksploracji danych*, BTC, Legionowo, 2013, 217-231