

WYKORZYSTANIE SIECI NEURONOWYCH W PROCESACH FOTOGRAMETRYCZNYCH

APPLICATION OF NEURAL NETWORKS TO PHOTOGRAMMETRIC PROCESSES

Sławomir Mikrut¹, Zbigniew Mikrut²

¹Katedra Geoinformacji, Fotogrametrii i Teledetekcji Środowiska, AGH w Krakowie

²Katedra Automatyki, AGH w Krakowie

SŁOWA KLUCZOWE: sztuczne sieci neuronowe, spasowanie obrazów, fotogrametria cyfrowa

STRESZCZENIE: W niniejszym artykule poruszono problem wykorzystania sztucznych sieci neuronowych (SSN) w geoinformatyce obrazowej, ze szczególnym uwzględnieniem procesów fotogrametrycznych. Przedstawiono wyniki przeglądu literatury światowej oraz zaprezentowano rezultaty badań prowadzonych w ramach projektu dotyczącego użycia sieci do spasowania fotogrametrycznych zdjęć lotniczych. W oparciu o literaturę, przeanalizowano wyniki prac wykorzystujących sieci neuronowe do: klasyfikacji obrazów wielospektralnych, wydobywania cech, kalibracji kamer oraz spasowania obrazów. Zaprezentowano również wyniki własnych eksperymentów, bazujących na idei wykorzystania sieci opierającej się na wyborze specjalnej reprezentacji, która następnie jest wykorzystywana do spasowania obrazów fotogrametrycznych dla dwóch wybranych typów terenu. W badaniach wykorzystano sieci impulsujące ICM (*Intersecting Cortical Model*), będące jedną z wersji sieci PCNN (*Pulse Coupled Neural Network*), przy pomocy których wygenerowano tzw. podpisy obrazów (*signatures*), czyli kilkudziesięcioelementowe wektory, opisujące strukturę fragmentu obrazu. Wyniki badań częściowo potwierdzają słuszność przyjętych założeń, mimo występujących problemów związanych ze specyfiką obrazów fotogrametrycznych.

1. WSTĘP

Rozwój sztucznych sieci neuronowych (SSN) został zapoczątkowany w latach czterdziestych badaniami nad matematycznym opisem komórki nerwowej oraz powiązaniem go z przetwarzaniem informacji przez komórkę. Za początek przyjmuje się rok 1943 związany z opracowaniem matematycznego modelu sztucznego neuronu przez McCullocha i Pittsa (Korbicz *et al.*, 1994). Jednak intensywne badania nad SSN przypadły na lata osiemdziesiąte z racji rozpowszechniania się komputerów i ułatwienia procesów obliczeniowych, towarzyszących sieciom. Istotnym osiągnięciem było opracowanie przez Hopfielda w 1982 r. sieci ze sprzężeniem zwrotnym, którą wykorzystano do rozwiązywania zadań optymalizacyjnych (np. problem komiwojażera) oraz odtwarzania obrazów z ich fragmentów.

SSN nie ominęły również zastosowań w geoinformatyce obrazowej. Koniec lat osiemdziesiątych i początek dziewięćdziesiątych to pierwsze zastosowania sieci neuronowych w tej dziedzinie nauk. W niniejszym opracowaniu autorzy skupili się na

przykładach wykorzystania sieci w geoinformatyce obrazowej (rozdział 2), ze szczególnym uwzględnieniem zastosowań fotogrametrycznych. W rozdziałach 3 i 4 zaprezentowano dokonania autorów w temacie pasowania obrazów zdjęć lotniczych przy pomocy sygnatur.

2. SIECI NEURONOWE W FOTOGRAMETRII – HISTORIA I DZIEŃ DZISIEJSZY

Zarys technik sieci neuronowych można znaleźć w literaturze (Tadeusiewicz, 1993; Tadeusiewicz *et al.*, 2007). Autorzy niniejszego artykułu również dokonali pewnych analiz w poprzednich publikacjach (Mikrut *et al.*, 2007; Mikrut, 2007).

Przykłady wykorzystania sztucznych sieci neuronowych w zadaniach związanych z geoinformatyką obrazową mają już prawie dwudziestoletnią historię. Pierwszymi zastosowaniami SSN było wykorzystanie ich do klasyfikacji obrazów wielospektralnych. Pierwsze wzmianki o próbach klasyfikowania obrazów pochodzą z początku lat 90-tych, kiedy to zaczęto klasyfikować zobrażenia satelity LANDSAT TM (Bishop *et al.*, 1992), następnie SPOT (Dreyer, 1993). Sieci wykorzystano również do opracowania danych radarowych (Hara, 1994). Badania te potwierdziły możliwości skutecznego działania SSN jako jednej z metod klasyfikacji obrazów. Kolejne zastosowania to wydobywanie cech (ang. *feature extraction*), kalibracja kamer, spasowanie obrazów (ang. *matching*).

2.1. Pojawienie się sieci neuronowych w tematyce kongresów ISPRS

Analizując historię kongresów Międzynarodowego Towarzystwa Fotogrametrii i Teledetekcji (ISPRS) można zauważyć wzrost liczby publikacji obejmujących tę tematykę. Ilość artykułów, zawierających słowa „*neural network*” w tytule, wzrastała z kongresu na kongres. Tematyka ta przedstawiona była już w ośmiu artykułach w 1996 w Wiedniu, tyle samo było prezentowanych w Amsterdamie (2000 r.), jedenaście w Istambule (2004 r.) oraz osiemnaście w Pekinie (2008 r.). Wiele było również artykułów dotyczących problematyki wykorzystania sieci neuronowych, nie zawartych jednak w tytule. Jak wspomniano wcześniej, początki to zastosowanie SSN do klasyfikacji obrazów wielospektralnych. Miało to odzwierciedlenie również w publikacjach kongresowych. Przykładowe opracowania można znaleźć w literaturze (Barsi, 1996; Bock, 1996; Chen *et al.*, 1996; Barsi, 2000; Kamiya, 2000; Aria *et al.*, 2004; Cetin *et al.*, 2004; Zhou *et al.*, 2008).

2.2. Sztuczne sieci neuronowe w zagadnieniach fotogrametrycznych

SSN zaczęto również stosować do rozwiązywania typowych problemów fotogrametrycznych. Poniżej dokonano próby tematycznego ich usystematyzowania. Badania literaturowe podkreśliły potrzebę dokonania takiej analizy, z racji coraz powszechniejszego ich wykorzystania, a także skuteczności działania.

2.2.1. Wykorzystanie SSN do identyfikacji wybranych obiektów

Wykorzystanie sieci typu FFN (*feed-forward network*) zastosowano do operacji na ogół wykonywanej manualnie przez operatora stacji fotogrametrycznej tj. do identyfikacji punktów sygnalizowanych (Kepuska *et al.*, 1991). W tym celu zastosowano sieć typu FFN-BP (*back-propagation*) do rozpoznawania obiektów na zeskanowanych zdjęciach lotniczych. Dane przygotowane zostały z postaci obrazów o różnej orientacji, tle, skali i z wprowadzonymi pewnymi zniekształceniami geometrycznymi. Aby wyeliminować nadmiarowość i zminimalizować rozmiar danych testowych przygotowano odpowiedni schemat treningowy. Po dwóch iteracjach treningu uzyskano zadawalające rezultaty: model FFN-BP uzyskał rozpoznanie na poziomie 85% punktów, znajdujących się na 34 obrazach. Zważywszy na rok przeprowadzonych testów (lata 90-te, początki stosowania sieci), należy podkreślić osiągnięcie stosunkowo dobrych rezultatów.

2.2.2. SSN a wydobywanie cech z obrazu (*feature extraction*)

SSN zaczęto również wykorzystywać do wydobywania cech (*feature extraction*). W badaniach dokonano porównania metod wydobywania cech pomiędzy transformacją Karhunen-Löwe (K-L) a SOM (*Self Organizing Maps*) Kohonena (Törma, 96). Uzyskane w trakcie badań różnice pomiędzy metodami były bardzo małe. W niektórych przypadkach błędy metody K-L były mniejsze, jednak proces działający przy użyciu SOM był szybszy. Różnice pomiędzy metodami powstawały wówczas, gdy liczba próbek na klasę była mała, co jest typowe dla sieci, tzn. przy zbyt małej liczbie próbek sieci wykazują małą skuteczność.

Również przy pomocy SSN opracowano ekstrakcję węzłów drogowych z obrazów cyfrowych (Barsi *et al.*, 2002). Neuronowy operator, oparty o sieć typu *feed-forward*, przeszukuje oknem czarno-białe obrazy w średniej rozdzielczości decydując czy skrzyżowanie zawiera 3 czy 4 odnogi. Najlepsze rezultaty otrzymano dla sieci trójwarstwowej. Do uczenia wykorzystano warstwę wektorową dróg.

Kilka innych przykładów użycia SSN do wydobywania cech z obrazu zaprezentowano na kongresie ISPRS w Pekinie (Ahadzadeh *et al.*, 2008; Mokhtarzade *et al.*, 2008).

2.2.3. Kalibracja kamer przy pomocy SSN

Procedurę kalibracji kamer przy wykorzystaniu sieci *back-propagation* (BP) opisano w publikacji (Mendonca *et al.*, 2002). Dzięki jej zastosowaniu nie musimy posiadać danych takich jak stała kamery czy dystorsja. Sieci są skuteczne w rozwiązywaniu nieliniowych zagadnień i mogą „nauczyć” się odpowiednich relacji pomiędzy współrzędnymi obrazowymi a współrzędnymi przedmiotowymi (obiektu), bez potrzeby określania konkretnego modelu matematycznego. W opisaney przez autora metodyce dokonano porównania wyznaczonych współrzędnych pola testowego (520 punktów) metodą najmniejszych kwadratów oraz SSN. Dokładności uzyskane przy pomocy obu metod były porównywalne. Według autora zaletami metody neuronowej są: 1) niekonieczna znajomość modelu matematycznego, 2) nie wymagana jest również znajomość przybliżonych (wejściowych) parametrów kalibracji, 3) metodyka może być wykorzystana dla różnego typu kamer, 4) może być użyta dla kilku kamer o różnych

elementach orientacji, 5) SSN mogą rozpoznać specyficzne ułożenia kamer, co ma znaczenie przy dynamicznych systemach pracujących w czasie rzeczywistym.

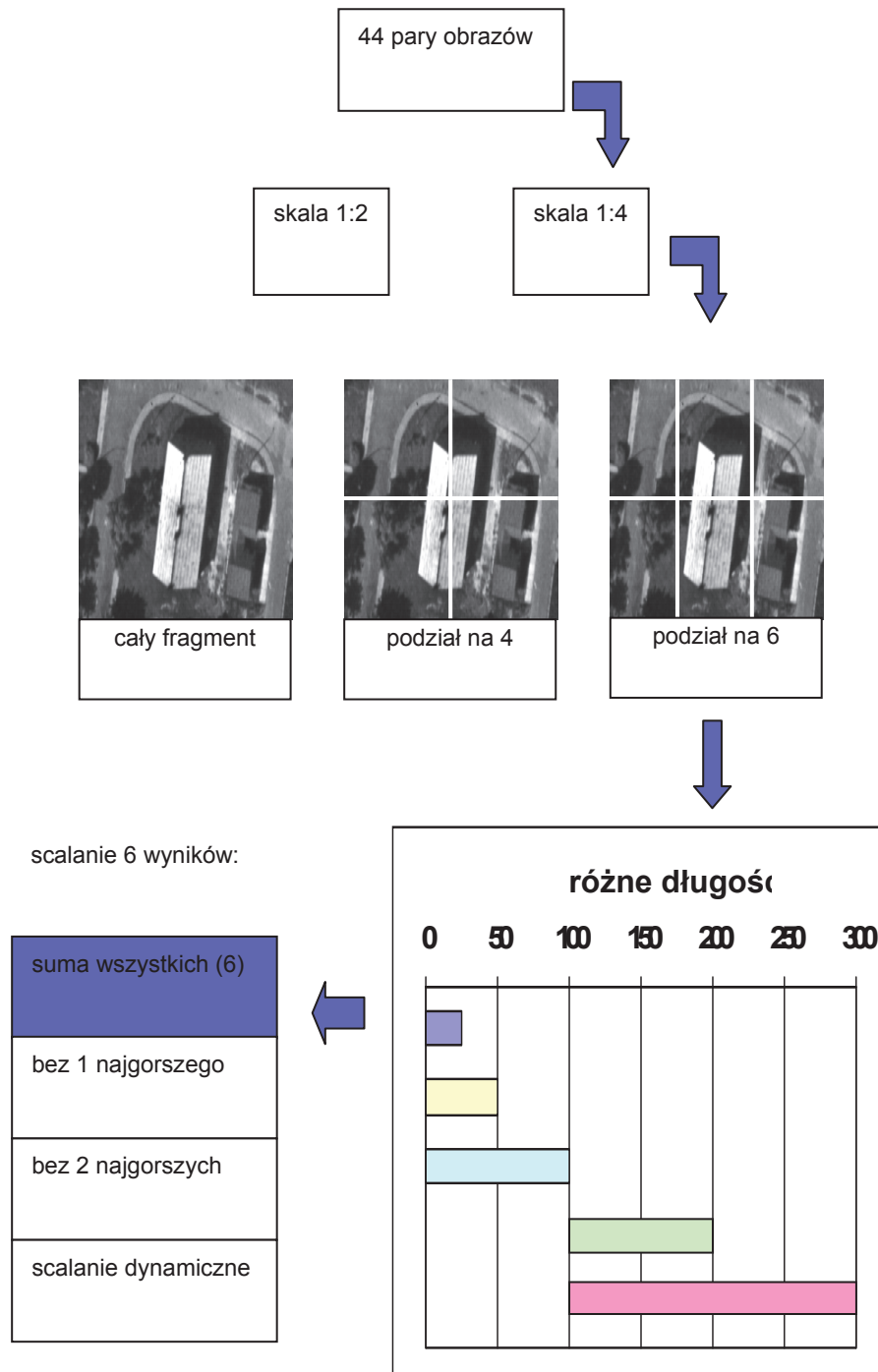
2.2.4. Spasowanie obrazów (*matching*)

SSN znalazły zastosowanie w spasowaniu obrazów w systemach stereowizyjnych (Wang *et al.*, 1999). Wykorzystano je do określania pozycji dowolnego punktu w przestrzeni 3D w oparciu o zdjęcia naziemne wykonane z dwóch kamer. Użyto dwóch typów sieci: VC (*vitality conservation*) oraz BP (*back-propagation*). Pierwsza sieć, jako wektora uczącego, używa wybranych parametrów (jasność, zmienność, orientację, x , y) dla każdego piksela oraz dzieli obraz wejściowy na kilka klastrów, a rezultaty są wykorzystane przez drugą sieć (BP), celem osiągnięcia dokładnego dopasowania. Po treningu sieć BP jest wykorzystana do generowania tzw. mapy dysparycji. W procesie dopasowania, aby zredukować ewentualne błędy oraz ograniczenia obszaru przeszukiwania, korzysta się z dodatkowych informacji takich jak linie epipolarne, uporządkowanie, geometria i ciągłość.

Przykładem podobnego zastosowania SSN, ale z innym rozwiązaniem (model sieci) jest stereowizyjny system zbudowany przez zespół laboratorium węgiersko-japońskiego (Resko *et al.*, 2003). Stereowizja jest realizowana poprzez biologicznie zainspirowaną, pięciowarstwową, sztuczną sieć neuronową, o prostych komponentach. Zadaniem sieci jest porównywanie symbolicznych cech (*symbolic feature*) w obrębie obydwu obrazów, celem wyznaczenia kierunków obu kamer (określenia ich orientacji). System funkcjonujący w oparciu o ten model działa w czasie rzeczywistym.

Wykorzystanie sieci neuronowych oraz zależności fotogrametrycznych (komplanarność, współczynnik korelacji) do spasowania wieloobrazowego (*multi-image matching*) zostało zaprezentowane podczas kongresu ISPRS w Pekinie w roku 2008 (Elaksher, 2008). Autor zaprezentował metodykę użycia sieci w odniesieniu do zdjęć naziemnych przedstawiających ratusz w Zurichu, udostępnianych przez ISPRS jako materiał testowy. Wykorzystując dowolną liczbę zdjęć, proces startuje od orientacji wzajemnej na podstawie warunku komplanarności pomiędzy wszystkim parami zdjęć. Dla każdej pary otoczeń punktów liczony jest współczynnik korelacji dla lokalnych jasności obrazu. Do określania prawidłowego dopasowania użyto sieci typu *feed-forward*. Uzyskano bardzo zadowalające rezultaty. Współczynnik detekcji (ang. *detection rate*) na poziomie 95% - 98% oraz współczynnik niedopasowania (ang. *false alarm rate*) od 4% do 7%.

Przedstawione pozycje literaturowe odnoszące się do spasowania obrazów obejmują badania prowadzone na obrazach zdjęć naziemnych. Jest to zagadnienie o tyle nieskomplikowane, że obrazy z kamer naziemnych charakteryzują stosunkowo dużym podobieństwem (z reguły robione są z krótkiej bazy). Problem staje się bardziej złożony w sytuacji, gdy mamy do czynienia ze zdjęciami lotniczymi. Rozdrobnienie informacji oraz duża paralaksa na przykład w przypadku wysokich budynków znacznie komplikuje zagadnienie. Autorzy niniejszego artykułu dokonali próby rozwinięcia zagadnienia w oparciu o własne doświadczenia prezentowane w publikacji (Mikrut *et al.*, 2007).



Rys. 1. Dopasowywanie fragmentów obrazów przy użyciu sygnatur.
Kolorem niebieskim oznaczono ciąg przekształceń prowadzący do najlepszego wyniku

3. EKSPERYMENTY Z WYKORZYSTANIEM SSN W ODNIESIENIU DO SPASOWANIA ZDJĘĆ LOTNICZYCH

Koncepcję badań nad wykorzystaniem SSN do spasowania zdjęć lotniczych można znaleźć w publikacji (Mikrut *et al.*, 2007). Autorzy kontynuowali badania nad wykorzystaniem sygnatur generowanych przez impulsującą sieć neuronową (Mikrut, 2007). Korzystano z bazy danych opisanej we wspomnianej publikacji. Z bazy wybrano podzbiór 44 obrazów w stopniach szarości (składowe R, G i B zostały scalone). Na kolejnym zdjęciu lotniczym manualnie zlokalizowano te obszary i po odpowiednim obramowaniu zapisano je w celu przeprowadzenia testów spasowania. Należy podkreślić, że w omawianych badaniach sprawdzano możliwości dopasowywania reprezentacji podobrazów, czyli obliczano *sygnatury* podobrazu ze zdjęcia chronologicznie pierwszego, następnie obliczano sygnatury wszystkich podobrazów wchodzących w skład fragmentu drugiego (kolejnego) zdjęcia lotniczego i te wektory porównywano. Samo porównywanie wektorów realizowano przy użyciu sumy modułów różnic poszczególnych elementów (patrz Dodatek). Schemat eksperymentów przedstawiono na rysunku 1.

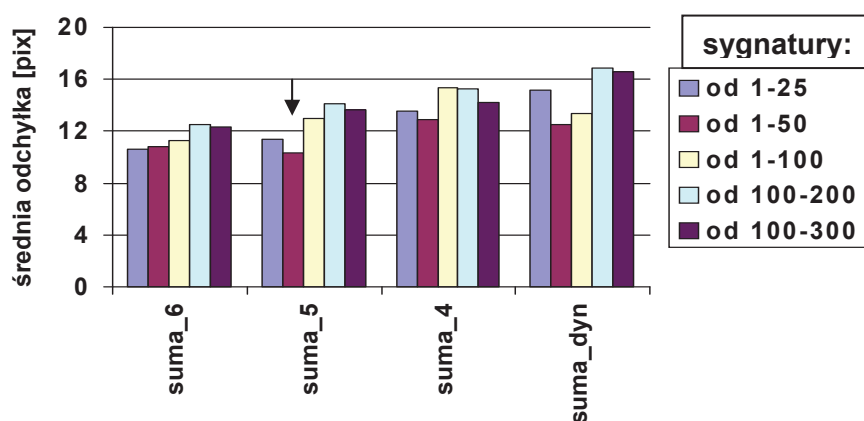
Pary obrazów, na które składały się: fragment obrazu o rozmiarach 240x160 pikseli i odpowiadający mu obszar z kolejnego zdjęcia, obramowany dookoła pasem o szerokości 150 pikseli, zmniejszono dwu i czterokrotnie. Następnie, przy pomocy sieci impulsującej ICM, generowano 300. elementowe sygnatury dla całego fragmentu obrazu oraz dla jego czterech lub sześciu podobrazów, dla każdego podobrazu osobno. Po drugim zdjęciu przesuwano okno odpowiadające rozmiarom fragmentu poprzedniego obrazu. Po każdym jednopikselowym przesunięciu obliczano sygnatury w wariantach podziału na podobrazy, opisanych wyżej. Cały ten proces schematycznie przedstawiono na rysunku 1 w górnych trzech rzędach. Formalny opis algorytmu zamieszczono w Dodatku. W celu zachowania przejrzystości rysunku nie umieszczono na nim strzałek pokazujących zrealizowane kombinacje przetwarzania. Strzałki powinny obrazować połączenia „każdy z każdym”, idące od bloków górnych do dolnych.

Wstępna analiza rezultatów dopasowań pozwoliła na wyznaczenie sekwencji przekształceń, prowadzących do najlepszych wyników. Kolejność ta została zaznaczona na rysunku 1 strzałkami blokowymi. Przekształcenia prowadzące do najlepszych wyników, to operowanie na obrazach czterokrotnie zmniejszonych, dla których dopasowywano fragmenty powstałe z podziału obrazu na 6 części. Na tej podstawie zrealizowano kolejny etap badań, przedstawiony na rysunku 1 w dolnym (ostatnim) rzędzie.

Fakt posiadania 300 elementowych sygnatur wykorzystano do porównywania wektorów sygnatur o różnych długościach i różnych położeniach. Jak pokazano na rysunku 1, do porównań używano sygnatur 25, 50 i 100 elementowych pobieranych od umownego momentu startu oraz fragmentów sygnatur od pozycji 100 do 200 oraz od pozycji 100 do 300. Podział dopasowywanego fragmentu obrazu na 4 lub 6 podobrazów był zainspirowany dwoma czynnikami. Pierwszym z nich był fakt uzyskiwania lepszych wyników dopasowywania sygnatur dla mniejszych wycinków obrazu. Drugim czynnikiem była potrzeba ewentualnego uwzględnienia występowania na obrazie obiektów, które na obrazie „późniejszym” nieco zmieniłyby położenie względem poziomu gruntu, a być może także kształt. Przykład występowania takich obiektów pokazano na rysunku 1. W przypadku gorszego dopasowania niektórych podobrazów istnieje możliwość ich pominięcia podczas procesu scalania cząstkowych

wyników dopasowywania. Podczas analizy wyników uwzględniono 4 możliwości scalania, pokazane na rysunku 1. Były to: sumowanie wszystkich sześciu wyników cząstkowych, sumowanie pięciu i czterech najlepszych oraz dynamiczne odrzucanie gorszych wyników. Ta ostatnia metoda polegała na badaniu różnic pomiędzy uporządkowanymi rosnąco wynikami i odrzucaniu tych, które przekroczyły zadany próg. W każdym z wymienionych przypadków suma wyników cząstkowych była normalizowana: dzielono ją przez liczbę wziętych pod uwagę składników (por. wzory D.3 i D.4). Na rysunku 2 porównano średnie odchyłki pasowania (por. wzór D.6), obliczone dla wyżej zdefiniowanych sygnatur i czterech metod scalania wyników cząstkowych.

Dla metody sumowania wszystkich 6 wyników cząstkowych (na rysunku 2 oznaczenie *suma_6*) średnie odchyłki są najbardziej skupione i niskie. Minimum wystąpiło dla metody z odrzuceniem jednej wartości (*suma_5*) dla sygnatury 50 elementowej (por. czarna strzałka na rysunku 2). Dla sygnatury o tej długości średnie odchyłki w następnym procesie scalania (*suma_4* i *suma_dyn*) także osiągają minimalne wartości. W związku z wysokimi średnimi odchyłkami obserwowanymi dla sygnatur 100 i 200 elementowych porównywanych od setnego elementu w dalszej analizie sygnatury te zostaną pominięte. Oprócz wartości średnich, które charakteryzują jakość dopasowania w sposób ogólny, przyjęto kilka progów odchyłek dopasowań (δ - por. wzór D.5) i obliczono procentowo, ile odchyłek lokuje się poniżej. Rezultaty uzupełniono – dla porównania – wartościami średnimi i zestawiono w tabeli 1. W każdej kolumnie, odpowiadającej różnym progom odchyłek, zaznaczono wartość maksymalną. Większość maksimum jest zlokalizowana w wierszach odpowiadających sygnaturom 1-25 i 1-50, przy czym w obu przypadkach wyniki końcowe dopasowania obliczono sumując wszystkie 6 wyników cząstkowych (bez odrzuceń – por. ostatnią kolumnę tabeli 1). Ta metoda prezentacji wyników ma za zadanie zorientowanie się w precyzji pasowania, uzyskiwanej przy użyciu różnych sygnatur. Wartości uzyskane dla sygnatur krótkich (od 1-25 i od 1-50) metodą sumowania wszystkich 6 elementów są do siebie zbliżone.

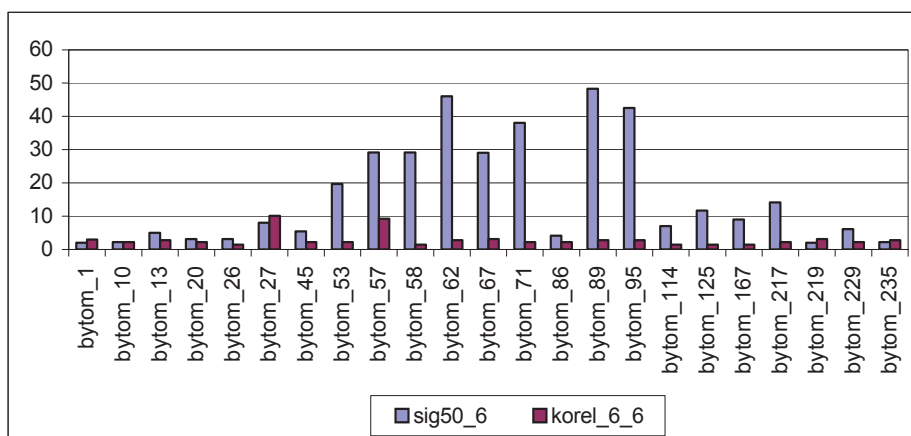


Rys. 2. Porównanie średnich błędów pasowania dla kilku długości sygnatur i czterech metod scalania wyników. Strzałką zaznaczono wartość minimalną

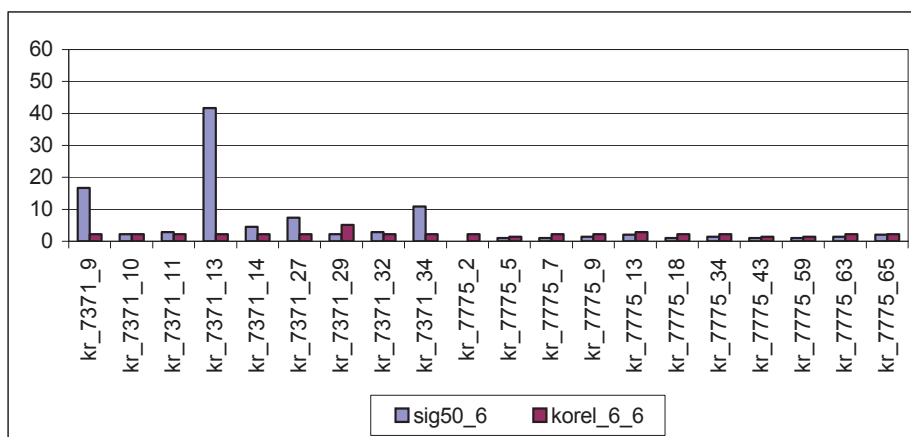
Tabela 1. Procentowe dopasowania dla kilku dopuszczalnych odchyłek δ

sygnatura	średnia odchyłka [piksele]	$\delta < 3$	$\delta < 5$	$\delta < 10$	$\delta < 30$	liczba odrzuceń
od 1 – 25	10,59	47,73%	54,55%	70,45%	86,36%	0
od 1 – 50	10,77	45,45%	56,82%	70,45%	88,64%	0
od 1 – 100	11,28	34,09%	54,55%	65,91%	88,64%	0
od 1 – 25	11,38	34,09%	47,73%	65,91%	88,64%	1
od 1 – 50	10,31	34,09%	43,18%	68,18%	90,91%	1
od 1 – 100	12,99	31,82%	45,45%	59,09%	84,09%	1

Dla sygnatury o najniższej średniej odchyłce (od 1-50, sumowanie 5 elementów – patrz przedostatni wiersz tabeli 1) precyzja dopasowania dla małych odchyłek jest o wiele niższa. Spasowanie przy pomocy tej metody wykonano również dla dwóch różnych typów obrazów „Bytom” i „Krakow”. Wyniki zaprezentowano na rysunkach 3 i 4, na których porównano odchyłki pasowania uzyskane przy pomocy sygnatur (sig50_6) z metodą klasyczną (kor_6_6). Jak widać metoda sygnatur jest bardzo czuła na rodzaj obrazu wejściowego. W przypadku „Bytomia” mamy do czynienia ze zdjęciami z centrum miasta, a więc o dużej liczbie bloków mieszkalnych, z występującymi tu znacznymi przesunięciami radialnym pomiędzy obrazami. Obrazy z „Krakowa” charakteryzują się większym podobieństwem z racji lokalizacji poza miastem, o nieodróżnianej rzeźbie terenu. Zauważalne na rysunku 3 znaczne różnice pomiędzy odchyłkami wynikają głównie z faktu występowania wysokich bloków na obrazie. Klasyczna korelacja stosunkowo dobrze sobie „radzi” z tym problemem. Sieci, z racji swojej specyfiki działania są wrażliwe zarówno na przesunięcia radialne jak i na zmiany oświetlenia, stąd dokładność spasowania spada (obrazy Bytom_53 do Bytom_95). Podobnie jest w przypadku obrazów kr_7371_9 i kr_7371_13.



Rys. 3. Wyniki spasowania dla obrazu „Bytom”



Rys. 4. Wyniki spasowania dla obrazu „Kraków”

4. PODSUMOWANIE I WNIOSKI

W artykule dokonano przeglądu literatury dotyczącej wykorzystania sztucznych sieci neuronowych (SSN) w geoinformatyce obrazowej, ze szczególnym uwzględnieniem procesów fotogrametrycznych. Autorzy dokonali podsumowania zastosowań SSN w fotogrametrii, udowadniając tym samym coraz powszechniejsze ich wykorzystanie, poczynając od wydobywania cech z obrazu, przez kalibrację kamer, a na spasowaniu obrazów kończąc. Te ostatnie obejmowały głównie prace związane ze stereowizją. Brak w literaturze eksperymentów związanych z pasowaniem zdjęć lotniczych, z racji złożoności zagadnienia, doprowadził do podjęcia przez autorów prób badawczych w tym kierunku. W punkcie 3 artykułu przedstawiono kolejny etap badań - serie eksperymentów związanych z pasowaniem reprezentacji fragmentu obrazu w postaci sygnatur, generowanych przez impulsującą sieć neuronową ICM. Wyniki nie są w pełni zadowalające, aczkolwiek w blisko 1/3 przypadków efekty uzyskane przy pomocy sygnatur są takie same lub lepsze od rezultatów, otrzymanych po zastosowaniu klasycznej korelacji (por. rys. 3 i 4). Autorzy widzą sens kontynuowania badań w dwóch kierunkach:

1. pogłębionej analizy sytuacji (podobrazów) powodujących duże różnice odchyłek dopasowania obu porównywanych metod i wprowadzenia ewentualnych korekt na etapie wstępnego przetwarzania obrazu,
2. zastosowania sieci impulsujących do wykrywania obiektów, które następnie posłużą do generowania sygnatur i pasowania.

Pracę zrealizowano w ramach badań własnych AGH nr 10.10.120.783 oraz badań statutowych AGH nr 11.11.150.949/08.

5. LITERATURA

Ahadzadeh S., Valadanouzj M., Sadeghian S., Ahmadi S., 2008. Detection of Damaged Buildings after an Earthquake Using Artificial Neural Network Algorithm. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Beijing, China, Vol. XXI, Part B8, s. 369-372.

Aria E.H., Saradjian M.R., Amini J., Lucas C., 2004. Generalized cooccurrence matrix to classify IRS-1D images using neural network. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Istanbul, Turkey, Vol. XX, Part B7., s. 114-116.

Barsi A., 1996. Thematic classification of Landsat image using neural network. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vienna, Austria, Vol. XVIII, Part B3, s.48-52.

Barsi A., 2000. The impact of data compression and neighborhood information on the classification accuracy of artificial neural networks. *Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Amsterdam, Holland, Vol. XIX, Part B7, s. 140-147.

Barsi A., Heipke C., Willrich F., 2002. Junction extraction by artificial neural network system – JEANS, *IntArchPhRS XXXIV/3b*, Graz , s.18-21

Bischof H., Schneider W., Pinz A., 1992. Multispectral classification of Landsat images using neural network, *IEEE Trans. on Geoscience and Remote Sensing*, Vol. 30, No 3, s. 482-490.

Bock S., 1996. A region-based approach to land-use classification of remotely –sensed image data using artificial neural network. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vienna, Austria, Vol. XVIII, Part B7, s.71-76.

Cetin M., Kavzoglu T., Musaoglu N., 2004. Classification of multi-spectral, multi-temporal and multi-sensor images using principal components analysis and artificial neural networks: beykoz case. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Istanbul, Turkey, Vol. XX, Part B4, s. 951-956.

Chen C., Chen S., Shyn S., 1996. Classification of remote sensing imagery using an unsupervised neural network. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vienna, Austria, Vol. XVIII, Part B7, s. 112-117.

Dreyer P., 1993. Classification of land cover using optimized neural nets on SPOT data, *Photogrammetric Engineering and Remote Sensing*, Vol. 59, No. 5, s. 617-621.

Elaksher A., 2008. Multi-image matching using neural networks and photogrammetric conditions. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Beijing, China, Vol. XXI, Part B3, s. 39-44.

Hara Y., Atkins R., Yueh S., Shin R., Kong J., 1994. Application of neural networks to radar image classification, *IEEE Trans. on Geoscience and Remote Sensing*, Vol. 32, No. 1, s. 100-109.

- Kamiya I., 2000. Image classification by spatial shift invariant neural network. *Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Amsterdam, Holland, Vol. XIX, Part B7, s. 636-639.
- Kepuska V., Mason, S., 1991. Automatic signalized point recognition with feed-forward neural network, *IEE Conference Publication, 2-nd International Conference on Artificial Neural Networks*, London, Nr 349, s. 359-363.
- Korbicz J., Obuchowicz A., Uciński D., 1994. *Sztuczne sieci neuronowe. Podstawy i zastosowania*. Akademicka Oficyna Wydawnicza PLJ, Warszawa.
- Lindblad T., Kinser J.M., 2005: *Image Processing Using Pulse-Coupled Neural Networks*. Springer-Verlag, Berlin, Heidelberg
- Mendonca M., da Silva I., Castanho J., 2002. Camera calibration using neural networks. *Journal of WSCG - poster abstract at 10-th Int. Conf. Central Europe on Computer Graphics, Visualization and ComputerVision (WSCG 2002)*, Vol. 10, nr 1-3.
- Mikrut S., Mikrut Z., 2007. Sieci neuronowe w procesach automatycznej korelacji zdjęć lotniczych. *Archiwum Fotogrametrii, Kartografii i Teledetekcji*, Vol 17b, s. 505-515.
- Mikrut Z., 2007. Tworzenie reprezentacji obszarów zdjęć lotniczych za pomocą sieci neuronowych klasy PCNN. *Uczelniane Wydawnictwo Naukowo-Dydaktyczne Kraków – Półrocznik Automatyka*, Tom 11, s. 355 – 364.
- Mokhtarzade M., Valadan Zoej M., Ebadi H., 2008. Automatic Road Extraction from High Resolution Satellite Images Using Neural Networks, Texture Analysis, Fuzzy Clustering and Genetic Algorithms. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Beijing, China, Vol. XXI, Part B3b, s. 549-556.
- Resko B., Baranyi P., Hashimoto H., 2003. Camera Control with Disparity Matching in Stereo Vision by Artificial Neural Networks. *In Workshop on Intelligent Solutions in Embedded Systems*, Vienna, Austria.
- Tadeusiewicz R., 1993. *Sieci neuronowe*. AOW Warszawa.
- Tadeusiewicz R., Gąciarz T., Borowik B., Leper B., 2007. *Odkrywanie właściwości sieci neuronowych przy użyciu programów w języku C\#*. Polska Akademia Umiejętności. Międzywydziałowa Komisja Nauk Technicznych, Kraków.
- Törma M, 1996, Self-organizing Neural Network in Feature Extraction. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vienna, Austria, Vol. XVIII, Part B2, s.374-379.
- Wang J., Hsiao C., 1999. On Disparity Matching in Stereo Vision via a Neural Network Framework, *Proc. Natl. Sci. Counc. ROC(A)*, Vol. 23, No. 5, s. 665-678.
- Zhou L., Yang X., 2008. Use of Neural Networks for Land Cover Classification from Remotely Sensed Imagery. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Beijing, China, Vol. XXI, Part B7, s. 575-578.

DODATEK: FORMALNY OPIS ALGORYTMU DOPASOWYWANIA SYGNATUR

Sieci impulsujące ICM (*Intersecting Cortical Model*) generują – na podstawie obrazu w stopniach szarości - sekwencje obrazów *binarnych*, na których w kolejnych krokach obliczeniowych uwypuklane są różne fragmenty obrazu wejściowego. W oparciu o tę własność sieci w 1994 roku Johnson wprowadził pojęcie sygnatury (Lindblad *et al.*, 2005). Podstawą obliczania sygnatur jest funkcja czasowa $G(i)$, otrzymywana w wyniku sumowania wyjść neuronów („białych pikseli”) w każdym kroku obliczeniowym:

$$G(i) = \sum_{g,h} B_{gh}(i) \quad (D.1)$$

przy czym $B_{gh} \in [0,1]$ oznacza piksel binarnego obrazu o współrzędnych (g,h) , $i=1...N$.

W celu sprawdzenia efektywności dopasowywania fragmentów obrazów na podstawie ich sygnatur wybrano $K=44$ fragmenty obrazów ze zdjęcia lotniczego a następnie zlokalizowano ich odpowiedniki na zdjęciu następnym. Z tego ostatniego zdjęcia wycięto większe fragmenty, które zawierały zlokalizowane uprzednio podobrazy otoczone dość dużym marginesem. Współrzędne lewego górnego rogu zlokalizowanego podobrazu oznaczono wartościami $(0, 0)$. Po fragmencie obrazu przesuwano ramkę o wielkości równej szukanemu pobobrazowi. Ramkę dzielono na $J=4$ lub $J=6$ równych części, i dla każdej z nich wyznaczano sygnatury, które następnie porównywano z sygnaturami, otrzymanymi w analogiczny sposób dla podobrazu poprzedniego zdjęcia (poszukiwanego wzorca). Dla każdej pary odpowiadających sobie sygnatur obliczano wskaźnik:

$$q_{xyj} = \sum_{i=M}^{i=O} |G_j(i) - G_{xyj}(i)| \quad (D.2)$$

gdzie M i O określają początek i koniec sygnatury ($O-M < N$), (x, y) oznacza przesunięcie analizowanej ramki względem lokalizacji prawidłowej $(0,0)$, a $j=1...J$ jest numerem części,

na które został podzielony podobraz (por. rys. 1).

Sumaryczny wskaźnik dopasowania sygnatur dla całej ramki wynosił:

$$Q_{xy}^0 = \frac{1}{J} \sum_{j=1}^{j=J} q_{xyj} \quad (D.3)$$

przy uwzględnieniu wszystkich J części (indeks 0 oznacza brak – czyli zero - części nieuwzględnionych). W przypadku odrzucenia jednego najgorszego wskaźnika:

$$Q_{xy}^1 = \frac{1}{J-1} \left(Q_{xy}^0 - \max_j(q_{xyj}) \right) \quad (D.4)$$

Po osiągnięciu granic zmienności przesunięć ramki (x, y) oblicza się, w którym miejscu dopasowanie całych ramek było najlepsze, czyli wyznacza się współrzędne:

(X^0, Y^0) takie, że $D(X^0, Y^0) = \min_{x,y} Q_{xy}^0$ w przypadku sumowania wszystkich J części, lub

(X^l, Y^l) takie, że $D(X^l, Y^l) = \min_{x,y} Q^l_{xy}$ w przypadku sumowania z odrzuceniem najgorszego wskaźnika.

Miarą dopasowania dla k -tej pary fragmentów obrazów z bazy danych jest odchyłka lokalizacji lewego górnego narożnika podobrazu od punktu $(0,0)$, czyli:

$$\delta_k^z = \sqrt{(X^z)^2 + (Y^z)^2} \quad (\text{D.5})$$

przy czym $z \in [0,1]$ jest indeksem odpowiadającym liczbie odrzuconych wyników cząstkowych.

Dla całej bazy danych średnia odchyłka lokalizacji jest średnią z odchyłek cząstkowych:

$$\bar{\delta}^z = \frac{1}{K} \sum_{k=1}^{k=K} \delta_k^z \quad (\text{D.6})$$

APPLICATION OF NEURAL NETWORKS TO PHOTOGRAMMETRIC PROCESSES

KEY WORDS: artificial neural network, image matching, digital photogrammetry

Summary

The paper discusses the use of artificial neural networks in geoinformatics, particularly in photogrammetric image analysis. It reviews the relevant international publications (including the ISPRS congress proceedings) and discusses the outcome of research on the use of networks for matching photogrammetric images. The paper shows also results of tests, described in the literature, in which neural networks were applied to perform tasks such as feature extraction, multispectral image classification, camera calibration and matching.

The idea of using neural networks is based on the selection of special representations. The essence of the neural networks-based methodology consists of preparing suitable representations of image fragments and of using them to classify various types of neural networks. One of the methods adopted was based on the distribution and direction of image gradient module value.

The research was conducted on forty four sub-images, taken from aerial photographs of two Polish cities: Bytom and Cracow. The areas shown in those images differed in their terrain cover. The images were divided into three categories: full sub-images, sub-images divided into 4 parts, and sub-images divided into 6 small parts.

The research involved the Intersecting Cortical Model (ICM), a version of the Pulse Coupled Neural Network (PCNN), with which the so-called image signatures, i.e., a few dozen-element vectors that describe the image structure were generated. The preliminary results partially confirm the correctness of the approach adopted, despite problems resulting from the complex nature of photogrammetric images.

dr inż. Sławomir Mikrut
e-mail: smikrut@agh.edu.pl
tel. (12) 617 23 02

dr inż. Zbigniew Mikrut
e-mail: zibi@agh.edu.pl
tel. (12) 617 38 53