

Edward Wiśniewski

Spoleczna Wyższa Szkoła Przedsiębiorczości i Zarządzania w Łodzi

ZASTOSOWANIE METOD ROZPOZNAWANIA OBRAZÓW W BADANIACH STATYSTYCZNYCH JEDNOSTEK TERYTORIALNYCH

Streszczenie: Klasyfikacja i grupowanie jednostek terytorialnych mają istotne znaczenie dla prowadzenia badań przestrzennych i analiz porównawczych. Szerokie zastosowanie znajdują tu statystyczne metody wielowymiarowej analizy porównawczej, a w szczególności metody taksonomiczne i czynnikowe. Autor sygnalizuje możliwość wykorzystania metod rozpoznawania obrazów. Obraz określany jest tu jako ilościowy opis obiektu, zjawiska czy procesu. W zadaniu rozpoznawania obrazów chodzi o rozpoznanie przynależności obiektów do pewnych klas. Artykuł zawiera przegląd podstawowych metod rozpoznawania obrazów, istotę zadania rozpoznawania, bardziej szczegółowo omawia metody minimalnoodległościowe oraz analizę przypadku z zastosowaniem omawianej metodologii dla nowych miast województwa zachodniopomorskiego.

Słowa kluczowe: statystyka, taksonomia, jednostka terytorialna.

1. Wstęp

Pojęcie *obraz* kojarzy się najczęściej z dwuwymiarową ilustracją lub trójwymiarową sceną. Polska nazwa *rozpoznawanie obrazów* niezbyt precyzyjnie oddaje zakres tematyczny omawianych zagadnień; lepszy jest anglojęzyczny termin *pattern recognition*, oznaczający rozpoznawanie wzorców [Tadeusiewicz, Flasiński 1991, s. 7]. Rozpoznawaniem obrazów nazywa się podejmowanie decyzji, która polega na ustaleniu przynależności nowego (dotychczas nie rozpoznanego) obiektu lub zjawiska do danej klasy obiektów (obrazu) przez porównywanie właściwości danego obiektu z cechami znanymi i zbadanymi. Klasa (obraz) to zbiorowość obiektów lub zjawisk, które charakteryzują się pewnymi wspólnymi właściwościami, bliskimi ze względu na zbiór pewnych cech [Rozin 1979, s. 62].

Ilościową miarą obrazu jest miara podobieństwa, to jest właściwość, na podstawie której obiekty i zjawiska są grupowane i tworzą obraz. W większości przypadków formalizacja pojęcia miary podobieństwa opiera się na hipotezie spoistości, która zakłada spełnienie dwóch warunków: a) wewnątrz danego obrazu możliwe jest płynne przejście od jednego zjawiska do drugiego; wszystkie pośrednie zjawiska są

postrzegane jako należące do tego samego obrazu; b) zjawiska przy niedużych zmianach nie wykraczają poza granice danego obrazu [Zagorujko 2004, s. 27-28].

Praktyka wypracowała trzy podstawowe sposoby ilościowego wyrażenia miary podobieństwa: współczynniki asocjacji, współczynniki korelacji i wskaźniki odległości.

W rozpoznawaniu obrazów możliwe są dwa podejścia [Jajuga 1990, s. 32-33]:

- podejście całościowe, polegające na wzięciu pod uwagę wszystkich cech rozpoznawanego obiektu i podjęciu decyzji o jego przynależności w jednym akcie decyzyjnym; dominują tu metody oparte na pojęciu odległości w przestrzeni cech, metody oparte na aproksymacji funkcji przynależności, metody probabilistyczne;
- podejście strukturalne, polegające na wyróżnieniu określonych elementów oraz ustaleniu ich wzajemnych relacji; właściwe rozpoznawanie dokonywane jest na podstawie opisu strukturalnego.

2. Zadanie rozpoznawania obrazu

Niech D będzie zbiorem obiektów lub zjawisk podlegających rozpoznawaniu. Na zbiorze tym zdefiniowana jest relacja równoważności $K \subset D \times D$, która określa rozbięcie zbioru D na kolekcję klas $\{D^i\}$, odpowiadających poszczególnym obrazom. Relacji K nadaje się nazwę *klasyfikacji*.

Jeśli I jest liczbą klas generowanych przez relację K , a zbiór indeksów klas oznaczmy I , wówczas:

$$D = \bigcup D^i, \quad (1)$$

$$i \in I,$$

$$\forall \mu, \nu \in I [D^\mu \cup D^\nu] = 0, \quad (2)$$

$$\forall d^\mu, d^\nu \in D [\langle d^\mu, d^\nu \rangle \in K \Rightarrow \exists_{i \in I} (d^\mu \in D^i) \wedge (d^\nu \in D^i)]. \quad (3)$$

Z opisu relacji K i zbioru I wynika istnienie odwzorowania:

$$A: D \rightarrow I \quad (4)$$

o właściwościach:

$$\forall_{d \in D} [A(d) = i = d \in D^i]. \quad (5)$$

W zadaniu rozpoznawania dąży się do tego, aby skonstruować algorytm realizujący odwzorowanie:

$$A: D \rightarrow I \cup \{i_0\} \quad (6)$$

takie, aby pewna miara $Q(A, A')$ nazywana oceną jakości algorytmu rozpoznawania A' była minimalna. Jednoelementowy zbiór $\{i_0\}$ symbolizuje brak odpowiedzi (decyzja typu „nie wiem”).

Odwzorowanie A jest realizowane jako złożenie trzech odwzorowań: [Tadeusiewicz, Flasiński 1991, s. 46-49]:

$$A = F^* C^* B. \quad (7)$$

Początkową częścią każdego algorytmu rozpoznającego jest pomiar cech wszystkich obiektów, zarówno wzorcowych należących do ciągu uczącego, jak i rozpoznawanych. Określenie cech prowadzi do zamiany obiektów $d \in D$ w punkty pewnej przestrzeni, której struktura jest arbitralna i zdeterminowana przez możliwości pomiarowe. W odwzorowaniu B zakłada się, że elementami przestrzeni cech X są wektory n -elementowe $\underline{x} = \langle x_1, x_2, \dots, x_n \rangle \in X$. Składowe tych wektorów traktuje się jako liczby $x_v \in R$, określające ilościową miarę określonej cechy. Przestrzeń X traktowana jest jako n -wymiarowa przestrzeń euklidesowa $X \subseteq R^n$.

W odwzorowaniu C należy ustalić pewną miarę podobieństwa nieznanego obiektu $d \in D$ do poszczególnych klas D_i . Na podstawie określonego wektora cech \underline{x} oblicza się funkcje przynależności $C^i(\underline{x})$, $i = 1, 2, \dots, L$. Wartości tych funkcji określają miarę przynależności nieznanego obiektu d do klas D^i .

W literaturze proponuje się rozmaite definicje funkcji przynależności $C^i(\underline{x})$ (m.in. [Chen 1973, James 1988, Kurzyński 1997]); sposób realizacji odwzorowania C rozróżnia metody rozpoznawania.

Odwzorowanie F służy do ustalenia ostatecznej decyzji: rozpoznanie lub jego brak. Zazwyczaj wykorzystuje się regułę majoryzacyjną:

$$\forall \underline{x} \in X \left[F(C^1(\underline{x}), C^2(\underline{x}), \dots, C^L(\underline{x})) = i \equiv \forall_{\substack{h \in I \\ h \neq i}} [C^h(\underline{x}) < C^i(\underline{x})] \right], \quad (8)$$

tzn. podejmowana jest decyzja o przynależności obiektu $d \in D$, opisywanego wektorem $\underline{x} \in D$, do tej klasy $i \in I$, dla której wartość funkcji przynależności $C^i(\underline{x})$ jest maksymalna.

Twórca algorytmu A' dysponuje zazwyczaj jedynie wiedzą na temat *ciągu uczącego*, ma natomiast znikomy dostęp do informacji na temat odwzorowania A .

Funkcja przynależności opiera się na wykorzystaniu ciągu uczącego U , który można zdefiniować jako zbiór par:

$$U = \{ \langle \underline{x}^k, i^k \rangle, k = 1, 2, \dots, N \}, \quad (9)$$

$$\text{gdzie: } \underline{x}^k = B(d^k) \wedge d^k \in D, \quad (10)$$

$$i^k \in I \wedge i^k = A(d^k). \quad (11)$$

Elementy zbioru U nazywa się przykładami. Każdy przykład składa się z kompletnego wektora cech \underline{x}^k pewnego obiektu d^k oraz z informacji na temat numeru klasy i^k , do którego obiekt powinien być zaliczony. Zbiór U można zdekomponować na L podzbiorów $U^i (i = 1, 2, \dots, L)$ w ten sposób, aby w i -tym podzbiorze znajdowały się wyłącznie obiekty należące do i -tej klasy.

Wybór elementów d^k należących do ciągu uczącego U powinien zapewnić jego reprezentatywność. W praktyce ciąg ten stanowi próbę losowo pobraną ze zbioru D .

3. Metody rozpoznawania obrazów

Objętość literatury z dziedziny rozpoznawania obrazów powoduje, że liczba różnorodnych metod rozpoznawania sięga setek (m.in. [Andrews 1972; Duda, Hart 1973; Tadeusiewicz 1985; Bobrowski 1987; Patrick 1988; Walesiak 2002]).

Podziału metod rozpoznawania obrazów można dokonać na podstawie klasyfikacji trzech odwzorowań F , C i B , składających się na rozpoznanie A' . Tabela 1. prezentuje podział metod według różnych kryteriów. Ogólnie można wyróżnić metody całościowe i strukturalne. W obrębie tych klas metod wydziela się różnorodne podklasy.

Najwięcej ciekawych możliwości podziału metod rozpoznawania obrazów wynika ze sposobu realizacji odwzorowania C , rozpatrywanego przy założeniu, że odwzorowanie F jest zadane jako jeden ze sposobów. Główną charakterystyką stopnia złożoności odwzorowania C jest liczba rozpoznawanych klas L , określająca moc zbioru, np. C_2 – zadanie dychotomizacji. Klasyfikacja oparta na odwzorowaniu B jest trudniejsza: recepcja jest zawsze silnie uzależniona od właściwości rozpoznawanych obiektów $d \in D$.

4. Metody minimalnoodległościowe

W metodach minimalnoodległościowych odwzorowanie C wiąże się z pojęciem odległości w przestrzeni X , która w tym celu musi być związana z metryką, zapisaną ogólnie:

$$\rho: X * X \rightarrow R_*. \quad (12)$$

Problem wyboru właściwej metryki rozwiązuje się empirycznie (metoda prób i błędów, decyzje arbitralne itp.), aczkolwiek dla metod minimalnoodległościowych jest to problem kluczowy, mający wpływ na rezultat rozpoznawania. Najbardziej naturalną jest metryka euklidesowa definiowana.

$$\rho(\underline{x}^u, \underline{x}^v) = \sqrt{\sum_{v=1}^n (x_v^u - x_v^v)^2}. \quad (13)$$

Inne, bardziej popularne metryki to: metryka Manhattan, metryka Czebyszewa, metryka Mahalanobisa, metryka Bhattacharyya [Sobczak, Malina 1978].

W metodach minimalnoodległościowych należy wybrać jako rozpoznanie i -tę klasę, do której należy obiekt $x^{i,k} \in U$ najbliższy (w myśl przyjętej metryki) rozpoznawanemu obiektowi d , czyli reprezentującemu go wektorowi X . Odwzorowanie można zapisać w postaci:

$$C^i(\underline{x}) = \frac{1}{\rho(\underline{x}, \underline{x}^{i,k}) + \varepsilon}, \quad i = 1, 2, \dots, L. \tag{14}$$

Element $\underline{x}^{i,k}$ należy do podzbioru U^i , a mała stała dodatnia ε wprowadzona jest dla zapewnienia warunku $C^i(\underline{x}) < \infty$, gdyż teoretycznie możliwe jest uzyskanie pokrycia punktów, tj. $\rho(\underline{x}, \underline{x}^{i,k}) = 0$.

Tabela 1. Klasyfikacja metod rozpoznawania obrazów

Podział ogólny	Kryteria		
	Podjęcie decyzji – F	Funkcja przynależności – C	Recepcja – B
CAŁOŚCIOWE	i – identyfikacja k – klasteryzacja j – jednoznaczne $F: R^L \rightarrow I$ o – odmowa $F: R^L \rightarrow I \{i_0\}$ w – wieloznaczne $F: R^L \rightarrow 2^I$ e – etapowe t – totalne	J – liczba rozpoznawanych klas g – gotowa funkcja przynależności w – wybrana p – probabilistyczna p_z – znany rozkład prawdopodobieństwa p_n – nieznany rozkład prawdopodobieństwa	n – rozmiar przeszerzenia cech i – cechy ilościowe (liczby) j – cechy jakościowe (kody wartości) o – cechy opisowe (porównania) b – cechy binarne (obecność lub brak właściwości)
	l – lingwistyczne π – z gramatyką łańcuchową d – z gramatyką drzewkową g – z gramatyką grafową r – relacyjne p – z programowaniem dynamicznym s – quasi-semantyczne	u – z uczeniem n – uczenie z „nauczycielem” k – uczenie bez „nauczyciela” p – uczenie parametryczne b – uczenie nieparametryczne m – minimalno odległościowe h – heurystyczne	

Źródło: opracowanie własne.

4.1. Metoda NN – najbliższy sąsiad

W wariancie tym wybór elementu $\underline{x}^{i,k}$ we wzorze () jest dokonywany zgodnie z regułą:

$$\rho(\underline{x}, \underline{x}^{i,k}) = \min_{x^u \in U^i} (\underline{x}, x^u), \tag{15}$$

Algorytm tej metody można opisać stosowaną w literaturze informatycznej notacją pascalopodobnej, przyjmując założenia definiujące tablice, zmienne i funkcje [Tadeusiewicz, Flasiński 1991]:

numclass – liczba rozpoznawanych klas (L),
dim – wymiar przestrzeni klas (n),
num – liczba obiektów ciągu uczącego (N),
sampl [1...num][1...dim+1] – ciąg uczący (U),
rec – identyfikator rozpoznawanego obrazu (i),
obj [1...dim] – rozpoznawany obiekt (X),
dist(sampl(k),obj) – funkcja licząca odległość ρ między k -tym elementem ciągu uczącego a rozpoznanym obiektem.

Algorytm:

```

procedure NN rec (obj; var rec);
begin
rec := 0; min := Max Real
for k := 1 to num do
if dist(sampl[k],obj) < min then
begin
min := dist (sampl[k], obj);
rec := sampl[k] [dim+1];
end
end

```

Metoda NN jest bardzo wrażliwa na błędy ciągu uczącego U .

4.2. Metoda α NN

W metodzie tej wprowadza się parametr α , którego wybór jest arbitralny, jednak tak, aby

$$\alpha \ll \min_{i \in I} N^i, \quad (16)$$

W praktyce α jest małą liczbą całkowitą. Po określeniu α i pojawieniu się obiektu rozpoznawania \underline{x} oblicza się wartości odległości tego obiektu od wszystkich obiektów ciągu uczącego $\rho(\underline{x}, \underline{x}^k)$, $k = 1, 2, \dots, N$. Następnie dokonuje się uporządkowania ciągu uczącego według rosnących odległości, wprowadzając numery $v = 1, 2, \dots, n$:

$$U = \{ \langle \underline{x}^v, i^v \rangle, v = 1, 2, \dots, n \}, \quad (17)$$

przy czym:

$$\forall_{v,\mu} = [v \succ \mu \Rightarrow \rho(\underline{x}, \underline{x}^v) \succ \rho(\underline{x}, \underline{x}^\mu)]. \quad (18)$$

Prawdopodobieństwo poprawnego rozpoznania zależy od wartości parametru α : $\alpha = 1$ $p = 0,923$; $\alpha = 2$ $p = 0,962$; $\alpha = 3$ $p = 0,974$; $\alpha = 4$ $p = 0,963$; $\alpha = 5$ $p = 0,961$ [Miształ 2002].

Wybiera się α początkowych obiektów ciągu uczącego, tworząc podzbiór:

$$U^\alpha = \{ \langle \underline{x}^v, i^v \rangle, v = 1, 2, \dots, \alpha \}, \quad (19)$$

rozbijając na podzbiory związane z poszczególnymi klasami:

$$U^{\alpha,i} = \{ \langle x^v, i^v \rangle, v < \alpha \ i^v = i \}. \quad (20)$$

Funkcję przynależności dla metody α NN wyznacza się na podstawie liczebności podzbiorów $U^{\alpha,i}$:

$$C^i(\underline{x}) = \# U^{\alpha,i}, i = 1, 2, \dots, L. \quad (21)$$

4.3. Metoda j_N NN

Istota tego algorytmu polega na określeniu przynależności nieznanego obiektu \underline{x} do tej klasy i , do której należy j_n -ty w kolejności element zbioru U , uporządkowanego według reguł określonych wzorami (17) i (18). Funkcję przynależności określa się wzorem:

$$C_i(x) = \delta_{ij_N}, \quad (22)$$

gdzie delta Kroneckera δ jest określona: $\delta = 0$ dla $\mu \neq v$ oraz $\delta = 1$ dla $\mu = v$, natomiast i^{j_N} oznacza wskaźnik przynależności j_n -tego elementu uporządkowanego ciągu uczącego (19).

W literaturze można znaleźć dowody identyczności własności algorytmów j_n NN oraz α NN dla $j_N = \text{ent}[(\alpha+1)/2]$, gdzie $\text{ent}(\mu)$ oznacza część całkowitą liczby rzeczywistej μ [Kulikowski 1968]. Można też zauważyć, że metody α NN i j_n NN sprowadzają się do podstawowego algorytmu NN dla $\alpha = 1$ ($j_N = 1$).

5. Zastosowanie metody rozpoznawania obrazów

– analiza przypadku: nowe miasta w województwie zachodniopomorskim

W celu skonstruowania ciągu uczącego wykorzystano wyniki badań A. Harańczyk dotyczące miast polskich [Harańczyk 1998, s. 57-69], w zakresie oceny przekształceń strukturalnych w nich w okresie transformacji ustrojowej. W badaniach tych zastosowano procedurę analizy zmiennych diagnostycznych oraz klasyfikacji i grupowania 860 miast polskich. W ostatniej dekadzie stany przekształceń uległy zmianom, które zostały uwzględnione.

Do badania przekształceń strukturalnych przyjęto siedem zmiennych diagnostycznych:

- X_1 – liczba przedsiębiorstw usługowych na 1000 osób pracujących,
- X_2 – liczba mieszkań oddanych do użytku na 1000 mieszkańców,
- X_3 – liczba sprzedanych mieszkań komunalnych na 10 000 mieszkańców,
- X_4 – liczba spółek z udziałem kapitału zagranicznego na 10 000 mieszkańców,
- X_5 – liczba fundacji,
- X_6 – liczba podmiotów gospodarczych na 1 km²,
- X_7 – liczba podmiotów gospodarczych na 1000 mieszkańców.

Dla wszystkich miast w Polsce, w ramach pięciu grup wielkości według liczby mieszkańców, obliczono wskaźniki syntetyczne W_i . W celu oceny przestrzennego zróżnicowania przekształceń strukturalnych zachodzących w miastach polskich otrzymane wyniki klasyfikacji w grupach wielkościowych (najmniejsze: do 5 tys., małe: 5-20 tys., średnie: 20-50 tys., średnie większe: 50-100 tys., duże: powyżej 100 tys. mieszkańców) podzielono na cztery klasy według stanu zaawansowania analizowanego problemu [Harańczyk 1998, s. 62]. Ilościowy podział oraz jego strukturę przedstawia tabela 2.

Tabela 2. Rozkład ilościowy miast polskich według standaryzowanych sum

Grupy miast	Klasa 1		Klasa 2		Klasa 3		Klasa 4	
	liczba	%	liczba	%	liczba	%	liczba	%
I	222	51,7	40	12,6	5	5,2	3	16,7
II	142	33,1	168	53,0	43	44,8	6	33,3
III	45	10,5	77	24,2	15	15,6	1	5,6
IV	12	2,8	18	5,7	15	15,6	6	33,3
V	8	1,9	14	4,5	18	18,8	2	11,1
	429	49,9	317	36,9	96	11,2	18	2,0

Źródło: opracowanie własne.

W Polsce co pewien czas kilka miejscowości otrzymuje prawa miejskie. W celu weryfikacji metodologii rozpoznawania obrazów wybrano trzy z najmłodszych miast województwa zachodniopomorskiego: Dziwnów (od 1 stycznia 2004 r.), Tychowo (od 1 stycznia 2010 r.) oraz Gościno (z dniem 1 stycznia 2011 r.), włączając je do ww. oceny. Nowe miasta z natury rzeczy należą do najmniejszych: Dziwnów – miejscowość wypoczynkowa nad Bałtykiem z małym portem morskim: 2904 mieszkańców, Tychowo – na pograniczu Pojezierza Drawskiego i Równiny Białogardzkiej: 2492 mieszkańców, oczekujące na prawa miejskie Gościno w powiecie kołobrzeskim: 2456 mieszkańców [Województwo zachodniopomorskie... 2007]. Zastosowano algorytmy minimalnoodległościowe z wykorzystaniem metryki euklidesowej, realizując obliczenia programem STATISTICA. Ostateczne wyniki prezentuje tabela 3.

Tabela 3. Minimalne wartości odległości miasta \underline{x} i klasy $\underline{x}^{i,k}$

Algorytm	Miasto		
	Dziwnów	Tychowo	Gościno
NN	$\rho_2^i = 0,241$	$\rho_1^i = 0,422$	$\rho_1^i = 0,401$
α NN	$\rho_3^i = 0,337$	$\rho_1^i = 0,311$	$\rho_1^i = 0,418$
j_N NN	$\rho_2^i = 0,135$	$\rho_1^i = 0,301$	$\rho_1^i = 0,435$

Źródło: opracowanie własne.

Nowe miasta w województwie zachodniopomorskim znajdujące się w grupie I miast (ciągu uczącego) wykazują zróżnicowanie w poziomie zmian strukturalnych: pewną przewagę wykazuje Dziwnów, a Tychowo i Gościno są w fazie początkowej procesu zmian. Wartości w tabeli 3. są najmniejsze z kompletnej konfrontacji par miasto \underline{x} i klasy $\underline{x}^{i,k}$. Dziwnów został zaliczony do klasy 2., Tychowo i Gościno do klasy 1. grupy I miast polskich.

6. Podsumowanie

1. Przedstawione metody rozpoznawania obrazów znajdują większe możliwości zastosowań w różnych dziedzinach życia społeczno-gospodarczego i badaniach naukowych. W ogromnym tempie wzrasta liczba monografii i publikacji dotyczących tej problematyki. Interesujące są periodyki, np. „Pattern Recognition”, „Computer Visio” czy „PAMI-IEEE Transactions on Pattern Analysis and Machine Intelligence”. Również w Polsce stale publikuje się prace poświęcone rozpoznawaniu obrazów, np. w „Postęпах Cybernetyki” czy „Podstawach Sterowania”.

2. Omawiana metodologia może także znaleźć też zastosowanie w problemach badań jednostek terytorialnych. Istnieje szeroka gama metod rozpoznawania obrazów oraz możliwości stosowania metod mieszanych. Wydaje się, że większe możliwości występują w zakresie rozpoznawania zjawisk i procesów w obszarze rozwoju regionalnego i lokalnego, gdyż barierą formalną jest „konstrukcja” nowych obiektów terytorialnych.

3. Autor podjął próbę rozpoznawania trzech nowych miast w województwie zachodniopomorskim, stosując określony ciąg uczący na podstawie danych z prac A. Harańczyk. Uzyskane rezultaty relatywnie dobrze odzwierciedlają sytuację rzeczywistą tych miast. Zastosowane algorytmy minimalnoodległościowe dla badanego przypadku okazały się efektywne.

Literatura

- Andrews H.C., *Introduction to Mathematical Technique in Pattern Recognition*, John Wiley, New York 1972.
- Bobrowski L., *Dyskryminacja symetryczna w rozpoznawaniu obrazów. Teorie, algorytmy, zastosowania*, Ossolineum, Wrocław 1987.
- Chen C.H., *Statistical Pattern Recognition*, D.C. Hayden Book Co., Washington 1973.
- Duda O., Hart P., *Pattern classification and scene analysis*, John Wiley, New York 1973.
- Harańczyk A., *Miasta Polski w procesie globalizacji gospodarki*, PWN, Warszawa 1998.
- Jajuga K., *Statystyczna teoria rozpoznawania obrazów*, PWN, Warszawa 1990.
- James M., *Pattern Recognition*, John Wiley, New York 1988.
- Kulikowski J., Kurzyński M., *Rozpoznawanie obrazów. Metody statystyczne*, Oficyna Wydawnicza Politechniki Wrocławskiej, Wrocław 1997.
- Misztal M., *O zastosowaniu statystycznych metod rozpoznawania obrazów do wspomagania procesów podejmowania decyzji w diagnostyce medycznej*, StatSoft, Kraków 2002.
- Panek T., *Statystyczne metody wielowymiarowej analizy porównawczej*, Szkoła Główna Handlowa, Warszawa 2009.
- Patrick E.A., *Fundamentals of Pattern Recognition*, Prentice-Hall, Detroit 1988.
- Rozin B.B., *Teoria rozpoznawania obrazów w badaniach ekonomicznych*, PWN, Warszawa 1979.
- Sobczak W., Malina W., *Metody selekcji informacji*, Wydawnictwa Naukowo-Techniczne, Warszawa 1978.
- Strahl D. (red.), *Metody oceny rozwoju regionalnego*, Wydawnictwo Akademii Ekonomicznej we Wrocławiu, Wrocław 2006.
- Tadeusiewicz R., *Rozpoznawanie obrazów – zarys teorii*, Uniwersytet Jagielloński, Kraków 1985.
- Tadeusiewicz R., Flesiński M., *Rozpoznawanie obrazów*, PWN, Warszawa 1991.
- Walesiak M., *Uogólniona miara odległości w statystycznej analizie wielowymiarowej*, Wydawnictwo Akademii Ekonomicznej we Wrocławiu, Wrocław 2002.
- Województwo zachodniopomorskie. Podregiony, powiaty, gminy – 2007*, Urząd Statystyczny w Szczecinie, Szczecin 2008.
- Zagarujko N.G., *Klasyfikacja zadań rozpoznawania obrazów*, „Wyczystlielnyje Sistiemy” 1971, nr 44.

APPLICATION OF PATTERN RECOGNITION METHODS IN STATISTICAL SURVEYS OF TERRITORIAL UNITS

Summary: Classification and grouping of territorial units are important for conducting spatial research and comparative analysis. There is a wide application of statistical methods for multidimensional comparative analysis, in particular taxonomic and factorial methods. The author indicates the possibility of using pattern recognition methods. The pattern here is defined as the quantitative description on an object, event, phenomenon or process. In general, the pattern recognition task comes down to the identification of an object belonging to certain classes. The article contains an overview of basic pattern recognition methods, the essence of the recognition task, and in details discusses minimal-distance methods and the case study using this methodology for new towns of West Pomerania Voivodeship.