

Anna Czechowicz*, Zbigniew Mikrut**

Wykorzystanie sieci Kohonena do selekcji podobrazów na potrzeby dopasowania zdjęć lotniczych***

1. Wprowadzenie

Wyznaczenie wzajemnego położenia zdjęć tworzących stereogram w momencie ich naświetlania, zwane orientacją wzajemną modelu, to jeden z podstawowych etapów przygotowania modelu fotogrametrycznego. Etap ten jest szczególnie ważny, gdyż poprzez usunięcie paralaksy poprzecznej (różnicy współrzędnych tłowych obrazów tego samego punktu dla zdjęć lotniczych, wskazującej na pochylenie obydwu lub jednego zdjęcia albo na różną wysokość lotu podczas wykonywania tych zdjęć [5]), możliwa jest obserwacja stereoskopowa.

W praktyce zagadnienie to rozwiązywane jest najczęściej manualnie. Obserwator wybiera promienie homologiczne (niosące obraz tego samego punktu w obu wiązkach), w sześciu obszarach zdjęcia, tzw. rejonach Grubera. Poprzez doprowadzenie do przecięcia minimum pięciu par promieni, określanych jest pięć parametrów kąto-liniowych informujących o wzajemnym położeniu pary zdjęć stereogramu.

Automatyzacja tego procesu jest jednym z kluczowych problemów fotogrametrii. W typowej procedurze można wymienić osiem etapów [1]:

1. pozyskanie piramidy obrazów dla obu zdjęć,
2. wyznaczenie pokrycia oraz przybliżonych wartości obrotu i różnicy skali między zdjęciami na najwyższym poziomie piramidy,
3. pozyskanie cech,
4. dopasowanie tych cech,
5. wyznaczenie zgrubnych wartości parametrów orientacji wzajemnej,
6. powtórzenie procedur pozyskania i dopasowania cech oraz określenia parametrów na coraz niższych poziomach piramidy obrazów w celu zwiększenia dokładności,
7. określenie położenia punktów homologicznych z podpikselową dokładnością,
8. obliczenie elementów orientacji wzajemnej.

* Zakład Fotogrametrii i Informatyki Teledetekcyjnej, Akademia Górniczo-Hutnicza w Krakowie; eire02@op.pl

** Katedra Automatyki, Akademia Górniczo-Hutnicza w Krakowie; zibi@agh.edu.pl

*** Praca finansowana z grantu Komitetu Badań Naukowych nr 4T12E 001 27

Wśród tych etapów najistotniejsze i najtrudniejsze są: etap pozyskania cech na obrazie oraz poszukiwanie ich odpowiednika na pozostałych obrazach. Pozyskiwanymi cechami są punkty, linie oraz powierzchnie. Wśród stosowanych metod pasowania wyróżnić można najpopularniejszą korelację krzyżową, pasowanie metodą najmniejszych kwadratów (*Least Squares Matching*), *Feature Based Matching* lub *Relational Matching* [1].

Celem niniejszej pracy jest zaproponowanie reprezentacji fragmentu obrazu opartej na informacji o rozkładzie wartości gradientu oraz jego kierunku i sprawdzenie jej przydatności do selekcji podobrazów przy pomocy sieci neuronowej Kohonena na potrzeby pasowania zdjęć lotniczych. Uzasadnieniem wykorzystania sieci Kohonena jest fakt, iż ma ona zdolność do grupowania sygnałów wejściowych na kategorie z zachowaniem topologii, co powoduje bliskie umiejscowienie na mapie cech elementów podobnych – w tym przypadku obszarów o podobnych reprezentacjach [7]. Przejrzysta wizualizacja wyników ułatwi obserwację grupowania obszarów należących do tych samych klas. Tym samym pozwoli stwierdzić czy proponowana reprezentacja jest wystarczająca by na zdjęciu lotniczym wskazać korzystne pod względem wyszukiwania cech do orientacji wzajemnej obszary.

W badaniach wykorzystano zdjęcia lotnicze okolic Krakowa. Doświadczenia przeprowadzono przy pomocy pakietu MATLAB 7.1 [4] rozszerzonego o SOM (*Self Organizing Maps*) Toolbox Teuvo Kohonena [6].

2. Przygotowanie reprezentacji obrazu

Dla obrazu cyfrowego wyznaczana jest wartość amplitudy gradientu oraz jego kierunek. Wartości te obliczane są zgodnie z algorytmem Canny'ego [3], na który składają się dwa etapy:

1. wstępny, obejmujący redukcję szumów funkcją Gaussa,
2. właściwy, w którym wyznaczany jest gradient odszumionego obrazu.

Macierz konwolucji filtru uśredniającego Gaussa dla wartości odchylenia standardowego 1.0 przedstawia rysunek 1a [4].

a)	b)	c)																											
<table style="width: 100%; border-collapse: collapse;"> <tr><td style="padding: 2px 10px;">0,0751</td><td style="padding: 2px 10px;">0,1238</td><td style="padding: 2px 10px;">0,0751</td></tr> <tr><td style="padding: 2px 10px;">0,1238</td><td style="padding: 2px 10px;">0,2042</td><td style="padding: 2px 10px;">0,1238</td></tr> <tr><td style="padding: 2px 10px;">0,0751</td><td style="padding: 2px 10px;">0,1238</td><td style="padding: 2px 10px;">0,0751</td></tr> </table>	0,0751	0,1238	0,0751	0,1238	0,2042	0,1238	0,0751	0,1238	0,0751	<table style="width: 100%; border-collapse: collapse;"> <tr><td style="padding: 2px 10px;">0.3679</td><td style="padding: 2px 10px;">0</td><td style="padding: 2px 10px;">-0.3679</td></tr> <tr><td style="padding: 2px 10px;">0.6065</td><td style="padding: 2px 10px;">0</td><td style="padding: 2px 10px;">-0.6065</td></tr> <tr><td style="padding: 2px 10px;">0.3679</td><td style="padding: 2px 10px;">0</td><td style="padding: 2px 10px;">-0.3679</td></tr> </table>	0.3679	0	-0.3679	0.6065	0	-0.6065	0.3679	0	-0.3679	<table style="width: 100%; border-collapse: collapse;"> <tr><td style="padding: 2px 10px;">0.3679</td><td style="padding: 2px 10px;">0.6065</td><td style="padding: 2px 10px;">0.3679</td></tr> <tr><td style="padding: 2px 10px;">0</td><td style="padding: 2px 10px;">0</td><td style="padding: 2px 10px;">0</td></tr> <tr><td style="padding: 2px 10px;">-0.3679</td><td style="padding: 2px 10px;">-0.6065</td><td style="padding: 2px 10px;">-0.3679</td></tr> </table>	0.3679	0.6065	0.3679	0	0	0	-0.3679	-0.6065	-0.3679
0,0751	0,1238	0,0751																											
0,1238	0,2042	0,1238																											
0,0751	0,1238	0,0751																											
0.3679	0	-0.3679																											
0.6065	0	-0.6065																											
0.3679	0	-0.3679																											
0.3679	0.6065	0.3679																											
0	0	0																											
-0.3679	-0.6065	-0.3679																											

Rys. 1. Współczynniki macierzy konwolucji: a) filtru uśredniającego Gaussa; b) gradientu w kierunku poziomym; c) gradientu w kierunku pionowym

Po usunięciu zakłóceń z obrazu wyznaczany jest gradient w dwóch prostopadłych do siebie kierunkach. Aproksymacja wartości pierwszej pochodnej obrazu w kierunku poziomym i pionowym odbywa się z wykorzystaniem filtrów przedstawionych kolejno na rysunkach 1b i 1c [4]. W wyniku konwolucji odszumionego obrazu z odpowiednią maską filtru-

jąca powstają obrazy gradientu poziomego i pionowego, oznaczone w dalszej części jako S_x oraz S_y .

Wyznaczenie wartości amplitudy gradientu i jej kierunku to proste operacje punktowe obliczane zgodnie z równaniami (1) oraz (2):

$$M(k, l) = \sqrt{S_x(k, l)^2 + S_y(k, l)^2} \quad (1)$$

$$\Theta(k, l) = \arctg S_y(k, l) / S_x(k, l) \quad (2)$$

gdzie:

- M – amplituda gradientu,
- Θ – kierunek amplitudy gradientu,
- k, l – współrzędne piksela obrazu.

2.1. Dwuwymiarowy histogram gradientu

Reprezentacja obrazu w postaci dwuwymiarowego histogramu gradientu \mathbf{W} sporządzana jest dla wszystkich pikseli obrazu, które są pikselami krawędzi. Przyjęto, że krawędź występuje w miejscach, dla których wartość amplitudy gradientu jest większa od 60. Pojedyncze pole dwuwymiarowego histogramu gradientu odpowiada liczbie pikseli na obrazie, dla których wartości gradientu oraz jego kierunku mieszczą się w zadanym przedziale

$$w_{ij} = \sum_{k=1}^{k=K} \sum_{l=1}^{l=L} p_{ij}(k, l) \quad i = 1, 2, \dots, n \quad j = 1, 2, \dots, m \quad (3)$$

przy czym

$$p_{ij}(k, l) = \begin{cases} 1 & \text{gdy } M(k, l) = i \wedge \Theta(k, l) = j \\ 0 & \text{w przeciwnym przypadku} \end{cases} \quad (4)$$

gdzie

- w_{ij} – wartość dwuwymiarowego histogramu gradientu,
- n – liczba wierszy dwuwymiarowego histogramu gradientu,
- m – liczba kolumn dwuwymiarowego histogramu gradientu,
- K, L – rozmiar obrazu wejściowego.

Na podstawie tak przygotowanego histogramu wyznaczany jest wektor cech w postaci:

- wektora wartości maksymalnych gradientu – równanie (5),
- wektora wartości średnich gradientu – równanie (6),
- wektora wartości maksymalnych kierunku – równanie (7),
- wektora wartości średnich kierunku – równanie (8)

$$g_{\max}(i) = \max_{j=1, m} (w_{ij}), \quad i = 1, 2, \dots, n \quad (5)$$

$$g_{\dot{s}r}(i) = \frac{\sum_{j=1}^m (w_{ij})}{m}, \quad i = 1, 2, \dots, n \quad (6)$$

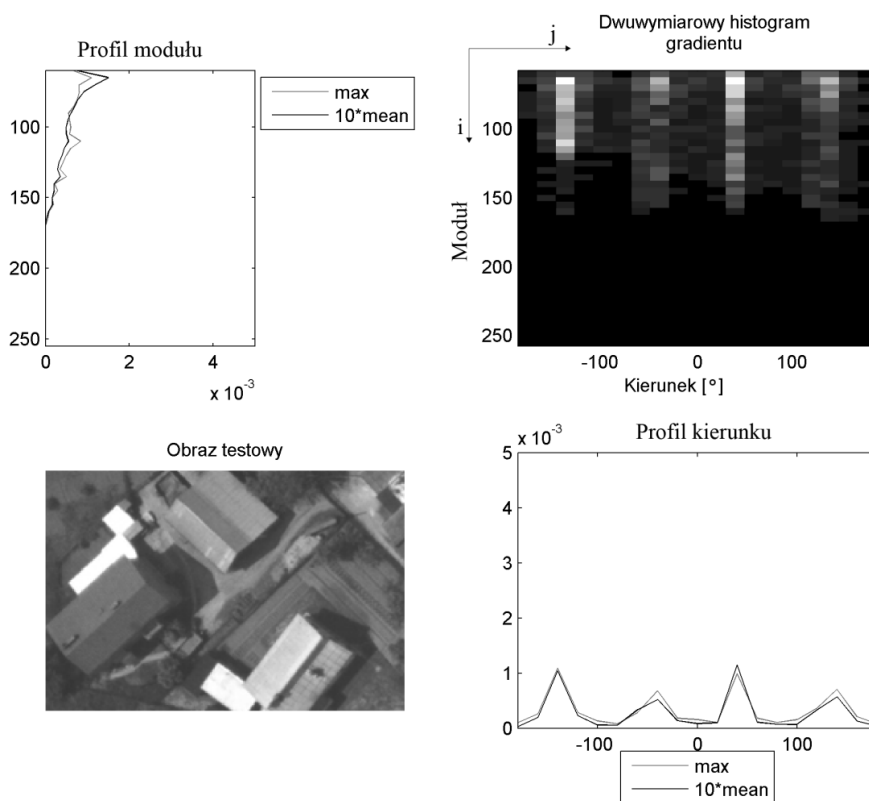
$$d_{\max}(j) = \max_{i=1, n} (w_{ij}), \quad j = 1, 2, \dots, m \quad (7)$$

$$d_{\dot{s}r}(j) = \frac{\sum_{i=1}^n (w_{ij})}{n}, \quad j = 1, 2, \dots, m \quad (8)$$

gdzie:

$g_{\max}, g_{\dot{s}r}$ – wartość maksymalna i średnia gradientu,
 $d_{\max}, d_{\dot{s}r}$ – wartość maksymalna i średnia kierunku.

Przykładowy histogram wraz z wizualizacją wektorów cech przedstawiono na rysunku 2.



Rys. 2. Przykładowy dwuwymiarowy histogram gradientu i wizualizacja wektorów cech

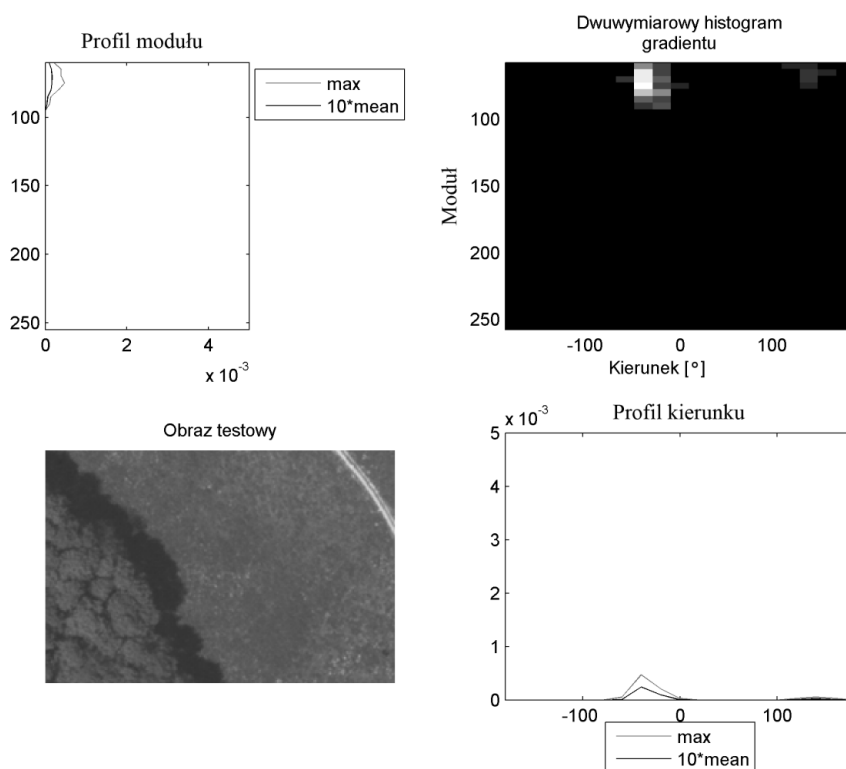
2.2. Wstępne badanie reprezentacji dla kilku obszarów

W celu wykrycia ewentualnych różnic w przebiegu reprezentacji dla obszarów różniących się pokryciem terenu, pozyskano wektory cech dla kilku fragmentów zdjęć lotniczych.

Przykładowe obszary to fragmenty zdjęć okolic Krakowa w skali szarości o wymiarze 240×160 pikseli. Obrazy podzielono na dwie grupy:

- 1) prezentujące obszary korzystne,
- 2) niekorzystne pod względem wyszukiwania cech do orientacji wzajemnej zdjęć.

Wektory cech pozyskano dla 25 fragmentów – jedenastu z obszarami korzystnymi i czternastu z obszarami niekorzystnymi. Obszary korzystne to takie, które charakteryzują się dużą ilością przebiegających pod różnymi kątami elementów liniowych oraz elementów kontrastowych. Są to np. tereny zabudowane, miejsca skrzyżowań dróg (rys. 2). Obszary niekorzystne zaś to te, na których znajdują się równoległe linie lub nie występują kontrastowe elementy. Są to np. obrazy pól uprawnych, łąk, lasów, zbiorników wodnych (rys. 3).



Rys. 3. Histogram i wektor cech dla obszaru „niekorzystnego”

Charakter reprezentacji kierunku różni się w większości przypadków dla obu typów obszarów. Zaobserwowano co najwyżej dwa maksima dla obszarów niekorzystnych (rys. 3) oraz co najmniej trzy maksima dla obszarów korzystnych (rys. 2). W przypadku wykresów

wartości gradientu nie stwierdzono wyraźnych różnic w przebiegu reprezentacji. Na tym etapie badań przyjęto więc jedynie kątowe zależności, jako wystarczające do próby podziału obszarów obrazu na korzystne i niekorzystne pod względem wyszukiwania punktów do orientacji wzajemnej. Dodatkowo spośród wektora wartości maksymalnych i średnich kierunku wybrano pierwszy, ponieważ lepiej odwzorowuje maksimum dwuwymiarowego histogramu gradientu.

Przebadano również wpływ zmniejszenia rozdzielczości obrazu i zwiększenia przedziału próbkowania na osi kierunku dwuwymiarowego histogramu gradientu na przebieg reprezentacji. Dokonano sprawdzenia przebiegu reprezentacji dla dwóch rozdzielczości (600 dpi i 300 dpi) oraz trzech wielkości przedziałów agregacji kierunku, tj. 5° , 10° oraz 20° . W wymienionych przypadkach sporządzano osobny dwuwymiarowy rozkład gradientu i na jego podstawie pozyskano wektory cech. Wielkość przedziałów agregacji kierunku związana jest z długością wektora cech. Dla przedziału 5° wektor cech składa się z 73 elementów, dla przedziału 10° z 37 elementów, a dla 20° jest wektorem 19-elementowym. Zmniejszenie rozdzielczości oraz zwiększenie przedziału agregacji kierunku ma niewielki wpływ na ogólny charakter reprezentacji – ilość pików nie ulega zmianie.

Reprezentacja obrazu w postaci 19-elementowego wektora wartości kierunku gradientu (por. wzór (5)), pozyskana z obrazu o rozdzielczości 600 dpi, została uznana za odpowiednią do próby klasyfikacji obszarów siecią Kohonena. Charakter takiej reprezentacji nie ulega znaczącej zmianie dla obrazu o dwukrotnie mniejszej rozdzielczości.

3. Krótkie wprowadzenie do sieci Kohonena

Sieć Kohonena, zwana inaczej samoorganizującym odwzorowaniem (*Self Organizing Map* – SOM), jest jedną z najpopularniejszych sieci grupujących. Zazwyczaj ma postać regularnej siatki dwuwymiarowej, w której węzłach w sposób uporządkowany przechowywane są modele pewnych obserwacji, tak iż podobne znajdują się w bliskim sąsiedztwie [2].

W sieci tej mamy do czynienia z uczeniem konkurencyjnym. Oznacza to, że po każdym podaniu sygnału wejściowego neurony rywalizują ze sobą i wygrywa ten, którego wagi najmniej różnią się od danych wejściowych. Zwycięski neuron podlega dalszemu uczeniu. Adaptacji wag mogą podlegać również sąsiednie neurony jeżeli dla sieci zdefiniowane jest sąsiedztwo.

Testowane sieci powstały w środowisku MATLAB 7.1 z wykorzystaniem funkcji biblioteki SOM Toolbox utworzonej w Laboratory of Computer and Information Science Uniwersytetu w Helsinkach¹⁾. Biblioteka umożliwia:

- tworzenie sieci o dowolnym wymiarze;
- dla sieci dwuwymiarowych definicję ich kształtu (wybór między płaską, w kształcie cylindra lub pierścienia) i układu neuronów (siatka kwadratowa lub sześciokątna);
- zdefiniowanie sąsiedztwa za pomocą kilku funkcji, np. 0–1 (*bubble*), Gaussa, ograniczonej Gaussa;

¹⁾ SOM Toolbox dostępny jest na stronie internetowej: www.cis.hut.fi/projects/somtoolbox/

- inicjalizację wag w sposób losowy (*random initialization*) i liniowy (*linear initialization*), gdzie wektory wag są dobierane w uporządkowany sposób wzdłuż podprzestrzeni liniowej rozciągniętej na dwóch głównych wektorach własnych przestrzeni danych wejściowych;
- dobór jednej z trzech funkcji zmian współczynnika uczenia: liniowej, hiperbolicznej i szeregu potęgowego;
- określenie liczby epok w dwóch fazach uczenia sieci, tj. w fazie porządkowania oraz fazie dostrajania;
- uczenie sieci dwoma algorytmami: sekwencyjnym (w każdym kroku uczenia sieci prezentowana jest wybrana w sposób losowy próbka ze zbioru uczącego, wagi zwycięzcy i neuronów sąsiednich aktualizowane są zgodnie z równaniem (9) i „batch” (w przypadku tego algorytmu mapa jest aktualizowana po prezentacji całego zbioru uczącego (po każdej epoce) zgodnie z równaniem (10))

$$m_i(t+1) = m_i + a(t)h_{ci}(r(t))[x(t) - m_i(t)] \quad (9)$$

gdzie:

- m_i – wektor wag neuronu o numerze i ,
- i – numer neuronu,
- t – czas,
- $a(t)$ – współczynnik uczenia,
- $h_{ci}(r(t))$ – otoczenie wokół neuronu zwycięzcy,
- $r(t)$ – promień sąsiedztwa,
- $x(t)$ – próbka zbioru uczącego,

$$m_i(t+1) = \frac{\sum_{j=1}^n h_{ic(j)}(t)x_j}{\sum_{j=1}^n h_{ic(j)}(t)} \quad (10)$$

gdzie:

- t – czas,
- $c(j)$ – wektor zwycięzcy dla próbki x_j ,
- $h_{ic(j)}$ – otoczenie wokół neuronu zwycięzcy,
- n – liczba próbek zbioru uczącego z tym samym neuronem zwycięzcą,
- x_j – próba zbioru uczącego;

- obliczenie współczynników kwantyzacji (qe) i topografii (te) określających efektywności nauczonej sieci. Współczynnik kwantyzacji to średnia odległość pomiędzy zbiorem próbek i jego wektorem zwycięzcy, współczynnik topografii jest proporcją wszystkich wektorów, dla których pierwszy i drugi neuron zwycięzcy nie są sąsiadami [6].

4. Proces uczenia – próba klasyfikacji obszarów

Źródłem danych są dwa zdjęcia lotnicze obszaru Krakowa w skali szarości (składowa G z obrazu kolorowego RGB) o rozdzielczości 600 dpi. Zdjęcia te tworzą stereogram lotniczy o pokryciu podłużnym 60%. Z każdego z obrazów wybrano po 452 fragmenty o wymiarze 240×160 pikseli, dla których wyznaczono wektor wartości maksymalnych dla kierunku gradientu z agregacją 20° . Każdy z fragmentów obrazu zaklasyfikowano do jednej z trzech grup: obszarów korzystnych, niekorzystnych oraz pośrednich. Uzyskano w ten sposób 904 próbki, które podzielono na dwa zbiory: uczący oraz testowy. Wektory danych znormalizowano do przedziału $0 \div 1$.

4.1. Struktura testowanych sieci Kohonena

Testy przeprowadzono dla trzech wielkości sieci płaskich o układzie heksagonalnym i następującej liczbie neuronów:

- mała – 7×4 neuronów,
- średnia – 14×8 neuronów,
- duża – 29×15 neuronów.

Dla wszystkich sieci przyjęto następujące parametry uczenia: metrykę Gaussa dla funkcji sąsiedztwa, algorytm uczenia typu *batch*, inicjalizację wag funkcją liniową oraz funkcję hiperboliczną dla współczynnika uczenia (omówione w rozdz. 3). Wartość początkowa współczynnika uczenia w fazie porządkowania wyniosła 0,5 – w fazie dostrajania natomiast 0,05. Promień sąsiedztwa malał od wartości początkowej w fazie porządkowania przyjętej na poziomie 0,25 dłuższego wymiaru sieci, do wartości 1,0 w fazie dostrajania.

4.2. Wybór rozmiaru sieci oraz parametrów uczenia

Pierwszy etap testów obejmował analizę wartości błędu kwantyzacji (*qe*) oraz topografii (*te*) (wyjaśnione w rozdz. 3) w zależności od liczby epok w poszczególnych fazach uczenia. Wyniki testów zawarte są w tabelach 1 oraz 2.

Na podstawie wyników stwierdzono, że maksymalna efektywność sieci osiągnięta jest już w fazie porządkowania po około 20 epokach dla sieci 7×4 , około 200 epokach dla sieci 14×8 oraz 400 dla 29×15 (por. zaznaczone szarym kolorem pola w tab. 1 oraz 2). W fazie dostrajania współczynnik kwantyzacji nie ulega zmianie. Współczynnik topografii zmienia się w sposób nieregularny w czasie uczenia. Analiza wartości wag neuronów sieci oraz rozkładu odpowiedzi dla próbek korzystnych, niekorzystnych i pośrednich zbioru uczącego i testowego pozwala stwierdzić, że wystarczającym jest trenowanie sieci zgrubnie w fazie porządkowania.

Dalsze testy przeprowadzono na sieci o najmniejszym wymiarze tj. 7×4 , grupującej obszary obrazu w 28 klasach. Dokonano sprawdzenia dwóch innych funkcji współczynnika uczenia tj. funkcji liniowej i szeregu potęgowego, oraz algorytmu sekwencyjnego (omówione w rozdz. 3), ucząc sieć w fazie porządkowania w 25 epokach. Analizując rozkład trzech kategorii próbek na otrzymanej mapie we wszystkich przypadkach stwierdzono brak znaczącej poprawy w postaci wyodrębnienia neuronów reagujących jedynie na próbki obszarów korzystnych, niekorzystnych lub pośrednich.

Tabela 1
Błąd kwantyzacji i topografii w fazie porządkowania w zależności od liczby epok

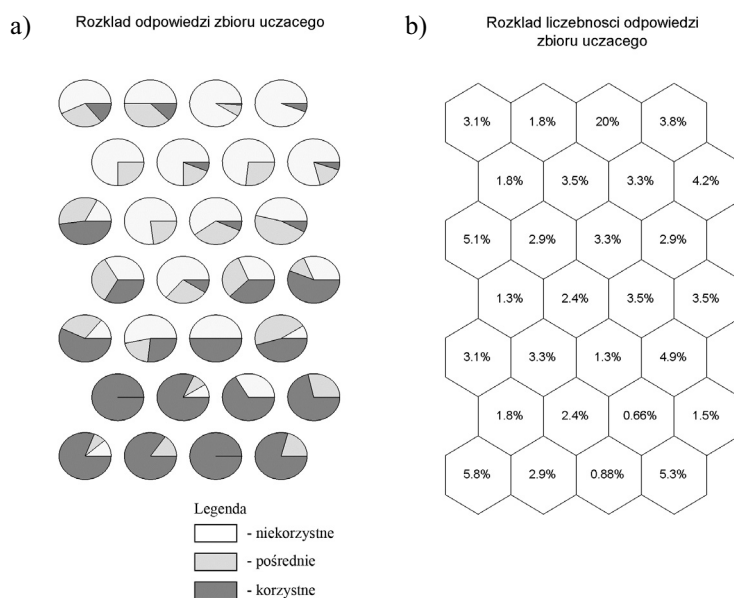
Rozmiar	Błąd	Liczba epok w fazie porządkowania								
		10	20	25	100	200	300	400	600	800
7×4	<i>qe</i>	0,267	0,267	0,267	–	–	–	0,267	0,267	–
	<i>te</i>	0,033	0,027	0,027	–	–	–	0,027	0,027	–
14×8	<i>qe</i>	0,211	0,210	0,210	0,207	0,207	–	0,206	0,207	0,206
	<i>te</i>	0,033	0,022	0,020	0,024	0,013	–	0,018	0,035	0,022
29×15	<i>qe</i>	0,158	0,155	0,154	–	0,152	0,152	0,152	0,152	0,153
	<i>te</i>	0,022	0,013	0,027	–	0,007	0,009	0,004	0,013	0,022

Tabela 2
Błąd kwantyzacji i topografii w fazie dostrajania w zależności od liczby epok

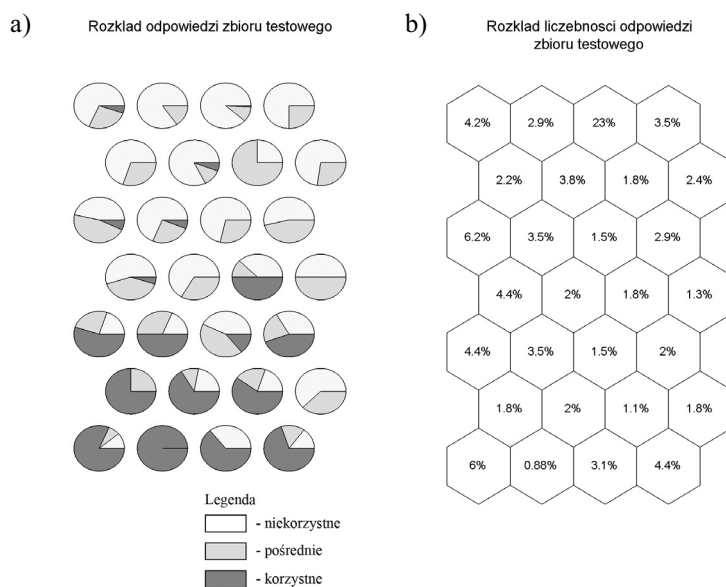
Rozmiar	Błąd	Liczba epok w fazie porządkowania \ dostrajania			
		25 \ 100	200 \ 800	400 \ 1600	400 \ 3000
7×4	<i>qe</i>	0,267	–	–	–
	<i>te</i>	0,027	–	–	–
14×8	<i>qe</i>	–	0,207	–	–
	<i>te</i>	–	0,013	–	–
29×15	<i>qe</i>	–	0,152	0,152	0,152
	<i>te</i>	–	0,007	0,004	0,004

W postaci grafu kołowego przedstawiono rozkład próbek zbioru uczącego (rys. 4a) i zbioru testowego (rys. 5a) dla sieci 7×4. Liczba próbek z poszczególnych klas wynosiła:

- w zbiorze uczącym – 142 korzystnych, 89 pośrednich oraz 221 niekorzystnych (ogółem 452 próbki);
- w zbiorze testowym – 99 korzystnych, 103 pośrednie i 250 niekorzystnych (ogółem 452 próbki).



Rys. 4. Prezentacja rozkładu próbek zbioru uczącego: a) w postaci grafu kołowego w trzech klasach: obszarów korzystnych, pośrednich i niekorzystnych; b) liczebność odpowiedzi zbioru uczącego



Rys. 5. Prezentacja rozkładu próbek zbioru testowego: a) w postaci grafu kołowego w trzech klasach: obszarów korzystnych, pośrednich i niekorzystnych; b) liczebność odpowiedzi zbioru testowego

4.3. Wybór neuronów odpowiadających na obszary korzystne

Na podstawie rozmieszczenia odpowiedzi neuronów dla próbek korzystnych zbioru uczącego oraz dwóch współczynników wyznaczanych dla każdego z neuronów (określających liczbę odpowiedzi próbek korzystnych do wszystkich odpowiedzi neuronu ($ws1$) oraz liczbę odpowiedzi korzystnych do całego zbioru próbek korzystnych ($ws2$)) wytypowano neurony o identyfikatorach 7, 13 oraz 28 jako rozpoznające obszary korzystne (por. tab. 3 i rys. 6). Wpływ podjętej decyzji sprawdzono na zbiorze testowym (tab. 3). Na podstawie analizy rozmieszczenia wykrytych podobrazów na oryginalnym zdjęciu, wytypowano dodatkowe neurony z wyższych poziomów mapy cech, tj. neurony o identyfikatorach 3, 11 oraz 27 (por. rys. 6). Neurony te rozpoznały jedynie dwie próbki klasy korzystnej zbioru testowego, większość próbek na które odpowiedziały to próbki klasyfikowane wcześniej jako niekorzystne (tab. 4). Powodem różnic między wynikami zbioru testowego a wynikami na zdjęciach lotniczych jest fakt, iż w niektórych przypadkach reprezentacje podobrazów z obszarami korzystnymi i niekorzystnymi są do siebie bardzo podobne. Należy również dodać, że podobrazy grupowane przez neurony 7, 13 oraz 28 to przede wszystkim tereny zabudowane, które na zdjęciach lotniczych niekoniecznie znajdują się w wymaganych sześciu rejonach Grubera, i które w przypadku dużego bloku zdjęć stanowią niewielki procent pokrycia terenu. Wprowadzenie więc dodatkowych neuronów, które wskazują obszary o mniej wyraźnych bądź mniej licznych obiektach, powinno zapewnić uniwersalność wyboru podobrazów.

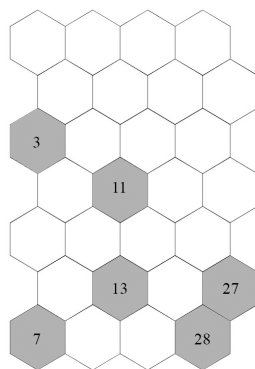
W przeprowadzonych testach została wyodrębniona klasa obszarów pośrednich zawierających niewielką ilością elementów liniowych. Biorąc pod uwagę przytoczony powyżej argument obszary te można uznać za korzystne. W dalszej części więc zarówno obszary korzystne jak i pośrednie traktowane będą jako korzystne.

Ostatecznie przyjęto, że neurony o identyfikatorach 7, 28, 13, 27, 3 oraz 11 odpowiadają na obszary korzystne pod względem wyszukiwania cech do orientacji wzajemnej zdjęć (rys. 6 oraz 4 i 5).

Tabela 3

Liczba odpowiedzi próbek korzystnych do wszystkich odpowiedzi neuronu ($ws1$) oraz liczba odpowiedzi korzystnych do całego zbioru próbek korzystnych ($ws2$)

Neuron	Zbiór uczący		Zbiór testowy	
	$ws1$ [%]	$ws2$ [%]	$ws1$ [%]	$ws2$ [%]
7	80,0	15,0	81,0	22,0
13	81,0	6,3	66,0	6,1
27	79,0	13,0	70,0	14,0



Rys. 6. Neurony odpowiadające na obszary korzystne pod względem wyszukiwania cech do orientacji wzajemnej zdjęć

4.4. Analiza otrzymanych rezultatów

Wśród wyselekcjonowanych przez wytypowane neurony (o identyfikatorach 3, 7, 11, 13, 27 oraz 28) podobrazów zbioru uczącego i testowego można wyróżnić odpowiedzi trzech typów:

- 1) *True Positive* (TP) – odpowiedź pozytywna, którą stanowią próbki klasy korzystnej i pośredniej rozpoznane przez sieć prawidłowo.
- 2) *False Positive* (FP) – odpowiedź negatywna dodatnio, którą stanowią próbki klasy korzystnej i pośredniej nie rozpoznane przez sieć.
- 3) *False Negative* (FN) – odpowiedź negatywna ujemnie, którą stanowią próbki klasy niekorzystnej rozpoznane przez sieć jako korzystne.

Na podstawie liczby odpowiedzi każdego typu obliczono dwa współczynniki opisujące jakość klasyfikacji podobrazów: współczynnik kompletności (kp) oraz poprawności (pr) opisane równaniami (11) oraz (12).

$$kp = \frac{TP}{(TP + FP)} \quad (11)$$

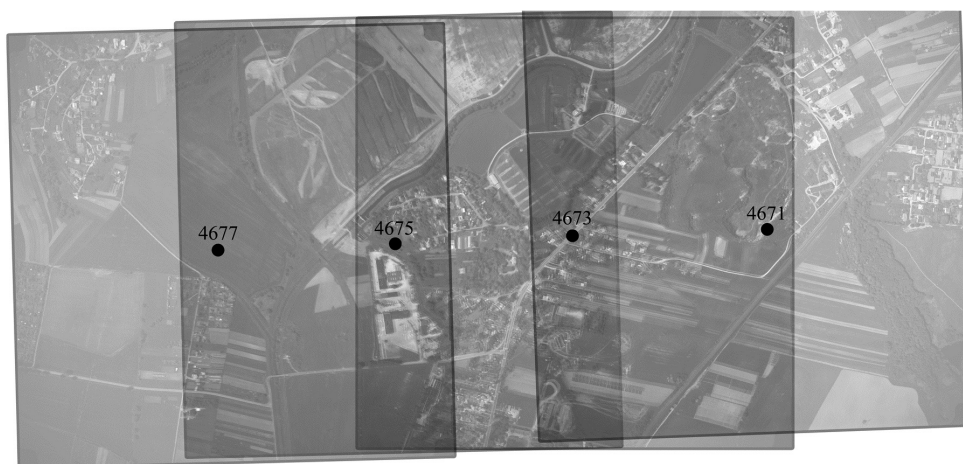
$$pr = \frac{TP}{(TP + FN)} \quad (12)$$

Wyniki zebrano w tabeli 4. Sieć w 68,3% poprawnie klasyfikuje próbki zbioru testowego jako korzystne (współczynnik pr). Próbki te stanowią 34,2% zbioru próbek klasy korzystnej i pośredniej (współczynnik kp). Z punktu widzenia fotogrametrii jest to dobry wynik. Przede wszystkim bowiem liczy się poprawność rozpoznania, a nie ilość. Zdecydowanie bardziej niepożądanym przypadkiem jest taki, gdy sieć zaklasyfikuje obszary niekorzystne jako korzystne, niż gdy obszary korzystne zostaną pominięte.

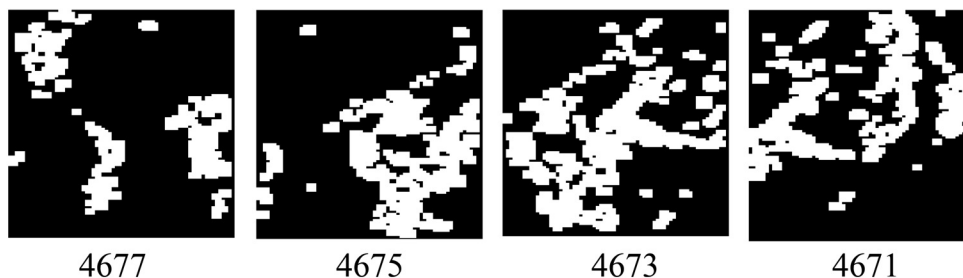
Tabela 4
Wyniki selekcji podobrazów zbioru uczącego i testowego neuronami 3, 7, 11, 13, 27 oraz 28

Neuron	Zbiór uczący			Zbiór testowy		
	<i>k</i>	<i>p</i>	<i>n</i>	<i>k</i>	<i>p</i>	<i>n</i>
3	11	8	6	2	13	13
7	21	2	3	22	2	3
11	1	3	7	0	3	6
13	9	3	1	6	1	2
27	5	2	0	0	3	5
28	19	5	0	14	3	3
TP	89		–	69		–
FP	142		–	133		–
FN	–		17	–		32
<i>kp</i>	38,5%			34,2%		
<i>pr</i>	84,0%			68,3%		

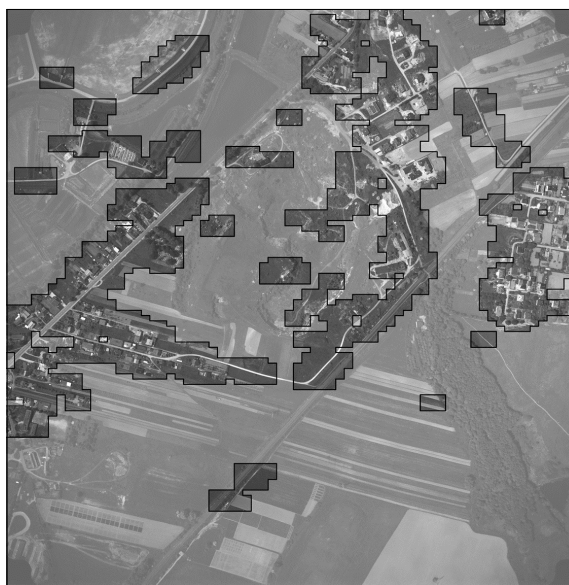
Skuteczność selekcji podobrazów sprawdzono wizualnie na czterech zdjęciach lotniczych stanowiących blok o pokryciu podłużnym 60% (rys. 7). Wyniki selekcji informacji nauczoną siecią Kohonena prezentują rysunki 8 i 9.



Rys. 7. Blok zdjęć lotniczych



Rys. 8. Zdjęcia lotnicze poddane selekcji informacji siecią Kohonena (białym kolorem oznaczono obszary korzystne wskazywane na kolejnych zdjęciach bloku)



Rys. 9. Zdjęcie lotnicze nr 4671 – wizualizacja obszarów korzystnych (oznaczonych na rysunku ciemniejszym odcieniem)

5. Wnioski

Zaproponowana reprezentacja obrazu w postaci wektora wartości maksymalnych dla osi kierunku dwuwymiarowego histogramu gradientu może być wykorzystana do selekcji podobrazów na zdjęciach. Opisane nią trzy klasy obszarów grupują się na mapie cech (rys. 4). Możliwe jest więc wskazanie neuronów odpowiadających na obszary korzystne pod względem wyszukiwania cech do orientacji wzajemnej. Uzyskane wyniki utrzymują się na poziomie 68,3% poprawnie klasyfikowanych próbek zbioru testowego.

Wyniki selekcji podobrazów siecią neuronową Kohonena można uznać za dobre. Na zdjęciach binarnych (rys. 8), można wskazać powtarzające się kształty złożone z wybranych obszarów. Świadczy to, że sieć Kohonena wybiera na kolejnych zdjęciach lotniczych podobne podobrazy, na których w dalszych etapach można poszukiwać odpowiedniości.

Dalsze badania ukierunkowane będą na poprawienie efektywności rozpoznania obszarów korzystnych (wyizolowania ich na mapie cech). Jedną z możliwości jest dodanie kolejnej warstwy sieci Kohonena do wytypowanych już neuronów (lub do innych) w celu bardziej precyzyjnego rozróżnienia odmiennych typów próbek. Inną jest zastosowanie uczenia metodą wstecznej propagacji błędów (*backpropagation*) [7]. Generowane na mapie Kohonena odpowiedzi mogą być sygnałami wejściowymi dla warstwy *backpropagation*, która będzie dokonywać ostatecznej klasyfikacji, ucząc się metodą nadzorowaną.

Literatura

- [1] Heipke C.: *Automation of interior, relative and absolute orientation*. ISPRS Journal of Photogrammetry & Remote Sensing, 52, 1–19
- [2] Kohonen T.: *Self-Organizing Maps*. New York, Springer-Verlag 1997
- [3] Parker J.R.: *Algorithms for Image Processing and Computer Vision*. Wiley 1996
- [4] *Image Processing Toolbox 5*. The MathWorks, Inc., 1993–2007
- [5] *Słownik Geoinformatyczny PAU*. sloownik.fotogrametria.agh.edu.pl
- [6] *SOM Toolbox Documentation*. <http://www.cis.hut.fi/projects/somtoolbox/documentation> www.cis.hut.fi/projects/somtoolbox/documentation
- [7] Tadeusiewicz R.: *Sieci neuronowe*. Warszawa, Akademicka Oficyna Wydawnicza RM 1993