P. 3361/83

ZESZYTY
NAUKOWE
POLITECHNIKI
ŚLĄSKIEJ

MECHANIKA

Z. 79
GLIWICE
1983

P 3361/83

POLITECHNIKA ŚLĄSKA

ZESZYTY NAUKOWE

Nr 764

WOJCIECH CHOLEWA

METODA DIAGNOZOWANIA MASZYN Z ZASTOSOWANIEM ZBIORÓW ROZMYTYCH

GLIWICE 1983

OPINIODAWCY

Prof. dr hab. Czesław Cempel
Doc. dr hab. inż. Ernest Czogała

KOLEGIUM REDAKCYJNE

Wiesław Gabzdyl (redaktor naczelny), *Józef Wojnarowski* (redaktor działu), *Elżbieta Stinzing* (sekretarz redakcji)

OPRACOWANIE REDAKCYJNE

Alicja Nowacka

Wydano za zgodą
Rektora Politechniki Śląskiej

PL ISSN 0434-0817

Dział Wydawnictw Politechniki Śląskiej
ul. Kujawska 3, 44-100 Gliwice

Nakł. 150+85 Ark. wyd. 6,82 Ark. druk. 7 Papier drukowy kl. V 70x100, 70 g
Oddano do druku 10.02.1983 Podpis. do druku 22.03.1983 Druk ukończ. w maju 1983
Zamówienie 257/83 Z-23 Cena zł 68,–

Skład, fotokopie, druk i oprawę
wykonano w Zakładzie Graficznym Politechniki Śląskiej w Gliwicach

SPIS TREŚCI

	Str.
Oznaczenia	5
1. WSTĘP	7
1.1. Cel pracy	8
1.2. Tezy pracy	8
1.3. Zakres pracy	9
2. OBIEKT BADAŃ	11
2.1. Własności obiektu badań	11
2.2. Czas	11
2.3. Sygnały	12
2.4. Cechy	15
2.5. Przekształcanie cech	16
2.6. Obrazy	21
3. MODEL DIAGNOSTYCZNY MASZYNY	26
3.1. Wejście, wyjście modelu maszyny	27
3.2. System	28
3.3. Obraz stanu maszyny	29
3.4. Działanie modelu	32
4. KONCEPCJA UKŁADU DIAGNOZUJĄCEGO	36
4.1. Działanie układu diagnozującego	36
4.2. Diagnozowanie	41
4.3. Weryfikacja układu diagnozującego	43
5. PRZYKŁAD ZASTOSOWANIA	45
5.1. Wyznaczenie klas stanów	48
5.2. Wyznaczenie obrazów stanów	49
5.3. Minimalizacja liczby wymiarów przestrzeni cech sygnałów diagnostycznych	51
5.4. Wyznaczenie obrazów sygnałów diagnostycznych	55
5.5. Wyznaczenie rozmytej relacji między obrazami sygnałów diagnostycznych i obrazami stanu	56
5.6. Weryfikacja układu diagnozującego	57
6. UWAGI KOŃCOWE	59

Str.

DODATKI

A. PORÓWNYWANIE ELEMENTÓW PRZESTRZENI WARTOŚCI CECH 62

A.1. Odległość elementów przestrzeni wartości cech 62

A.1.1. Odległość w przestrzeni wartości cech ilościowych . 63

A.1.2. Odległość w przestrzeni wartości cech jakościowych 65

A.2. Podobieństwo elementów przestrzeni wartości cech 67

A.2.1. Podobieństwo elementów w przestrzeni wartości cech ilościowych 68

A.2.2. Podobieństwo elementów w przestrzeni wartości cech jakościowych 68

A.2.3. Podobieństwo elementów w ogólnej przestrzeni wartości cech 69

A.2.4. Układy pomiarowe 70

A.3. Porównywanie elementów ze zbiorami elementów 71

A.4. Porównywanie zbiorów ze zbiorami 72

B. KRYTERIA GRUPOWANIA ELEMENTÓW PRZESTRZENI CECH 73

B.1. Kryteria wykorzystujące funkcję odległości 74

B.2. Kryteria wykorzystujące funkcję podobieństwa 77

B.3. Kryteria progowe 79

C. ALGORYTMY GRUPOWANIA ELEMENTÓW PRZESTRZENI CECH 80

C.1. Algorytmy poszukiwania ogólnego ekstremum funkcji kryterialnej 81

C.2. Algorytmy hierarchicznego podziału i grupowania 81

C.3. Algorytmy iteracyjne 83

C.4. Algorytmy wykorzystujące zbiory rozmyte 86

D. ZBIORY I RELACJE ROZMYTE 88

D.1. Pojęcie zbioru i relacji rozmytej 88

D.2. Działania 90

D.3. Relacyjne równania rozmyte 94

D.4. Układy równań rozmytych 99

D.4.1. Ograniczone rozwiązanie układu równań rozmytych ... 100

D.5. Zbiór reprezentantów 101

LITERATURA 104

STRESZCZENIA 109

OZNACZENIA (ZASADY OGÓLNE)

1. Małe litery oznaczają stałe, zmienne, funkcje, operatory i elementy zbiorów.
2. Cieniowane małe litery (np. $\tilde{v}$) oznaczają elementy przestrzeni i macierze jednowierszowe.
3. Duże litery oznaczają zbiory:
   - $\tilde{V}$ - zbiór rozmyty,
   - $\tilde{V}'$ - dopełnienie zbioru rozmytego $\tilde{V}$,
   - $q(\tilde{V})$ - zbiór reprezentantów zbioru rozmytego $\tilde{V}$,
   - $V_{>\alpha}$ - zbiór będący α-obcięciem zbioru rozmytego $\tilde{V}$,
   - $\overline{V}$ - moc zbioru V,
   - $\overline{\tilde{V}}$ - moc zbioru rozmytego $\tilde{V}$.
4. Podzbiór oznaczany jest przez dodanie prawego dolnego indeksu do oznaczenia zbioru - np.: $V_{i,j} \subset V_i \subset V$. Jeżeli nie prowadzi to do nieporozumień, to ciągi dolnych indeksów (lub ich początkowe elementy) mogą być pominięte - np.: V (lub V_j) zamiast $V_{i,j}$.
5. Elementy zbiorów oznaczane są małymi literami. Ciąg prawych dolnych indeksów jest równy ciągowi prawych dolnych indeksów w oznaczeniu zbioru, uzupełnionemu o indeks identyfikujący element w zbiorze, np.: $v_{i,j,k} \in V_{i,j}$. Jeżeli nie prowadzi to do nieporozumień, to ciągi dolnych indeksów (lub ich początkowe elementy) mogą być pominięte - np.: $v_{j,k} \in V_{i,j}$ lub $v_k \in V_{i,j}$ lub $v \in V_{i,j}$.
6. Cieniowane duże litery (np.: $\tilde{V}$) oznaczają przestrzenie, zbiory elementów przestrzeni i macierze prostokątne.
7. Element macierzy oznaczany jest przez zapisanie indeksów (identyfikujących ten element) w nawiasach kwadratowych za oznaczeniem macierzy - np.: $\tilde{b}[2]$, $A[1,2]$.
8. Wymiar przestrzeni zapisywany jest jako prawy górny indeks oznaczenia przestrzeni lub elementu - np.: $\tilde{v}^n \in \tilde{V}^n$. Stosowane jest również oznaczenie dim - np.: $n = \dim(\tilde{V})$. Wymiary macierzy (liczba wierszy, liczba kolumn macierzy prostokątnych oraz liczba elementów macierzy jednowierszowych) zapisywane są jako górne prawe indeksy oznaczenia macierzy - np.: $\tilde{b}^m$, $A^{k,l}$. Oznaczenia wymiarów mogą być pominięte.

9. Lewy górny indeks zapisany w nawiasach okrągłych oznacza krok interacji, w którym wyznaczono wielkość opisaną tym indeksem - np.: $^{(t)}g$ oznacza wielkość g po t-tym kroku iteracji. Indeks ten może być pominięty.

10. Lewy dolny indeks zapisywany w nawiasach okrągłych oznacza zbiór lub moc zbioru "danych" wykorzystanych do wyznaczenia wielkości opisanej tym indeksem - np.: $_{(n)}x$ oznacza rozwiązanie układu n równań. Indeks ten może być pominięty.

11. Zbiór wartości funkcji v oznaczany jest Cod(v). v(t) oznacza funkcję v argumentu t lub wartość funkcji v dla argumentu t. Wykres funkcji oznaczany jest $v(X) = \{(x,v(x)) : x \in X\}$ - np.: $v([t)$.

12. Nawiasy okrągłe są stosowane, między innymi, do zapisywania uporządkowanego układu elementów - np.: para uporządkowania (a,b).

13. Nawiasy okrągłe i kwadratowe są stosowane, między innymi, do oznaczania odcinków - np.: dla liczb rzeczywistych x $(a : b] = \{x : a < x \leqslant b\}$, $(a = \{x : a < x\}$.

14. Cieniowane nawiasy kwadratowe są stosowane do zapisywania klas abstrakcji - np.: $[\![v]\!]$.

15. Nawiasy klamrowe są stosowane do definiowania zbiorów; $\{v : R\}$ oznacza zbiór elementów v spełniających warunek R. $\{v\}$ może oznaczać jednoelementowy zbiór, do którego należy element v lub może oznaczać zbiór elementów v spełniających warunek wynikający domyślnie z tekstu.

16. Oznaczenia szczegółowe:

$\stackrel{\mathrm{df}}{=}$	- "równy wg definicji",
$\forall_x$	- "dla wszystkich x",
$\exists_y$	- "istnieje takie y, że",
x:=y	- "y zastępuje x",
tr(X)	- ślad macierzy X,
det(X)	- wyznacznik macierzy X,
dim(X)	- wymiar przestrzeni X,
int(x)	- część całkowita argumentu x,
$\subset$	- relacja zawierania zbiorów,
$\sim$	- relacja podobieństwa zbiorów rozmytych,
$\vee,\wedge,\alpha,\varepsilon$	- operatory działań na zbiorach rozmytych,
$\circ,\Delta,\mathcal{D},\mathcal{E}$	- operatory składania relacji rozmytych.

1. WSTĘP

Termin "diagnostyka" wywodzi się z greckiego słowa "diagnosis" oznaczającego rozpoznawanie, określanie. Wynikiem diagnozowania jest diagnoza, czyli określenie stanu chorego (diagnostyka medyczna) lub stanu układu technicznego (diagnostyka techniczna). Celem diagnostyki technicznej jest zwiększenie niezawodności układów technicznych. Podstawowym zadaniem diagnostyki technicznej jest rozpoznawanie stanu (klasy stanu) obiektu badań w warunkach ograniczonej ilości informacji [13].

Problemy omawiane w pracy dotyczą różnych klas obiektów badań. Ze względu na zainteresowania autora, ograniczono zakres rozważań do takich obiektów badań, które są maszynami i z tego powodu obiekt badań nazywany jest dalej maszyną. Dostępnym źródłem informacji o stanie maszyny są sygnały diagnostyczne. Z dużej liczby różnych możliwych sygnałów diagnostycznych maszyn szczególne znaczenie mają hałas i drgania wywołane ich działaniem. Ocenianie stanu maszyn na podstawie sygnałów diagnostycznych może być realizowane z różnym stopniem szczegółowości [20,30.79,66,31,32]. Spotykane obecnie układy diagnozujące są najczęściej prostymi układami elektronicznymi umożliwiającymi sprawdzanie czy wybrane cechy sygnałów diagnostycznych przekraczają przyjęte wartości krytyczne. Umożliwiają one, na podstawie zaobserwowanych przekroczeń, wnioskowanie o prawdopodobnym wystąpieniu stanu krytycznego wymagającego, dla uniknięcia awarii badanej maszyny, przeprowadzenia jej szczegółowych badań, przeglądu lub remontu. Urządzenia diagnozujące o takim przeznaczeniu mogą działać w sposób ciągły w postaci układów zabezpieczających lub wykorzystywane są przez obsługę okresowo. Brak jest ogólnie uznanych metod wyznaczania krytycznych wartości cech sygnałów. Często są one przyjmowane w wyniku arbitralnych rozstrzygnięć na podstawie wyników badań innych obiektów (realizowanych przez innych badaczy). Jest to przyczyną kwestionowania przez wielu badaczy zasadności wniosków wynikających z takich badań.

Odmienną grupę urządzeń stanowią złożone układy diagnozujące umożliwiające identyfikację stanu krytycznego maszyny. Układy takie wykorzystywane mogą być w stanowiskach kontroli technicznej maszyn produkowanych seryjnie, w stanowiskach kontroli technicznej warsztatów remontowych oraz można przewidywać ich wykorzystanie przez obsługę złożonych układów mechanicznych dla wnioskowania o wymaganym terminie i zakresie remontu. Podstawowe problemy w zakresie projektowania złożonych układów diagnozujących związane są z metodami wnioskowania o relacjach między cechami sygnałów diagnostycznych i stanem maszyny. Złożoność maszyn rozpatrywanych ja-

ko źródła sygnałów powoduje, że bezpośrednie, skuteczne przewidywanie tych relacji jest trudne (lub niemożliwe). Cechą charakterystyczną większości prac z tego zakresu jest przyjmowanie pewnego modelu relacji i badanie istotności przyjętego modelu lub poszukiwanie zależności funkcyjnych między ocenami sygnałów i stanem maszyny. Taki sposób postępowania zmusza do indywidualnego rozpatrywania układów: maszyna - oceny sygnałów i powoduje, że osiągnięcie zadowalających wyników badań jest w dużym stopniu zależne od intuicji badacza.

Badania autora związane z wibroakustyczną diagnostyką maszyn zapoczątkowane zostały pod kierunkiem Profesora J. Dietrycha i Docenta Z. Jaskóły. Część zadań objętych niniejszym opracowaniem wykonano podczas realizacji prac badawczych zleconych przez GIG w Katowicach, IPPT w Warszawie oraz COBR POLMATEX-CENARO w Łodzi. Na obecną postać pracy wpłynęły uwagi krytyczne poczynione przez Profesora C. Cempla i Profesora J.Wojnarowskiego, którzy zapoznali się szczegółowo z rękopisem, za co autor wyraża im szczególne podziękowanie.

1.1. Cel pracy

Celem pracy jest sformułowanie metody projektowania układów diagnozujących przeznaczonych do rozpoznawania klas stanów maszyn na podstawie obserwowanych sygnałów diagnostycznych.

1.2. Tezy pracy

A. Istnieje możliwość ograniczenia liczby wymiarów przestrzeni analizowanych w badaniach diagnostycznych maszyn dzięki przekształceniu przestrzeni wartości cech sygnałów diagnostycznych i przestrzeni wartości cech stanów w przestrzenie obrazów cech[1)] sygnałów i przestrzenie obrazów cech[1)] stanów. Stosowanie obrazów cech umożliwia uproszczenie procedur związanych z wyznaczaniem działania układów diagnozujących.

B. Modele klas maszyn wykonanych według wspólnej konstrukcji, opisywane rozmytymi relacjami między obrazami cech sygnałów diagnostycznych i obrazami cech stanów maszyn wyznaczają działanie układu diagnozującego, przeznaczonego do rozpoznawania klas stanów maszyn na podstawie obserwowanych sygnałów diagnostycznych.

[1)] Przestrzeń obrazów cech jest taką przestrzenią, w której wartości współrzędnych elementów równe są wartościom funkcji ich przynależności do rozmytych klas elementów podobnych (wyznaczanych w przestrzeni wartości cech), (por. rozdz. 2.6).

1.3. Zakres pracy

Pracę ujęto w trzech głównych rozdziałach (rozdz.2, 3, 4). Dla uzyskania zwartości tekstu tych rozdziałów część zagadnień opisano w dodatkach A, B, C, D. Prosty przykład zastosowania rozpatrywanej metody postępowania opisano w rozdz. 5.

W rozdziale 2 omówiono zagadnienia związane z obserwowaniem maszyny. Stwierdzono, że własności maszyny oraz ich zmiany w czasie obserwowane są za pośrednictwem sygnałów. Wprowadzono pojęcie cech sygnałów, cech oddziaływań i cech stanu maszyny, rozszerzając w ten sposób zakres znaczeniowy wprowadzonego przez J. Dietrycha (np. [33]) pojęcia "cecha konstrukcyjna". Przyjmując, że sygnały rozpatrywane są jako procesy stochastyczne określone w funkcji czasu "mikro" i "makro" założono, że cechy sygnałów zależne są wyłącznie od czasu "makro". Dla opisania własności przedmiotu poznania niezbędne jest określenie zbioru cech. Zbiór wszystkich możliwych cech orzekających o maszynie jest w praktyce zbiorem nieskończonym, nawet jeżeli maszyna charakteryzuje się bardzo małym stopniem złożoności. Dla umożliwienia praktycznej realizowalności badań konieczna jest obserwacja wyłącznie skończonego zbioru cech. Zaproponowano sposób ograniczania liczby rozpatrywanych cech polegający na wyróżnieniu klas podobnych wartości cech, co umożliwia opisywanie sygnałów za pomocą ich podobieństw do tych wyróżnionych klas. Wprowadzono pojęcie obrazu sygnału jako zbioru wartości podobieństw sygnału do wyróżnionych klas sygnałów. Obrazy sygnałów można interpretować jako wynik przekształcania wielowymiarowej przestrzeni wartości cech w przestrzeń obrazów cech charakteryzującą się małą liczbą wymiarów. Dla wyznaczenia obrazów cech konieczne jest rozwiązanie zadania grupowania elementów przestrzeni cech w zbiory elementów podobnych, co wymaga formalizacji sposobu określania podobieństw elementów. Zagadnienia porównywania (dla potrzeb grupowania i klasyfikacji) elementów z elementami, elementów ze zbiorami oraz zbiorów ze zbiorami w przestrzeniach cech ilościowych i jakościowych omówiono w dodatku A. W dodatku B omówiono kryteria stosowane podczas wyznaczania klas elementów podobnych, a w dodatku C omówiono stosowane algorytmy grupowania, ze szczególnym zwróceniem uwagi na algorytmy hierarchicznego podziału i grupowania oraz algorytmy iteracyjne. Omawiając metody grupowania i klasyfikacji pominięto, pozostające w związku z nimi, tzw. statystyczne metody wyboru wykorzystujące metody statystycznej weryfikacji hipotez, metodę ocen Bayesa oraz metody wykorzystujące bezpośrednio lub pośrednio miary ilości informacji (np. [13, 89, 27]).

W rozdziale 3 omówiono metodę modelowania działania zbioru maszyn. Konieczność modelowania działania zbioru maszyn, a nie jednej wybranej maszyny, wynika z założenia, że działanie układu diagnozującego ma być dobrane odpowiednio do własności klasy maszyn wykonanych wg wspólnej dla nich konstrukcji. Istotą proponowanej metody jest rozpatrywanie maszyny

jako "czarnej skrzynki" bez wnikania w działanie jej elementów. Maszyna będąca "czarną skrzynką" jest układem względnie odosobnionym z otoczenia. Podczas badań prowadzone są obserwacje zmieniających się w czasie oddziaływań otoczenia na maszynę i maszyny na otoczenie. Oddziaływania te oraz stan maszyn mogą być odpowiednio opisane przez zbiory wartości ich cech jakościowych i ilościowych. Zbiory wartości tych cech w chwili obserwacji nazywane są odpowiednio wejściem, wyjściem i stanem maszyny. Umożliwiają one wprowadzenie pojęcia systemu maszyny oraz jej wewnętrznego i zewnętrznego działania. Analizując własności klasy maszyn wykonanych wg wspolnej konstrukcji stwierdzono, że najczęściej system takiego abstrakcyjnego obiektu badań nie może być rozpatrywany jako system przyczynowo-skutkowy, co uzasadnia wprowadzenie opisu działania maszyny w postaci rozmytych relacji między zbiorami uwzględnianych elementów wejścia, wyjścia i stanu maszyny. Wykorzystując wprowadzone wcześniej pojęcie obrazu cech, określono wygodny dla zastosowań praktycznych model klasy maszyn oraz jego rozmyte działanie. Pojęcie zbiorów i relacji rozmytych omówiono w dodatku D. W dodatku tym określono podstawowe działania na zbiorach rozmytych. Szczególną uwagę zwrocono na zagadnienia związane z formułowaniem i rozwiązywaniem równań i układow rownań rozmytych.

W rozdziale 4 omówiono zagadnienia związane z wyznaczaniem działania układów diagnozujących. Dla opisania układu diagnozującego wykorzystano rozwinięte w rozdziale 3 pojęcie systemu, zakładając, że wyjściem układu diagnozującego jest obraz stanu maszyny, a jego wejściem są obrazy wejścia i wyjścia maszyny. Działanie układu diagnozującego określono jako parę relacji rozmytych będących rozwiązaniem odpowiedniego układu równań. Przyjęto, że dla sformułowania układu równań rozmytych, dostępne są wyniki tzw. "badań uczących" w postaci par (wartości cech sygnałów, wartości cech stanu) lub (obraz cech sygnałów, obraz stanu). Omówiono sposób weryfikacji otrzymanego rozwiązania. Wyznaczone działanie układu diagnozującego stanowić może podstawę do przyjęcia szczegółowych założeń konstrukcyjnych dla tego układu przewidzianego do wykonania jako proste w obsłudze urządzenie elektroniczne.

W pracy pominięto problemy dotyczące pomiarów, rejestracji, analizy sygnałów oraz gromadzenia wyników analiz i organizacji banku wyników. Zagadnienia te omówiono częściowo w innych opracowaniach [24, 25, 26].

2. OBIEKT BADAŃ

W rozdziale tym opisano proces obserwowania maszyny będącej obiektem badań.

2.1. Własności obiektu badań

Założono, że obiektami badań są j-te maszyny u_j będące elementami zbioru U maszyn wytworzonych według wspólnej dla nich konstrukcji. Maszyna u_j może być opisana zbiorem W_j jej własności[1)], który jest najczęściej zbiorem nieskończonym, nawet wtedy, gdy maszyna charakteryzuje się małym stopniem złożoności. Ze względu na potrzeby badań oraz ograniczone możliwości poznania dokonuje się wyboru uwzględnianego podczas badań k-tego podzbioru $W_{j,k}$ zbioru własności i właściwości W_j. Wyróżnianie tego podzbioru można zapisać jako przekształcenie rzutowe

$$\delta_k : W_j \longmapsto W_{j,k}. \qquad (2.1.1)$$

Ze względu na własności losowe procesów wytwarzania i procesów destrukcji wywołanych eksploatacją, własności maszyn rozpatrywane są jako procesy stochastyczne $w(\tau,t)$

$$W_{j,k} \stackrel{d}{=} \left\{w(\tau,t)\right\} \subset W_j, \qquad (2.1.2)$$

gdzie:

τ - czas "mikro",

t - czas "makro".

2.2. Czas

Wyjaśnienia wymaga uwzględnianie dwóch zmiennych τ, t w zapisie procesu stochastycznego $w(\tau,t)$.

---

1) Zakres znaczeniowy pojęcia własność obejmuje w tej pracy zakresy znaczeniowe pojęć własność i właściwość, rozróżnianych przez J. Dietrycha [33].

Czas $\tilde{\tau}$ ma znaczenie czasu procesu stochastycznego [72] i nazywany jest w tej pracy czasem "mikro". Czas ten nazywany jest różnie przez różnych autorów - np.: C. Cempel [20] nazywa go czasem dynamicznym. Określając chwile czasu "mikro", zakłada się najczęściej (jawnie lub domyślnie), że są one elementami odcinka czasu, w którym rozpatrywany proces może być uznany za stacjonarny.

Czas t należy interpretować jako identyfikator realizacji procesu stochastycznego $w(\tilde{\tau},t)$. Czas t nazywany jest w tej pracy czasem "makro". C. Cempel [20] nazywa go czasem życia obiektu i czasem ewolucji rozpatrywanego procesu stochastycznego.

Należy zwrócić uwagę na zakres znaczenia pojęcia "czas". Przyjęto, że zbiór czasu jest dowolnym, liniowo uporządkowanym przez relację $\leqslant$ zbiorem, którego elementy nazywane są chwilami czasu i mogą miec różne znaczenie. Elementami tymi mogą być na przykład:

- liczba godzin działania badanej maszyny od chwili jej uruchomienia,
- liczba obrotów wału badanej maszyny od chwili rozpoczęcia badań.

Szczególnym przypadkiem zbioru czasu jest zbiór czasu rzeczywistego będący zbiorem gęstym. Brak możliwości ciągłej obserwacji badanych wielkości jest przyczyną wyróżniania dyskretnego zbioru czasu "makro", definiowanego jako ciąg chwil czasu. Przyjęto, że chwile czasu $\tilde{\tau}$ oraz t mogą być elementami różnych zbiorów czasu.

2.3. Sygnały

Informacja o uwzględnianych własnościach maszyny dociera do obserwatora za pośrednictwem sygnałów. Pod pojęciem sygnału rozumie się dowolny materialny nośnik, najczęściej przebieg wielkości fizycznej, umożliwiający przenoszenie informacji.

Przyjęto, że sygnał f jest procesem stochastycznym

$$f : (\tilde{\tau},t) \longmapsto r \in R, \tag{2.3.1}$$

gdzie:

R - zbiór liczb rzeczywistych.

Dla wyróżnienia tych sygnałów, które obserwowane są we wspólnym dla nich czasie "mikro", zapisuje się je jako wielowymiarowy (n-wymiarowy) proces stochastyczny.

$$f = f^n : (\tilde{\tau},t) \longmapsto r \in R^n. \tag{2.3.2}$$

Zbiór F_j sygnałów $f(\vec{\tau},t)$ uwarunkowanych zbiorem W_j własności j-tego obiektu jest zbiorem nieskończonym. Może on być zapisany

$$F_j = \varphi(W_j), \tag{2.3.3}$$

gdzie:

φ - operator mający znaczenie transmitancji.

Podczas badań obserwowany jest wyłącznie skończony podzbiór tego zbioru. Wyróżnianie ze zbioru F_j, l-tego podzbioru obserwowanych sygnałów można zapisać jako przekształcenie rzutowe

$$\sigma_l : F_j \rightarrow F_{j,l}. \tag{2.3.4}$$

Oznaczając złożenie przekształceń

$$\varphi_l \stackrel{d}{=} \sigma_l \circ \varphi, \tag{2.3.5}$$

można l-ty zbiór obserwowanych sygnałów zapisać

$$F_{j,l} = (\sigma_l \circ \varphi)(W_j) = \varphi_l(W_j). \tag{2.3.6}$$

Określenie zbioru wszystkich tych własności maszyny, których zmiany są przyczyną zmian obserwowanych sygnałów, jest czynnością trudną. Przyjmuje się, że obserwowany sygnał $f(\vec{\tau},t)$ jest uwarunkowany przez uwzględniany k-ty podzbiór $W_{j,k}$ oraz przez nieznany k'-ty podzbiór $W_{j,k'}$ własności maszyny. Większość autorów uwzględnia ten fakt przez przyjęcie założenia, że obserwowany podczas badań przebieg wielkości fizycznej (sygnał) może być zapisany

$$f(\vec{\tau},t) = f_1(\vec{\tau},t) + f_2(\vec{\tau},t) \tag{2.3.7}$$

lub

$$f(\vec{\tau},t) = f_1(\vec{\tau},t) f_2(\vec{\tau},t), \tag{2.3.8}$$

gdzie:

$f_1(\vec{\tau},t)$ - sygnał uwarunkowany wyłącznie uwzględnianym podzbiorem $W_{j,k}$,

$f_2(\vec{\tau},t)$ - sygnał uwarunkowany nieznanym podzbiorem $W_{j,k'}$, sygnał ten nazywany jest szumem.

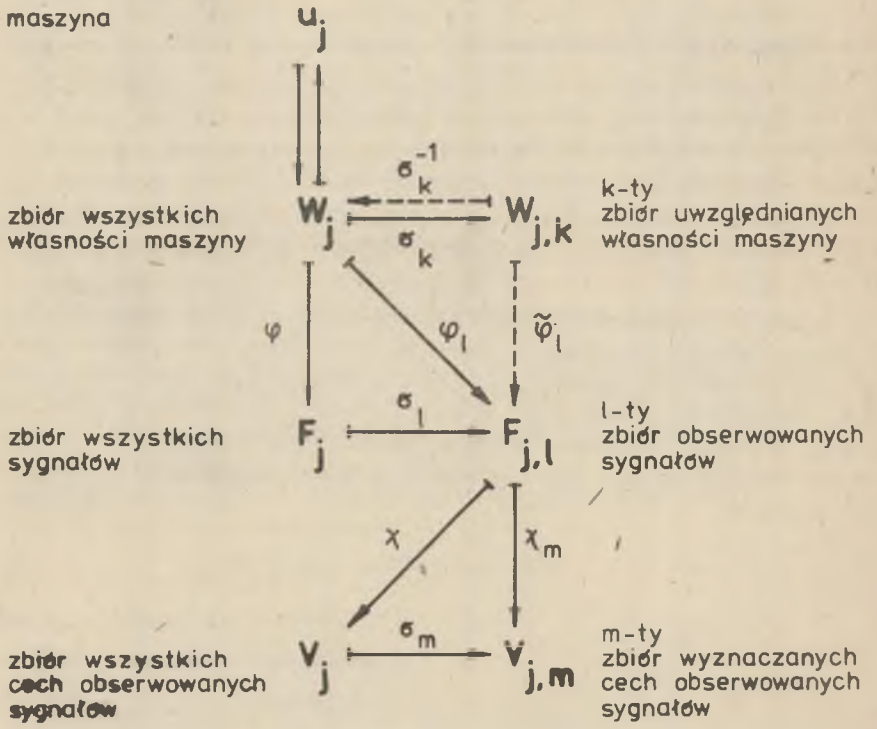
Innym sposobem uwzględniania tego faktu jest przyjęcie założenia, że obserwowany zbiór sygnałów może być zapisany jako

$$F_{j,l} = \tilde{\varphi}_l(W_{j,k}), \tag{2.3.9}$$

gdzie:

$\tilde{\varphi}_l$ - operator wieloznaczny.

Niejednoznaczność operatora $\tilde{\varphi}_l$ wynika z wieloznaczności przekształcenia σ_k^{-1} (por. rys. 2.1).



Rys. 2.1. Diagram przekształcenia zbioru uwzględnianych własności maszyny w zbiór wyznaczanych cech obserwowanych sygnałów

⟶ przekształcenie jednoznaczne,

⇢ przekształcenie wieloznaczne

2.4. Cechy

Dla opisania sygnału $f(\tau,t)$ wyznacza się wartości jego cech. Cechami sygnału rozpatrywanego jako proces stochastyczny są oceny (estymaty) tego procesu. Zakłada się, że cecha sygnału $f(\tau,t)$ jest funkcją $v(t)$ wyłącznie czasu "makro" t i nie zależy od czasu "mikro" τ. Cechy sygnałów jednowymiarowych nazywane są cechami własnymi tych sygnałów. Cechy sygnałów wielowymiarowych nazywane są cechami wzajemnymi. Rodzaje uwzględnianych cech sygnałów oraz metody wyznaczania ich wartości opisano wyczerpująco w dużej liczbie publikacji, np. [20, 95, 10, 71, 38, 43, 92, 51,97].

W ogólnym przypadku wartościami cech mogą być liczby (estymaty punktowe) lub funkcje (estymaty funkcyjne). Zakłada się, że rozpatrywane będą dalej wyłącznie cechy punktowe, co nie wprowadza ograniczeń zbioru możliwych do uwzględniania wartości cech, ponieważ wszystkie cechy funkcyjne mogą być (po ich dyskretyzacji) zastępowane ciągami cech punktowych. Przyjęto, że wartości cech będących elementami takiego ciągu zapisywane są w postaci macierzy (macierzy wartości cech). Wymaga to, w przypadku wartości nie określonych ilościowo (wartości jakościowe), przyjęcia kodu umożliwiającego zapisywanie ich w postaci liczb, np.:"jasny kolor oleju" = 1, "ciemny kolor oleju" = 3.

Podczas badań j-tej maszyny wyznaczane są wartości cech będących elementami skończonego m-tego podzbioru $V_{j,m}$ zbioru V_j wszystkich cech opisujących zbiór $F_{j,1}$ sygnałów.

$$V_{j,m} \triangleq \{v(t)\} \subset V_j. \tag{2.4.1}$$

Moc zbioru $V_{j,m}$ może być przedmiotem optymalizacji (por.rozdz.2.5).

Oznaczając złożenie przekształceń (por. rys. 2.1)

$$\chi_m \triangleq \sigma_m \circ \chi \tag{2.4.2}$$

gdzie:

σ_m - przekształcenie rzutowe

$$\sigma_m : V_j \longmapsto V_{j,m} \subset V_j, \tag{2.4.3}$$

χ - operator wyznaczania cech sygnałów

$$\chi : F_{j,1} \longmapsto V_j, \tag{2.4.4}$$

można zbiór wyznaczanych cech sygnałów zapisać

$$V_{j,m} = \chi_m(F_{j,1}). \tag{2.4.5}$$

Zbiór wyznaczanych cech $V_{j,m}$ jest interpretowany jako zbiór cech j-tej maszyny, dla której uwzględniany zbiór własności $W_{j,k}$ obserwowany jest za pośrednictwem zbioru sygnałów $F_{j,l}$ (por. rys. 2.1).

Należy odróżniać zbiór cech od zbioru wartości cech. Zbiór cech $V_{j,m}$ jest zbiorem funkcji $v(t)$. Zbiór wartości cech j-tej maszyny w chwili t jest zbiorem wartości funkcji $v(t) \in V_{j,m}$, które zapisywane są w postaci macierzy (o różnej liczbie elementów). Dla uproszczenia zapisu zbioru wartości cech przyjmuje się uporządkowanie elementów zbioru cech $V_{j,m}$ wyznaczające uporządkowanie ich wartości. Umożliwia to zapisywanie zbioru wartości cech j-tej maszyny w chwili t w postaci jednej macierzy $v_j^m(t)$. Macierz $v_j^m(t)$ interpretowana jest jako zbiór współrzędnych punktu w m-wymiarowej przestrzeni wartości cech V^m.

Zbiory wartości cech $v_j^m(t)$ są wyznaczane dla różnych maszyn u_j (będących elementami zbioru U badanych maszyn) oraz dla różnych chwil czasu. Tworzą one rodzinę V_u zbiorów wartości cech w przestrzeni wartości cech V

$$V_u = \left\{ v_j^m(t) \right\}_{j,t} \subset V^m. \qquad (2.4.6)$$

Przyjęto interpretację rodziny V_u zbiorów wartości cech $v_j^m(t)$, wyznaczanych dla maszyn $u_j \in U$ w różnych chwilach czasu, jako rodziny zbiorów wartości cech dla jednej abstrakcyjnej maszyny, zastępującej zbiór badanych maszyn (U), którego elementy (maszyny u_j) są równoważne ze względu na ich konstrukcję.

2.5. Przekształcanie cech

Wyznaczany podczas badań zbiór wartości cech orzekający o obserwowanej maszynie może być zbiorem bardzo licznym - zwłaszcza wtedy, gdy do wyznaczania cech sygnałów stosowane są programowane układy pomiarowe. Istnieje trudne do rozwiązania zadanie wyboru cech [58], które mogą być pominięte w dalszych rozważaniach (zadanie ograniczania liczby wymiarów przestrzeni wartości cech V^m).

Podejmowanie decyzji o ograniczeniu zbioru cech wymaga dużej ostrożności, ponieważ podczas sformalizowanego postępowania zmierzającego do ograniczenia liczby cech, uwzględniane są wyłącznie te związki pomiędzy cechami, które występują dla rozpatrywanej rodziny wartości cech (będącej wynikiem badań). Wybór ograniczonego zbioru cech może ulec zmianie po zmianie tej rodziny.

Wybór uwzględnianych cech może być dokonany arbitralnie przez osobę prowadzącą badania. Jawnie lub domyślnie przyjmowanym kryterium wyboru jest uwzględnianie wyłącznie tych cech, które są niezależne. Skuteczność takiego wyboru zależy głównie od wiedzy i intuicji badacza. Optymalne prze-

prowadzenie arbitralnego wyboru jest trudne, a nawet niewykonalne, jeżeli początkowa liczba cech jest duża.

Prosta sformalizowana metoda ograniczania liczby uwzględnianych cech polega na przekształcaniu rodziny wartości cech V_u^m (określonej w przestrzeni V^m) w rodzinę wartości cech głównych G_u^m (określoną w przestrzeni wartości cech głównych G^m) i następnie na określaniu ograniczonej przestrzeni wartości cech głównych G^h (np. [35, 74, 20])

$$V^m \xrightarrow{\mathfrak{A}} G^m \xrightarrow{\sigma_h} G^h. \qquad (2.5.1)$$

Przestrzeń wartości cech głównych G^m charakteryzuje się tym, że jej osie są osiami głównymi rozpatrywanego zbioru punktów określonego rodziną V_u wartości cech abstrakcyjnej maszyny. Przekształcenie $\mathfrak{A}$ może być wyznaczone na podstawie analizy macierzy kowariancji współrzędnych elementów rodziny w przestrzeni V^m. Celem analizy jest wyznaczenie wartości własnych i wektorów własnych tej macierzy. Wektory własne określają kierunki osi przestrzeni cech głównych czyli wyznaczają przekształcenie $\mathfrak{A}$. Ze zbioru wyznaczonych m wektorów własnych wybiera się $h \leq m$ wektorów własnych odpowiadających największym wartościom własnym. Wektory te określają osie ograniczonej przestrzeni cech głównych G^h. Liczbę wybranych wektorów własnych ustala się arbitralnie lub określa się ją na podstawie wyznaczonych wartości własnych wykorzystując kryterium

$$\lambda_h/\lambda_1 \geq \beta, \qquad (2.5.2)$$

gdzie:

λ_1, λ_h - pierwsza (największa uwzględniana) i h-ta (najmniejsza uwzględniana) wartość własna w uporządkowanym ciągu nierosnących wartości własnych,

β - stała przyjmowana arbitralnie; na podstawie badań autora, zaleca się

$$0{,}1 \geq \beta \geq 0{,}01. \qquad (2.5.3)$$

Dla ograniczenia liczby uwzględnianych cech można przeprowadzić ich grupowanie w klasy cech podobnych, będące pewnym uogólnieniem metody analizy kanonicznej [52].

Dla wyznaczonych klas określa się następnie cechy mające znaczenie reprezentantów klas. Ograniczonym zbiorem uwzględnianych cech jest zbiór reprezentantów klas. Zaletą takiego postępowania jest to, że wybrane cechy są elementami wyznaczanego podczas badań zbioru cech, a nie liniowymi kombinacjami tych cech. Ułatwia to interpretację znaczenia ograniczonego

zbioru cech. Wymienione zalety są przyczyną zalecania takiego sposobu postępowania (szczególnie podczas wstępnego ograniczania zbioru cech).

Opisane sposoby ograniczania liczby uwzględnianych cech charakteryzują się tym, że o wyborze ograniczonego zbioru cech decyduje rozkład rozpatrywanych elementów (rodzina V_u) w przestrzeni wartości cech. Dla potrzeb badań diagnostycznych ważne jest rozwiązanie zadania ograniczania liczby cech (liczby wymiarów m) przestrzeni wartości cech V^m, w której określono zbiór elementów V_u^m oraz rodzinę $\{\tilde{A}\} = \{\tilde{A}_k : k \in [1:l]\}$ zbiorów rozmytych $\tilde{A}_k = \{(v, a_k(v)) : v \in V_u\}$. (Pojęcie zbiorów rozmytych omówiono w dodatku D.1). Ograniczona przestrzeń wartości cech V^h ma umożliwiać określenie w niej "minimalnie zniekształconej" rodziny $\{\tilde{A}\}$ zbiorów rozmytych $\tilde{A}_i$. Zadanie jest trudne i autorowi nie są znane ogólne algorytmy prowadzące bezpośrednio do rozwiązania.

Dla rozwiązania zadania bada się różne ograniczone przestrzenie wartości cech V^h i ocenia ilościowo wrażliwość (czułość) $\varepsilon(V^h, \{\tilde{A}\})$ rozpoznawania (rozróżniania) w tych przestrzeniach elementów rodziny $\{\tilde{A}\}$ zbiorów rozmytych $\tilde{A}_k$. Proponuje się określanie wrażliwości $\varepsilon(V^h, \{\tilde{A}\})$ zgodnie z zależnością przyjętą na podstawie uogólnień metody analizy wariancji (opisanej w dodatku B.1)

$$\varepsilon(V^h, \{\tilde{A}\}) = \frac{tr(B)}{tr(T)} \in [0:1], \tag{2.5.4}$$

gdzie:

- sumaryczna macierz rozproszenia między l zbiorami $\tilde{A}_k$ rodziny $\{\tilde{A}\}$ w przestrzeni V^h

$$B^{h,h} = \sum_{k=1}^{l} \left(\sum_{j=1}^{k-1} T_{k,j}^{h,h} + \sum_{j=k+1}^{l} T_{k,j}^{h,h} \right) = T^{h,h} - R^{h,h}, \tag{2.5.5}$$

- sumaryczna macierz rozproszenia l zbiorów $\tilde{A}_k$ rodziny $\{\tilde{A}\}$ w przestrzeni V^h

$$R^{h,h} = \sum_{k=1}^{l} T_{k,k}^{h,h}, \tag{2.5.6}$$

- ogólna macierz rozproszenia l-elementowej rodziny $\{\tilde{A}\}$ zbiorów $\tilde{A}_k$ w przestrzeni V^h

$$T^{h,h} = \sum_{i=1}^{l} \sum_{j=1}^{l} T_{i,j}^{h,h}, \tag{2.5.7}$$

- cząstkowa macierz rozproszenia zbioru rozmytego

$\tilde{A}_i = \{(v, a_i(v)) : v \in V_u\}$ w zbiorze rozmytym

$\tilde{A}_j = \{(v, a_j(v)) : v \in V_u\}$ (określona dla zbioru V_u elementów przestrzeni V^h)

$$T^{h,h}_{i,j} = \frac{\sum_{v \in V_u} (v - \bar{v}_i)^T a_j(v)(v - \bar{v}_i)}{\sum_{v \in V_u} a_j(v)} \qquad (2.5.8)$$

(uwaga: w ogólnym przypadku $T_{i,j} \neq T_{j,i}$),

- reprezentant zbioru rozmytego $\tilde{A}_k$ (wg (D.5.7))

$$\bar{v}^h_k = \frac{\sum_{v \in V_u} v\, a_k(v)}{\sum_{v \in V_u} a_k(v)} \qquad (2.5.9)$$

Przyjmując zależność (2.5.8) założono, że ograniczona przestrzeń cech V^h ma umożliwić rozróżnianie zbiorów rozmytych $\tilde{A}_k$ niezależnie od ich mocy. Pomijając to założenie, tzn. w tych przypadkach, kiedy zbiory rozmyte o dużej mocy mają być "wyraźniej" rozróżniane w przestrzeni V^h niż zbiory rozmyte o małej mocy, należy w zależności (2.5.8) pominąć mianownik (mianownik = 1). Optymalizację, zmierzającą do ustalenia ograniczonej przestrzeni wartości cech, dla której wartości wrażliwości $\varepsilon(V^h_u, \{\tilde{A}\})$ są największe, można rozpocząć od badania jednowymiarowych przestrzeni cech w celu wyeliminowania z dalszych rozważań tych cech, które charakteryzują się wrażliwościami mniejszymi od wrażliwości wymaganej, a którą można przyjąć równą wrażliwości $\varepsilon(V^m_u, \{\tilde{A}\})$ rozpoznawania elementów rodziny zbiorów rozmytych w przestrzeni wartości cech V^m (przed jej ograniczeniem).

Drugim zadaniem związanym z przekształcaniem przestrzeni V^m, w której określono rodzinę $\{\tilde{A}\}$ zbiorów rozmytych $\tilde{A}_k$ jest zadanie wymagające znalezienia macierzy $\Pi^{h,m}$ przekształcenia transformującego przestrzeń V^m w przestrzeń V^h (dla $h < m$) posiadającą ortogonalny układ współrzędnych

$$v^m_j \longmapsto v^h_j = v^m\, \Pi^T, \qquad (2.5.10)$$

taką, że

$$\varepsilon(V^h_u, \{\tilde{A}\}) \longrightarrow \max. \qquad (2.5.11)$$

Niech B i T oznaczają macierze wg (2.5.5) i (2.5.7) dla przestrzeni V^m oraz niech B_1 i T_1 oznaczają odpowiednie macierze wyznaczone dla przestrzeni V^h

$$B_1^{h,h} = \Pi B^{m,m} \Pi^T, \tag{2.5.12}$$

$$T_1^{h,h} = \Pi T^{m,m} \Pi^T. \tag{2.5.13}$$

Pozwala to (po uwzględnieniu (2.5.4)) na zapisanie kryterium (2.5.11) w postaci

$$\varepsilon(V_u^h, \{\tilde{A}\}) = \frac{tr(\Pi B^{m,m} \Pi^T)}{tr(\Pi T^{m,m} \Pi^T)} \rightarrow \max. \tag{2.5.14}$$

Wiersze macierzy Π optymalnej ze względu na kryterium (2.5.14) są wektorami własnymi g_i^h odpowiadającymi kolejnym h największym wartościom własnym λ_i wyznaczanym jako rozwiązanie równania

$$g_i^h B^{h,h} - \lambda_i g_i^h T^{h,h} = 0^h, \tag{2.5.15}$$

po przyjęciu (dla uniknięcia nieoznaczoności równania (2.5.15), że wektory g_i są unormowane

$$g_i g_i^T = 1. \tag{2.5.16}$$

Niedogodnością tak wyznaczanego rozwiązania jest to, że w ogólnym przypadku wektory własne g_i nie są ortogonalne, co ze względu na (2.5.10) prowadzi do nieortogonalnego układu współrzędnych w przestrzeni V^h. Dla otrzymania macierzy Π o ortogonalnych wierszach można postępować wg następującego algorytmu:

- rozwiązując równanie charakterystyczne

$$\det(B - \lambda_1 T) = 0 \tag{2.5.17}$$

wyznacza się największą wartość własną λ_1,

- rozwiązując równanie

$$g_1(B - \lambda_1 T) = 0 \tag{2.5.18}$$

wyznacza się pierwszy wektor główny,

- dalsze wektory główne wyznacza się iteracyjnie metodą mnożników Lagrange'a, zakładając dla j-tego wyznaczanego wektora ε_j, że jest on ortogonalny względem wcześniej wyznaczonych wektorów

$$\forall_{i \in [1:j-1]} \quad \varepsilon_j \varepsilon_i^T = 0; \tag{2.5.19}$$

odpowiednie lagrangeany przyjmują postać

$$L = \frac{\varepsilon_j B \varepsilon_j^T}{\varepsilon_j T \varepsilon_j^T} - \varepsilon_j \sum_{i=1}^{j-1} \lambda_i \varepsilon_i^T. \tag{2.5.20}$$

2.6. Obrazy

Ciekawe wyniki w zakresie ograniczania mocy zbioru uwzględnianych cech można otrzymać, wykorzystując wprowadzoną przez L.A. Zadeha koncepcję zbiorów rozmytych [103, 55]. Pojęcie zbiorów rozmytych omówiono w dodatku D.1.

Istotą proponowanego sposobu postępowania jest określanie położenia elementu w przestrzeni wartości cech nie bezpośrednio za pomocą jego współrzędnych, lecz pośrednio za pomocą podobieństw tego elementu do wyróżnionych klas elementów w tej przestrzeni. Wartości takich podobieństw mogą być interpretowane jako współrzędne elementu w nowej przestrzeni, ktora nazywana będzie przestrzenią obrazów. Zaletą takiego postępowania jest znaczne ograniczenie liczby wymiarów przestrzeni przy zachowaniu możliwości skutecznego rozróżniania punktów.

W celu wyznaczenia rozmytych klas elementów podobnych w przestrzeni wartości cech, ustala się rodzinę reprezentantów klas. Reprezentantem klasy może być element lub zbiór elementów przestrzeni V. Rodzinę reprezentantów ustala się arbitralnie lub dla jej wyznaczenia przeprowadza się podział (grupowanie) elementów v rozpatrywanego zbioru V_u.

Rozwiązanie zadania grupowania polega na znalezieniu takiego rozłącznego podziału zbioru V_u na klasy elementów podobnych (równoważnych), który spełnia przyjęte kryterium jakości grupowania. Liczba wyznaczanych klas jest najczęściej przyjmowana arbitralnie.

Określenie relacji równoważności oraz wybór kryterium jakości grupowania, a zwłaszcza ich formalizacja, są trudne (por. dodatek B).

Dla wyróżnienia klas elementów podobnych konieczne jest porównywanie tych elementow, co wymaga wprowadzenia miar podobieństw (por. dodatek A). Przyjmuje się, że wartości podobieństw są liczbami rzeczywistymi zawartymi w przedziale $[0,0 : 1,0]$ i oznaczają:

1,0 - pełne podobieństwo,
0,0 - brak podobieństwa.

Po rozwiązaniu zadania grupowania elementów otrzymuje się (por. dodatek C) rodzinę reprezentantów klas $\{q(V_k)\}$.

Po wyznaczeniu rodziny reprezentantów (w wyniku rozwiązania zadania grupowania lub w wyniku arbitralnego ich ustalenia) przeprowadza się klasyfikację punktów przestrzeni cech, polegającą na ocenianiu podobieństw tych punktów do wyznaczonych reprezentantów. Wartości tak wyznaczonych podobieństw mogą być interpretowane jako wartości funkcji przynależności elementów do wyznaczonych klas, co prowadzi do określenia klas w postaci zbiorów rozmytych.

Rozmyta klasa $[\![q(V_k)]\!]_{\sim}$ elementu $q(V_k)$ (będącego reprezentantem klasy) określona w przestrzeni V jest zbiorem uporządkowanych par elementów $v \in V$ i wartości funkcji ich podobieństwa $p(v, q(V_k))$ do reprezentanta klasy interpretowanej jako wartość funkcji charakterystycznej zbioru rozmytego, czyli (por. (D.1.1))

$$[q(V_k)]_{\sim} = \{(v,\ p(v, q(V_k))) : v \in V\}. \qquad (2.6.1)$$

Wartości funkcji charakterystycznej $p(v,q(V_k))$ określone są dla wszystkich elementów $v \in V$. Każdy element $v \in V$ określany jest za pomocą uporządkowanego zbioru wartości cech. Dla ograniczenia liczby cech opisujących elementy v proponuje się określać je za pomocą ich podobieństw do wyróżnionych klas elementów podobnych. Oznacza to, że dla każdego elementu v wyznaczany jest zbiór rozmyty

$$\tilde{P}_v = \{(x,\ p(v, x)) : x \in \{q(V_k)\}\}. \qquad (2.6.2)$$

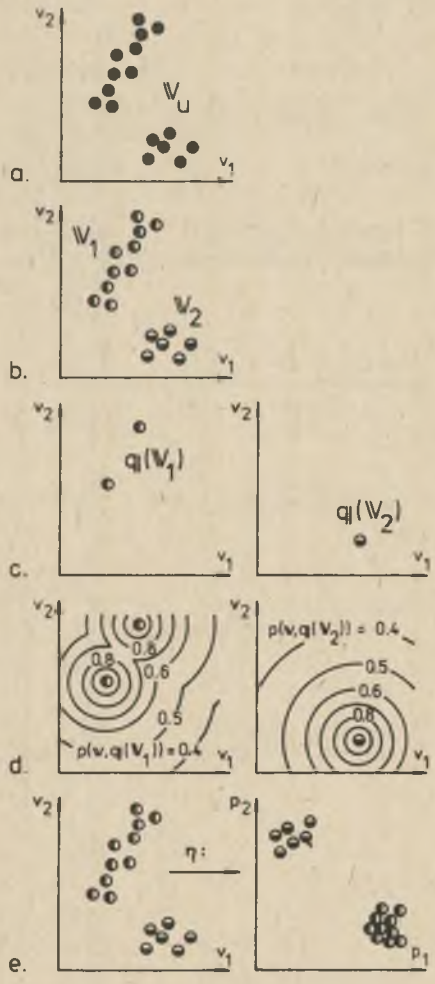
Wartości funkcji charakterystycznej zbioru rozmytego $\tilde{P}_v$ zapisywane są w postaci macierzy charakterystycznej

$$p = [p(v, q(V_1)), \ldots, p(v, q(V_l))]. \qquad (2.6.3)$$

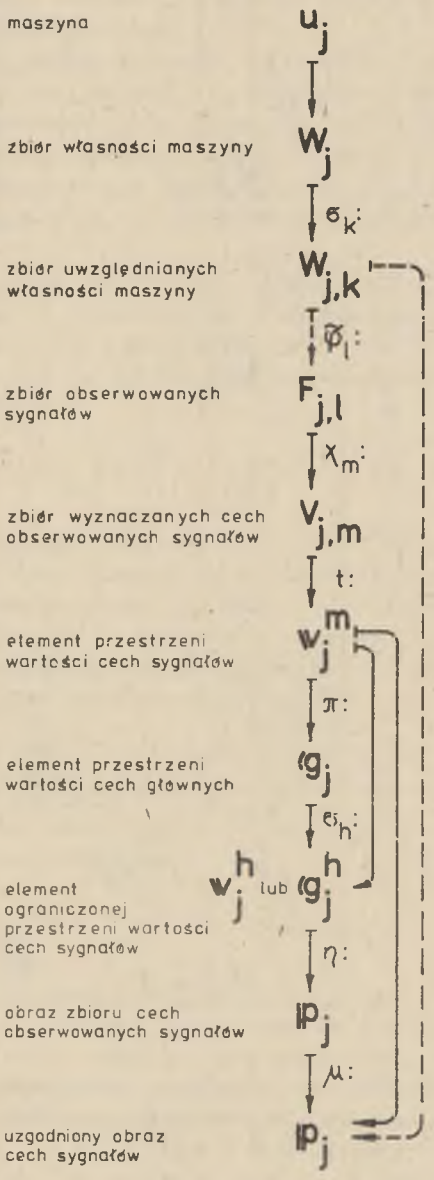
Przyjęto, że tak wyznaczony zbiór rozmyty $\tilde{P}_v$ jest nazywany obrazem elementu v na rodzinie rozmytych klas elementów podobnych. Wartości elementów macierzy p interpretowane są jako współrzędne wyznaczające punkt p w przestrzeni P nazywanej przestrzenią obrazów cech. Punkt p przestrzeni obrazów P jest obrazem punktu v przestrzeni wartości cech V

$$V \xrightarrow{?} P. \qquad (2.6.4)$$

Szkic algorytmu wyznaczania przestrzeni obrazów cech pokazano na rys.2.2. Element v przestrzeni cech jest wyznaczony przez wartości cech v(t),



Rys. 2.2. Uproszczony szkic algorytmu wyznaczania przestrzeni obrazów a) zbiór V_u^2 wartości cech obserwowanych sygnałów określony w dwuwymiarowej przestrzeni wartości cech (cechy: v_1, v_2), b) wyznaczenie rozłącznego podziału zbioru V_u na dwa podzbiory V_1 i V_2, c) rodzina reprezentantów klas $\{q(V_1), q(V_2)\}$, d) wyznaczenie wartości funkcji przynależności punktów przestrzeni wartości cech do klas V_1 i V_2, e) klasy punktów przestrzeni cech i zbiory ich obrazów



Rys. 2.3. Diagram przekształceń k-tego zbioru uwzględnianych własności j-tej maszyny w punkt p przestrzeni obrazów cech sygnałów

które są ocenami zbioru obserwowanych sygnałów $f(\tau,t)$. Punkt przestrzeni obrazów może być interpretowany jako obraz zbioru obserwowanych sygnałów (por. rys. 2.3).

Funkcje charakterystyczne zbiorów rozmytych mogą być określone w różny sposób (por. dodatek A i dodatek D). Dla zmniejszenia wpływu sposobu wyznaczania tych funkcji (na wartość współrzędnych obrazów) przeprowadza się (w co najmniej dwuwymiarowych przestrzeniach obrazów) przekształcenie nazwane uzgadnianiem obrazów. Podstawą uzgadniania obrazów jest założenie, że rozpatrywane rozmyte klasy $[x_k]$ rodziny $\{[x_k]\}$ mają charakter dychotomiczny (w znaczeniu rozmytym), czyli że dopełnieniem (wg (D.1.4)) sumy klas (wg (D.2.2)) rodziny $\{[x_k]\}$ jest iloczyn tych klas (wg (D.2.3)). Uzgadnianie 1-wymiarowych obrazów elementu v można zapisać jako przekształcenie

$$\mu : p^1[k] := \frac{p^1[k]}{\min\limits_{k\in[1:1]} p^1[k] + \max\limits_{k\in[1:1]} p^1[k]} \quad \forall v \in V. \qquad (2.6.5)$$

3. MODEL DIAGNOSTYCZNY MASZYNY

Podstawowe zadania diagnostyki maszyn związane są z metodami wnioskowania o relacjach między cechami sygnałów diagnostycznych i cechami stanu maszyny. Ogólne określenia pojęcia "stan" przyjmują, że stan maszyny jest zbiorem własności przysługujących jej w danej chwili, koniecznych i wystarczających do zidentyfikowania jej w tej chwili. Wykorzystując wprowadzone pojęcie cech można rozumieć cechy stanu jako cechy sygnałów uwarunkowanych wyłącznie tymi własnościami. Nie należy utożsamiać cech stanu z cechami sygnałów diagnostycznych, ponieważ sygnały te uwarunkowane są między innymi (a nie wyłącznie) stanem maszyny.

Procesy wytwarzania, kontroli i eksploatacji posiadają własności losowe powodujące, że cechy konstrukcyjne maszyn rozpatrywane są jako wielkości losowe o ściśle zadanych rozkładach lub dopuszczalnych granicach (tolerancjach), przyjętych ze względu na wymagany skutek (działanie maszyny), por. [33, 23]. Te własności losowe oraz złożoność maszyny powodują, że bezpośrednie przewidywanie relacji między cechami sygnałów i cechami stanu maszyny jest trudne. Relacje te wyznaczane są (po przyjęciu założeń upraszczających) na podstawie modelu maszyny. Najczęściej wykorzystywanymi modelami (por. np. [20, 80, 69]) są modele strukturalne odwzorowujące postać i wzajemne oddziaływania elementów maszyny. Modele takie umożliwiają skuteczne wnioskowanie o rodzaju wielkości fizycznych, których zmiany powinny być obserwowane podczas badań jako sygnały diagnostyczne (tzn. sygnały zależne od stanu obiektu), o rodzaju cech tych sygnałów, które prawdopodobnie są wrażliwe na zmiany stanu obiektu, o zalecanej lokalizacji czujników (przetworników) pomiarowych oraz o wymaganych warunkach działania obiektu podczas badań. Bezpośrednim skutkiem uproszczeń jest często brak zgodności szczegółowych wniosków dotyczących relacji między cechami sygnałów diagnostycznych i cechami stanu obiektu, wyznaczonych na podstawie analizy modeli, z wnioskami co do relacji wyznaczonych na podstawie eksperymentów.

Odmiennym sposobem postępowania jest wyznaczanie diagnostycznego modelu maszyny na podstawie relacji między ocenami stanu i ocenami sygnałów diagnostycznych wyznaczanymi podczas eksperymentów. Znane są sposoby takiego modelowania obiektów badań (w technice, medycynie, socjologii, ...) stosujące:

- statystykę matematyczną [6, 21, 68, 14, 9] i analizę regresji [34],
- metody adaptacyjne (teoria układów uczących się, teoria aproksymacji) [27, 36, 13] oraz teorię optymalnych decyzji [1, 89, 46, 14] i teorię informacji [91, 81, 20, 13],

- metody automatycznego wnioskowania wykorzystujące opracowany przez F. Rosenblatta model działania mózgu (perceptron) [77], metodę funkcji potencjalnych [5, 14] oraz lingwistykę teoretyczną [98] i teorię rozpoznawania obrazów [35, 73, 44, 96],
- teorię automatów skończonych i metody algebraiczne [16, 13], teorię grafów [101, 50, 64, 68],
- metody wnioskowania przyczynowego (causal inference) [81] oraz metody analizy czynnikowej i kanonicznej (korelacja kanoniczna) [18, 52, 74],
- teorię zbiorów i relacji rozmytych [37, 87, 86, 17, 70].

Dla ograniczenia czasu wymaganego do identyfikacji modelu obiektu badań można stosować postępowanie sekwencyjne [53, 83], podczas którego zakres badań jest optymalizowany w czasie ich realizacji. Pewnym uogólnieniem wymienionych metod jest ogólna teoria systemów [65]. Dla rozwiązania zadania wyznaczania relacji między cechami sygnałów diagnostycznych i cechami stanu maszyny, proponuje się wykorzystać systemowy model maszyny [93, 61, 2, 3], którego istotą jest traktowanie maszyny jako "czarnej skrzynki" [7] bez wnikania w jej strukturę. Maszyna będąca "czarną skrzynką" rozpatrywana jest jako układ względnie odosobniony, tzn. układ wyodrębniony myślowo z otoczenia i spełniający następujące warunki:

- oddziaływanie otoczenia na układ odbywa się wyłącznie za pośrednictwem wejść układu,
- oddziaływanie układu na otoczenie odbywa się wyłącznie za pośrednictwem wyjść układu,
- zbiór wejść układu i zbiór wyjść układu są zbiorami rozłącznymi.

3.1. Wejście, wyjście modelu maszyny

Definicja 3.1.1: Wejście (wyjście) maszyny jest zbiorem oddziaływań otoczenia na maszynę (maszyny na otoczenie).

Wejście i wyjście j-tej maszyny oznaczane są odpowiednio $W_{x,j}$, $W_{y,j}$. Elementami wejścia maszyny mogą być w rozumieniu def. 3.1.1 wszystkie oddziaływania otoczenia na maszynę, obejmując łącznie (jako równoważne) oddziaływania rozróżniane przez większość autorów (por. np. [20]) jako zasilanie i sterowanie. Elementem wyjścia maszyny może być w rozumieniu def. 3.1.1 każde jej oddziaływanie na otoczenie, bez względu na to czy oddziaływanie to jest celem działania maszyny, czy też jest niezamierzonym skutkiem jej działania (np. hałas, drgania).

Przyjęcie założenia, że maszyna rozpatrywana jest jako układ względnie odosobniony powoduje, że wszystkie oddziaływania między maszyną i otoczeniem interpretowane są jako wejście lub wyjście maszyny i pozwala na pominięcie w dalszych rozważaniach oddziaływań interpretowanych jako zakłócenia. Rezygnacja z uwzględniania zakłóceń jest jedną z przyczyn (nie

jedyną) rozpatrywania działania modelu maszyny jako modelu rozmytego (por. rozdz. 3.4).

Wejście i wyjście maszyny mogą być obserwowane podczas badań za pośrednictwem sygnałów. Dla sygnałów tych można wyznaczać cechy interpretowane jako cechy wejścia i wyjścia. Dla ograniczenia mocy zbiorów cech wyznacza się ich obrazy. Obrazy zbiorów wartości cech nazywane są odpowiednio obrazami wejścia i wyjścia maszyny, rozumianymi jako wejście i wyjście modelu maszyny.

Definicja 3.1.2: Wejściem (wyjściem) modelu maszyny jest obraz zbioru wyznaczanych wartości cech sygnałów, za pośrednictwem których obserwowane jest wejście (wyjście) maszyny. Wejście (wyjście) modelu maszyny interpretowane jest jako obraz wejścia (wyjścia) maszyny. Wartości wejścia i wyjścia modelu j-tej maszyny, oznaczane odpowiednio $p_{x,j}$, $p_{y,j}$, rozpatrywane są jako elementy przestrzeni $P_{x,j}$ wartości wejść oraz przestrzeni $P_{y,j}$ wartości wyjść tego modelu. Zakłada się, że przestrzenie obrazów wejść (wyjść) badanych maszyn $u_j \in U$ są wyznaczane po przyjęciu wspólnej dla nich rodziny wyróżnionych klas rozmytych w przestrzeni cech sygnałów (por. rozdz. 2.6).

3.2. System

Dla opisywania związków między wejściem i wyjściem maszyny wykorzystuje się pojęcie systemu. Przyjmując definicję systemu zgodną z ogólną teorią systemów [65], system maszyny definiuje się jako relację określoną na iloczynie kartezjańskim przestrzeni wejść i wyjść tej maszyny. Praktyczne wykorzystanie pojęcia systemu maszyny w prowadzonych badaniach jest trudne, ze względu na brak możliwości bezpośredniego wyznaczania jej wejścia i wyjścia. Uwzględniając możliwość rozpatrywania obrazów wejścia i wyjścia maszyny jako wejścia i wyjścia modelu maszyny, wprowadza się pojęcie systemu modelu maszyny.

Definicja 3.2.1: System modelu j-tej maszyny $u_j \in U$ jest relacją R_j określoną na iloczynie kartezjańskim przestrzeni $P_{x,j}$ wartości wejść i przestrzeni $P_{y,j}$ wartości wyjść modelu tej maszyny

$$R_j \subset P_{x,j} \times P_{y,j}. \qquad (3.2.1)$$

Dla opisywania zbioru maszyn wykonanych według wspólnej konstrukcji wprowadza się pojęcie systemu modelu klasy maszyn.

Definicja 3.2.2: System modelu zbioru U maszyn wykonanych według wspólnej konstrukcji jest relacją R określoną na iloczynie kartezjańskim przestrzeni P_x wartości wejść i przestrzeni P_y wartości wyjść modelu zbioru maszyn

$$R \subset P_x \times P_y, \tag{3.2.2}$$

gdzie przestrzeń wejść (wyjść) modelu zbioru maszyn zawiera przestrzenie wejść (wyjść) modeli maszyn $u_j \in U$

$$P_x = \bigcup_j P_{x,j}; \qquad P_y = \bigcup_j P_{y,j}, \tag{3.2.3}$$

z czego wynika, że:

$$R = \bigcup_j R_j. \tag{3.2.4}$$

Należy zwrócić uwagę na różnice pojęć "system R_j modelu maszyny" i "system R modelu zbioru maszyn". System R_j określony jest dla j-tej maszyny u_j (dla rzeczywistego obiektu badań). System R określony jest dla zbioru U badanych maszyn wykonanych według wspólnej konstrukcji, tzn. dla klasy maszyn równoważnych ze względu na konstrukcję (a nie dla j-tej maszyny). Oznacza to, że system modelu zbioru maszyn jest systemem modelu abstrakcyjnej maszyny zastępującej zbiór badanych maszyn. Można badać hipotezę, że abstrakcyjna maszyna zastępuje zbiór wszystkich maszyn wykonanych według rozpatrywanej konstrukcji i że model zbioru maszyn jest modelem klasy maszyn równoważnych ze względu na konstrukcję.

3.3. Obraz stanu maszyny

Ogólnie przyjmuje się (np. [13]), że stan maszyny (a właściwie stan modelu maszyny [104]) opisany jest zbiorem chwilowych wartości określających ją tzw. "parametrów". W większości prac autorzy przyjmują, że zbiór tych "parametrów" jest zbiorem wyłącznie takich elementów, które są "ważne" dla rozwiązywanego zadania (i występują wyraźnie w matematycznym opisie modelu maszyny [104]). Często (np. [20]) rozpatruje się wyłącznie te "parametry", które warunkują poprawność działania lub spełnienie norm (luzy, niewyrównoważenie, przebieg kilometrowy).

Wprowadzone pojęcie "system modelu maszyny" umożliwia formalne określenie pojęcia "stan modelu maszyny", które nie odpowiada definicji wprowadzonej przez B.W. Pawłowa stwierdzającej, że: "stan mechanizmu jest pełnym, minimalnym zbiorem parametrów charakteryzujących odchylenie struktury mechanizmu od struktury jego idealnego wzorca" [76 str. 6].

Definicja 3.3.1: Stan $p_{s,j}(t)$ modelu j-tej maszyny w chwili czasu t jest takim elementem przestrzeni $P_{s,j}$ stanów modelu tej maszyny, że istnieje funkcja Q nazywana funkcją odpowiedzi systemu

$$Q: (p_{x,j}([t), p_{s,j}(t)) \longmapsto p_{y,j}([t), \tag{3.3.1}$$

spełniająca warunek

$$\overset{\forall}{(p_{x,j}([t),\ p_{y,j}([t)) \subset R_j} \qquad \overset{\exists}{p_{s,j}(t) \in P_{s,j}}$$

$$[p_{y,j}([t) = \varrho(p_{x,j}([t),\ p_{s,j}(t))], \tag{3.3.2}$$

gdzie:

- wykres wejścia modelu maszyny na odcinku czasu od chwili t

$$p_{x,j}([t) \triangleq \left\{ p_{x,j}(t'):\ t' \geqslant t \right\}, \tag{3.3.3}$$

- wykres wyjścia modelu maszyny na odcinku czasu od chwili t

$$p_{y,j}([t) \triangleq \left\{ p_{y,j}(t'):\ t' \geqslant t \right\}. \tag{3.3.4}$$

Stan $p_{s,j}(t)$ modelu j-tej maszyny interpretowany jest jako obraz stanu j-tej maszyny (por. def. 3.1.2).

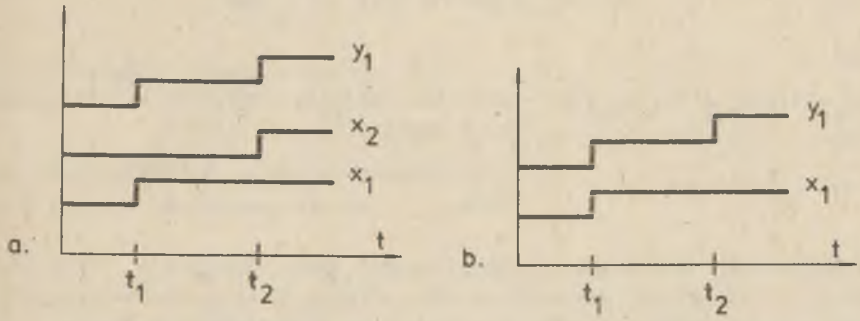
Z definicji 3.3.1 wynika, że stan $p_{s,j}(t)$ modelu maszyny w chwili t oraz wykres wejścia $p_{x,j}([t)$ tego modelu na odcinku czasu od chwili t (zbiór obrazów cech sygnałów, za pośrednictwem których obserwowane są oddziaływania otoczenia na maszynę, od tej chwili) wyznaczają wykres wyjścia $p_{y,j}([t)$ tego modelu (zbiór obrazów cech sygnałów, za pośrednictwem których obserwowane są oddziaływania maszyny na otoczenie) od tej chwili.

Z wprowadzonej definicji wynika, że przestrzeń stanów modelu maszyny oraz funkcja odpowiedzi nie są wyznaczalne jednoznacznie. Należy podkreślić, że są one ze sobą związane w tym sensie, iż przyjęcie przestrzeni stanów wyznacza funkcję odpowiedzi systemu wtedy, gdy spełniony jest warunek (3.3.2). Stan modelu maszyny, rozumiany zgodnie z przyjętą definicją, jest zbiorem niezbędnych informacji o zmianach maszyny, w zakresie, w jakim zmiany te warunkują wyjście modelu maszyny. Oznacza to, że tak definiowany stan obejmuje między innymi to, co w niektórych pracach z zakresu cybernetyki i diagnostyki maszyn określa się jako wejście wewnętrzne maszyny por. (np. [60]).

Rozpatrując maszynę jako układ cybernetyczny, przyjmuje się intuicyjnie usasadniony postulat tzw. determinizmu lokalnego zakładający, że wyjście maszyny w określonej chwili czasu jest wyznaczone przez równoczesne i wcześniejsze jej wejście i nie zależy od późniejszego jej wejścia. Z definicji stanu modelu maszyny wynika (zgodnie z tym postulatem), że stan modelu maszyny w chwili t nie zależy od późniejszych wejść, czyli jest określony przez wcześniejsze wejścia modelu maszyny oraz przez stan początkowy.

Dla ułatwienia wprowadzenia (w rozdz. 3.4) pojęcia "działanie modelu maszyny" przyjmujemy chwilowo założenie, iż system modelu maszyny jest systemem przyczynowo-skutkowym.

Przyjmowanie założenia o właściwościach przyczynowo-skutkowych systemu modelu maszyny wymaga ostrożności ze względu na różnice między systemem obiektu rzeczywistego, dla którego uwzględnia się wszystkie możliwe (a więc praktycznie nieskończenie wiele) elementy wejścia, wyjścia i stanu - a systemem modelu, dla którego uwzględnia się skończoną liczbę elementów wejścia, wyjścia i stanu. Uzasadnienie tego stwierdzenia pokazano na rysunku 3.1.



Rys. 3.1. Wykresy obserwowanych elementów wejścia X i wyjścia Y modelu maszyny o niezmieniającym się stanie, którego system

a) może być uznany jako system przyczynowo-skutkowy (zmiany x_1 w chwili t_1 oraz x_2 w chwili t_2 są przyczynami zmian y_1 w tych chwilach),

b) nie może być uznany jako system przyczynowo-skutkowy (nie jest znana przyczyna zmian y_1 w chwili t_2)

Definicja 3.3.2: System modelu maszyny jest systemem przyczynowo-skutkowym od chwili t_0, jeżeli istnieje funkcja odpowiedzi tego systemu spełniająca warunek

$$\overset{\forall}{T'} = [t : t']; \; t \geqslant t_0 \quad \overset{\forall}{p_{s,j}}(t) \quad \overset{\forall}{p_{x,j}}(T') \quad \overset{\forall}{p'_{x,j}}(T')$$

$$[p_{x,j}(T') = p'_{x,j}(T') \Rightarrow$$

$$\Rightarrow \varrho(p_{x,j}(T'), p_{s,j}(t)) = \varrho(p'_{x,j}(T'), p_{s,j}(t))]. \qquad (3.3.5)$$

Jeżeli system modelu maszyny jest systemem przyczynowo-skutkowym, to stan modelu maszyny w chwili t zawiera informacje o wcześniejszych wejściach modelu maszyny, niezbędne do określenia wyjścia maszyny w chwili t.

3.4. Działanie modelu

Definicja 3.4.1: Zewnętrzne działanie modelu j-tej maszyny u_j jest funkcją

$$z_j : (p_{x,j}(t), p_{s,j}(t)) \longmapsto p_{y,j}(t), \qquad (3.4.1)$$

czyli

$$z_j : P_{x,j} \times P_{s,j} \longrightarrow P_{y,j}, \qquad (3.4.2)$$

gdzie:

$p_{x,j}(t)$, $p_{y,j}(t)$, $p_{s,j}(t)$ - wejście, wyjście, stan modelu j-tej maszyny w chwili t,

$P_{x,j}$, $P_{y,j}$, $P_{s,j}$ - przestrzenie wejścia, wyjścia, stanu modelu.

Stosując schemat dowodu przeprowadzonego przez Mesarovića [65] można wykazać, że zewnętrzne działanie modelu maszyny istnieje jako funkcja dla każdego modelu, którego system jest systemem przyczynowo-skutkowym.

Definicja 3.4.2: Wewnętrzne działanie modelu j-tej maszyny na odcinku czasu $[t : t+\Delta)$ jest funkcją

$$w_{\Delta,j} : (p_{x,j}([t : t+\Delta)), p_{s,j}(t)) \longmapsto p_{s,j}(t+\Delta). \qquad (3.4.3)$$

Zewnętrzne działanie modelu maszyny umożliwia określenie wyjścia modelu (obrazu wyjścia maszyny) jako skutku wejścia i stanu modelu (obrazu wejścia i obrazu stanu maszyny). Wewnętrzne działanie modelu maszyny umożliwia określenie zmienionego stanu modelu w chwili $t+\Delta$ jako skutku stanu modelu w chwili t i wykresu wejścia na odcinku czasu $[t:t+\Delta)$.

Zewnętrzne działanie modelu maszyny może być wykorzystywane w badaniach diagnostycznych, umożliwiając wnioskowanie o prawdopodobnym obrazie stanu maszyny na podstawie obrazu wejścia i obrazu wyjścia maszyny. Będzie ono przedmiotem dalszych rozważań.

Wewnętrzne działanie modelu maszyny może być wykorzystywane w badaniach prognostycznych, umożliwiając wnioskowanie o prawdopodobnych zmianach obrazu stanu maszyny na podstawie obrazu stanu początkowego oraz wykresów obrazu wejścia maszyny. Działanie to nie będzie rozpatrywane szczegółowo.

Model j-tej maszyny można zapisać w postaci piątki uporządkowanej

$$u_j := (P_{x,j}, P_{s,j}, P_{y,j}, z_j, w_{\Delta,j}). \qquad (3.4.4)$$

Analizując sposób wyznaczania modelu maszyny należy stwierdzić,że brak podstaw do przyjmowania założenia, że system modelu maszyny jest systemem przyczynowo-skutkowym. Rezygnacja z założenia, że jest to system przyczynowo-skutkowy powoduje, że funkcja "zewnętrzne działanie modelu" z_j może nie istnieć. Proponuje się zastąpić tę funkcję relacją rozmytą $\tilde{z}_j$ (por. dodatek D.1) nazywaną "rozmyte zewnętrzne działanie modelu maszyny".

Definicja 3.4.3: Rozmyte zewnętrzne działanie modelu maszyny u_j jest relacją rozmytą $\tilde{z}_j$

$$(p_{x,j}(t),\ p_{s,j}(t))\ \tilde{z}_j(p_{y,j}(t)). \tag{3.4.5}$$

Relacja ta jest zbiorem rozmytym $\tilde{z}_j$

$$\tilde{z}_j \stackrel{d}{=} \left\{(p,\ z_j(p)) : p \in P = P_{x,j} \times P_{s,j} \times P_{y,j}\right\}, \tag{3.4.6}$$

gdzie funkcja charakterystyczna

$$z_j : (p_{x,j},\ p_{s,j},\ p_{y,j}) \mapsto r \in [0 : 1]. \tag{3.4.7}$$

Wprowadzone rozmyte zewnętrzne działanie modelu maszyny umożliwia określanie rozmytego wyjścia $\tilde{P}_{y,j}(t)$ modelu w chwili t, uwarunkowanego wejściem $p_{x,j}(t)$ i stanem $p_{s,j}(t)$ w tej chwili

$$\tilde{P}_{y,j}(t) = \left\{(p,\ z_j(p_{x,j}(t),\ p_{s,j}(t),\ p)) : p \in P_{y,j}\right\}. \tag{3.4.8}$$

Bezpośrednie wykorzystanie w praktyce tak określonego działania modelu jest kłopotliwe ze względu na potrzebę wyznaczania funkcji charakterystycznej z_j. Jej określenie w postaci analitycznej jest trudne. Funkcję z_j można określać przez określenie jej wykresu w postaci zbioru par (ARGUMENT, WARTOŚĆ), Postępowanie takie prowadzi do dużej mocy zbioru, którego elementy są parami określającymi wartość funkcji charakterystycznej. Na przykład, dla prostego przypadku, gdy rozpatruje się tylko czterowymiarowe przestrzenie obrazów P_x^4, P_s^4, P_y^4 po dyskretyzacji przekształcającej te przestrzenie w kraty umożliwiające rozróżnianie wyłącznie trzech różnych wartości współrzędnych (p = 0, p = 0,5, p = 1), do zapisania wartości funkcji charakterystycznej wymagany jest zbiór (tablica) o liczbie elementów

$$n = 3^4\ \ 3^4\ \ 3^4 = 3^{12} = 531441. \tag{3.4.9}$$

Podane w przykładzie liczba wymiarów przestrzeni obrazów oraz liczba rozróżnianych wartości współrzędnych są małe w porównaniu z potrzebami występującymi w praktyce. Obliczona duża wymagana moc zbioru wartości funkcji charakterystycznej stanowić może podstawę do stwierdzenia, że bezpośrednie wykorzystanie tej funkcji w praktyce jest niecelowe.

Dla uniknięcia opisanych niedogodności wprowadza się drugą definicję rozmytego zewnętrznego działania modelu maszyny, wykorzystującą relacyjne równanie rozmyte. W celu wprowadzenia tej definicji należy zwrócić uwagę na możliwość interpretowania zbioru współrzędnych punktu p w przestrzeni obrazów P jako zbioru wartości funkcji charakterystycznej zbioru rozmytego $\tilde{P}$

$$\tilde{P} = \left\{(k,\ p(k)) : k \in K\right\}, \tag{3.4.10}$$

gdzie:

K - zbiór klas wyznaczających przestrzeń obrazów (zbiór osi przestrzeni P).

Definicja 3.4.4: Rozmyte zewnętrzne działanie modelu maszyny u_j jest parą relacji rozmytych $(\tilde{Z}_{min,j},\ \tilde{Z}_{max,j})$ takich, że:

$$(\tilde{P}_{x,j}(t),\ \tilde{P}_{s,j}(t))\ \circ\ \tilde{Z}_{min,j} \subset \tilde{P}_{y,j}(t), \tag{3.4.11}$$

$$\tilde{P}_{y,j}(t) \subset (\tilde{P}_{x,j}(t),\ \tilde{P}_{s,j}(t))\ \Delta\ \tilde{Z}_{max,j}, \tag{3.4.12}$$

gdzie:

- $\tilde{A} \subset \tilde{B}$ oznacza zbiór rozmyty $\tilde{A}$ zawarty w zbiorze rozmytym $\tilde{B}$ (por. (D. 1.7)),
- $\circ$, Δ oznaczają operatory działań na zbiorach rozmytych (por. (D.2.14) i (D.2.15)).

Tak definiowane rozmyte zewnętrzne działanie modelu maszyny może być wyznaczane jako rozwiązanie odpowiedniego układu relacyjnych równań rozmytych (por. dodatek D.4 oraz (D. 3.5))

$$\tilde{Z}_{min,j} = (\tilde{P}_{x,j},\ \tilde{P}_{s,j})^T \propto \tilde{P}_{y,j}, \tag{3.4.13}$$

$$\tilde{Z}_{max,j} = (\tilde{P}_{x,j},\ \tilde{P}_{s,j})^T \in \tilde{P}_{y,j}. \tag{3.4.14}$$

Pomijając w dalszych rozważaniach wewnętrzne działanie modelu maszyny, tzn. rozpatrując maszynę wyłącznie w ustalonej chwili czasu (a nie na odcinku czasu), można zapisywać model maszyny u_j w postaci czwórki uporządkowanej

$$u_j := (P_{x,j},\ P_{s,j},\ P_{y,j},\ (\tilde{Z}_{min,j},\ \tilde{Z}_{max,j})), \qquad (3.4.15)$$

Dla opisywania zbioru maszyn wykonanych wg wspólnej konstrukcji wprowadza się pojęcie rozmytego działania modelu zbioru maszyn. Rozpatrywane będzie wyłącznie rozmyte zewnętrzne działanie modelu zbioru maszyn.

Definicja 3.4.5: Rozmyte zewnętrzne działanie modelu zbioru U maszyn u_j jest parą relacji rozmytych $(\tilde{Z}_{min},\ \tilde{Z}_{max})$ takich, że

$$\tilde{Z}_{min} = \bigwedge_j \tilde{Z}_{min,j}, \qquad \tilde{Z}_{max} = \bigvee_j \tilde{Z}_{max,j} \qquad (3.4.16)$$

gdzie: $(\tilde{Z}_{min,j},\ \tilde{Z}_{max,j})$ wg (3.4.13) i (3.4.14).

Model zbioru U maszyn wykonanych wg wspólnej konstrukcji można zapisać w postaci czwórki uporządkowanej

$$U := (P_x,\ P_s,\ P_y,\ (\tilde{Z}_{min},\ \tilde{Z}_{max})). \qquad (3.4.17)$$

4. KONCEPCJA UKŁADU DIAGNOZUJĄCEGO

Celem postępowania diagnostycznego jest ocenianie stanu maszyn. Stan maszyny może być oceniany "bezpośrednio" na podstawie badań jej elementów i / lub badań ich współdziałania oraz na podstawie arbitralnych opinii "ekspertów". Podstawową niedogodnością, związaną z ocenianiem stanu maszyny na podstawie badań jej elementów i ich współdziałania jest w większości przypadków konieczność demontażu maszyny oraz konieczność adaptacji elementów do badań. Adaptacja elementów (umożliwiająca badania) powoduje często zmianę warunków ich współdziałania. Z kolei wykorzystywanie opinii "ekspertów" związane jest najczęściej z trudnościami formalizowania sposobów ich wyznaczania. Wymienione sposoby postępowania mogą być stosowane w warunkach laboratoryjnych dla mało licznych zbiorów maszyn. W badaniach diagnostycznych maszyn będących elementami zbioru maszyn wykonanych wg wspólnej konstrukcji istnieje możliwość wykorzystywania metod "pośrednich" polegających na tym, że cechy stanu wyznaczane są na podstawie obserwacji sygnałów związanych z działaniem maszyny bez bezpośredniego oceniania stanu. Wyznaczenie tych ocen i wnioskowanie na ich podstawie o prawdopodobnym stanie maszyny realizowane jest przez tzw. układy diagnozujące, przekształcające cechy wejścia i wyjścia maszyny w obrazy cech jej stanu. Ograniczenia wynikające z możliwości realizacji układu diagnozującego w praktyce zmuszają do minimalizowania liczby uwzględnianych cech wejścia, wyjścia i stanu maszyny. Wykorzystując wnioski wynikające z rozważań dotyczących możliwości minimalizacji liczby wymiarów przestrzeni cech (por. rozdz. 2.5, 2.6) podjęto decyzję, że rozpatrywane będą układy diagnozujące uwzględniające oceny wejścia, wyjścia i stanu maszyny w postaci ich obrazów. Oznacza to, że układ diagnozujący określony dla zbioru U, maszyn wykonanych według wspólnej dla nich konstrukcji, wyznacza obraz stanu p_s wynikający z obrazu wejścia p_x i obrazu wyjścia p_y badanej maszyny. Tak wyznaczony obraz stanu jest (por. rozdz. 2.6) zbiorem wartości podobieństw stanu badanej maszyny do wyróżnionych klas stanów, określającym rozmyty zbiór $\tilde{P}_s$ klas stanów. Zbiór ten stanowi podstawę do wnioskowania o prawdopodobnej klasie stanu maszyny.

4.1. Działanie układu diagnozującego

Dla określenia poszukiwanego układu diagnozującego, rozpoznającego stan (wyznaczającego prawdopodobne klasy stanu) na podstawie cech oddziaływań otoczenia na maszynę i maszyny na otoczenie wykorzystane zostanie wprowa-

dzone (3.4.16) rozmyte zewnętrzne działanie modelu zbioru U maszyn u_j wykonanych wg wspólnej konstrukcji, umożliwiające zapisanie relacji między rozmytymi zbiorami klas wejścia, wyjścia i stanu maszyny w postaci pary równań rozmytych

$$(\tilde{P}_x(t), \tilde{P}_s(t)) \circ \tilde{Z}_{min} \subset \tilde{P}_y(t), \tag{4.1.1}$$

$$(\tilde{P}_x(t), \tilde{P}_s(t)) \triangle \tilde{Z}_{max} \supset \tilde{P}_y(t). \tag{4.1.2}$$

Przekształcanie tych równań w celu określenia relacji umożliwiającej wyznaczanie ograniczeń rozmytego zbioru klas stanów $\tilde{P}_s$, na podstawie rozmytych zbiorów $\tilde{P}_x$ i $\tilde{P}_y$, prowadzi do dużego rozmycia rozwiązania (wg (D.3.9)). Dla uniknięcia wynikających stąd niedogodności przyjęto, że układ diagnozujący będzie określany bezpośrednio na podstawie jego systemu wyznaczonego jako relacja R_d na iloczynie kartezjańskim zbiorów wartości wejścia X_d i wyjścia Y_d tego układu

$$R_d \subset X_d \times Y_d. \tag{4.1.3}$$

Pomijając złożone i nie rozpatrywane w tej pracy przypadki adaptacyjnych (uczących się) układów diagnozujących, zakłada się, że stan układu diagnozującego nie zmienia się w funkcji czasu, co oznacza, że działanie tego układu (działanie zewnętrzne) (por. (3.4.1)) może być opisane funkcją

$$d : X_d \rightarrow Y_d. \tag{4.1.4}$$

Uwzględniając przyjęte zadanie badań diagnostycznych, którym jest wyznaczanie obrazu stanu maszyny, założono, że wyjściem Y_d układu diagnozującego jest obraz p_s stanu badanej maszyny, natomiast wejściem X_d jest element iloczynu kartezjańskiego przestrzeni P_x obrazów wejścia i przestrzeni P_y obrazów wyjścia maszyny. Przestrzeń $P_w = P_x \times P_y$ nazywana jest przestrzenią obrazów sygnałów diagnostycznych.

Funkcję d wg (4.1.4) można zapisać

$$d : p_w \mapsto p_s. \tag{4.1.5}$$

Ze względu na przyjęte wejście układu diagnozującego, podział zbioru uwzględnianych cech obserwowanych oddziaływań (zbioru cech sygnałów diagnostycznych), na zbiór cech wejścia i zbiór cech wyjścia maszyny nie wpływa na wyznaczane działanie układu diagnozującego.

Uzasadniony w rozdz. 3.4 brak podstaw do zakładania, iż system modelu maszyny jest systemem przyczynowo-skutkowym powoduje, że brak jest również

podstaw do zakładania, iż system układu diagnozującego jest systemem przyczynowo-skutkowym. Prowadzi to do rozpatrywania działania układu diagnozującego jako działania rozmytego (podobnie jak rozmyte zewnętrzne działanie modelu maszyny - por. definicja 3.4.4).

Podstawą wyznaczania działania układu diagnozującego jest zbiór tzw. "danych uczących", będących wynikami badań maszyn $u_j \in U$. Elementami tego zbioru są pary obrazów $(p_{w,j}, p_{s,j})$, gdzie: $p_{w,j}$ - obraz sygnałów diagnostycznych, $p_{s,j}$ - obraz stanu maszyny wyznaczone na podstawie j-tego badania. W badaniach, których celem jest zgromadzenie zbioru danych uczących, obraz stanu wyznaczany jest metodami bezpośrednimi.

Jedną z możliwości wyznaczenia funkcji d (wg (4.1.5)) jest rozwiązanie zadania regresji. Ograniczając się do regresji liniowej, wyznaczenie funkcji

$$d : p_{w,j} \mapsto p_{s,j} = p_{w,j}\, E, \tag{4.1.6}$$

polega na rozwiązaniu układu równań

$$p_{s,j}^{l_s} = p_{w,j}^{l'_w}\, E^{l'_w, l_s} + e_j^{l_s} \qquad \forall j \in [1:n], \tag{4.1.7}$$

gdzie:

$p_{s,j}$, $p_{w,j}$, e_j - macierze jednowierszowe,

$p_{w,j}^{l'_w}$ - macierz obrazu sygnału $p_{w,j}^{l_w}$ uzupełniona dodatkowym jednostkowym elementem zapisywanym jako ostatni element macierzy $p_{w,j}[l'_w] = 1$, $l'_w = l_w + 1$,

l_w, l_s - liczby klas wartości cech sygnałów diagnostycznych i stanów,

n - liczba równań w układzie równań (4.1.7),

E - macierz prostokątna określająca poszukiwaną funkcję d.

Rozwiązanie układu równań (4.1.7) wyznacza się (por. [62], [34]) po przyjęciu kryterium minimum sumy norm macierzy odchyleń e_j:

$$E = C^{-1}\, B, \tag{4.1.8}$$

gdzie

$$C^{l'_w, l'_w} = \sum_{j=1}^{n} (p_{w,j}^T\ p_{w,j}), \tag{4.1.9}$$

$$B^{l'_w, l_s} = \sum_{j=1}^{n} (p_{w,j}^T\ p_{s,j}). \tag{4.1.10}$$

Po wyznaczeniu macierzy E określa się odchylenia elementów macierzy e_j, wyznaczając w tym celu macierz kowariancji

$$C_e^{l_s,l_s} = \frac{1}{n - l'_w} \sum_{j=1}^{n} (e_j^T \; e_j). \tag{4.1.11}$$

Oznaczając

$$A^{l_s,l_s} = \sum_{j=1}^{n} (p_{s,j}^T \; p_{s,j}), \tag{4.1.12}$$

można (4.1.11), po uwzględnieniu (4.1.7), zapisać

$$C_e = \frac{1}{n - l'_w} (A - E^T B). \tag{4.1.13}$$

Macierz E będąca rozwiązaniem układu równań (4.1.7), określana jest dla dalszych potrzeb za pomocą pary macierzy (E_{min}, E_{max}), posiadających elementy:

$$E_{min}^{l'_w,l_s}[i,j] = \begin{cases} E[i,j] & \forall i < l'_w, \\ E[i,j] - \delta \sqrt{C_e[j,j]} & \forall i = l'_w, \end{cases} \tag{4.1.14}$$

$$E_{max}^{l'_w,l_s}[i,j] = \begin{cases} E[i,j] & \forall i < l'_w, \\ E[i,j] + \delta \sqrt{C_e[j,j]} & \forall i = l'_w. \end{cases} \tag{4.1.15}$$

Wartości liczby δ są wyznaczane w wyniku optymalizacji pary macierzy (E_{min}, E_{max}) ze względu na kryterium maksymalnej wartości wagi ν układu diagnozującego (wg (4.3.2)) lub mogą być przyjmowane arbitralnie (np. z przedziału $[1:2]$).

Za pomocą pary macierzy (E_{min}, E_{max}) określane są ograniczenia obrazu stanu j-tej maszyny, zgodnie z zależnościami

$$\begin{aligned} p_{s,min,j} &= p_{w,j} \; E_{min}, \\ p_{s,max,j} &= p_{w,j} \; E_{max}. \end{aligned} \tag{4.1.16}$$

Ograniczenia te umożliwiają wnioskowanie o prawdopodobnej klasie stanu j-tej maszyny (por. rozdz. 4.2).

Należy zauważyć, że opisywanie relacji między obrazami sygnałów diagnostycznych p_w i obrazami stanów p_s za pomocą związków liniowych nie prowadzi w ogólnym przypadku do liniowych zależności między wartościami cech sygnałów i wartościami cech stanu, co wynika z nieliniowości przekształcenia: PRZESTRZEŃ CECH ⟶ PRZESTRZEŃ OBRAZÓW.

Proponuje się drugi sposób opisywania działania układu diagnozującego wykorzystujący jego model analityczny w postaci układu równań rozmytych.

Definicja 4.1.1. Rozmyte działanie układu diagnozującego jest parą relacji rozmytych $(\tilde{D}_{min}, \tilde{D}_{max})$ takich, że

$$\tilde{P}_w(t) \circ \tilde{D}_{min} \subset \tilde{P}_s(t) \subset \tilde{P}_w(t) \triangle \tilde{D}_{max}, \qquad (4.1.17)$$

gdzie:

$\tilde{P}_w(t) = (\tilde{P}_x(t), \tilde{P}_y(t))$ - rozmyty zbiór klas sygnałów diagnostycznych,

$\tilde{P}_s(t)$ - rozmyty zbiór klas stanów maszyny.

Zgromadzone wyniki badań (dane uczące) są wykorzystywane do wyznaczania rozmytego działania układu diagnozującego $(\tilde{D}_{min}, \tilde{D}_{max})$ jako rozwiązania układów równań rozmytych (por. dodatek D.4):

$$p_{w,j}^{l_w} \circ D_{min}^{l_w,l_s} \subset p_{s,j}^{l_s} \qquad \forall j \in [1:n], \qquad (4.1.18)$$

$$p_{w,j}^{l_w} \triangle D_{max}^{l_w,l_s} \supset p_{s,j}^{l_s} \qquad \forall j \in [1:n]. \qquad (4.1.19)$$

Do wyznaczenia relacji $\tilde{D}_{min}$ i $\tilde{D}_{max}$ będących rozwiązaniami tych układów równań wykorzystuje się zależności (por. (D.4.2) i (D.4.4))

$$D_{min} = \bigwedge_{j \in [1:n]} (p_{w,j}^T \curlywedge p_{s,j}), \qquad (4.1.20)$$

$$D_{max} = \bigvee_{j \in [1:n]} (p_{w,j}^T \curlyvee p_{s,j}), \qquad (4.1.21)$$

gdzie:

$\curlywedge, \curlyvee$ - operatory działań na zbiorach rozmytych (por. def. D.2.1).

Tak wyznaczone rozwiązanie układu równań rozmytych cechuje się często dużym stopniem rozmycia (Ψ wg (D.3.9)). Dla zmniejszenia stopnia rozmycia wyznacza się ograniczone rozwiązanie układu równań rozmytych (wg dodatku

D.4.1). Należy zwrócić uwagę na to, że ograniczone rozwiązanie układu równań powinno być weryfikowane dla danych uwzględniających pełny układ równań (por. rozdz. 4.3).

Para relacji rozmytych $(\tilde{D}_{min}, \tilde{D}_{max})$ zapisywanych przy pomocy macierzy (D_{min}, D_{max}) umożliwia określenie ograniczeń obrazu stanu j-tej maszyny zgodnie z zależnościami

$$p_{s,min,j} = p_{w,j} \circ D_{min},$$
$$p_{s,max,j} = p_{w,j} \triangle D_{max}. \qquad (4.1.22)$$

Wyznaczone ograniczenia obrazu stanu umożliwiają wnioskowanie o prawdopodobnej klasie stanu j-tej maszyny (por. rozdz. 4.2).

4.2. Diagnozowanie

Wyznaczone przez parę macierzy (D_{min}, D_{max}) działanie układu diagnozującego umożliwia diagnozowanie, tzn. wnioskowanie o prawdopodobnej klasie stanów maszyny na podstawie obrazów cech sygnałów diagnostycznych p_w. W wyniku działania układu diagnozującego wyznacza się ograniczenia obrazów stanów wg (4.1.16) lub (4.1.22). Ograniczenia obrazów stanów wyznaczają rodzinę możliwych rozmytych zbiorów stanu badanej maszyny

$$\{\tilde{P}_s\} = \{\tilde{P}_s : \tilde{P}_s = \{(k, p_s[k]) : p_{s,min}[k] \leq p_s[k] \leq p_{s,max}[k]; k \in K_s\}\}, \qquad (4.2.1)$$

gdzie:

k - klasa stanu,

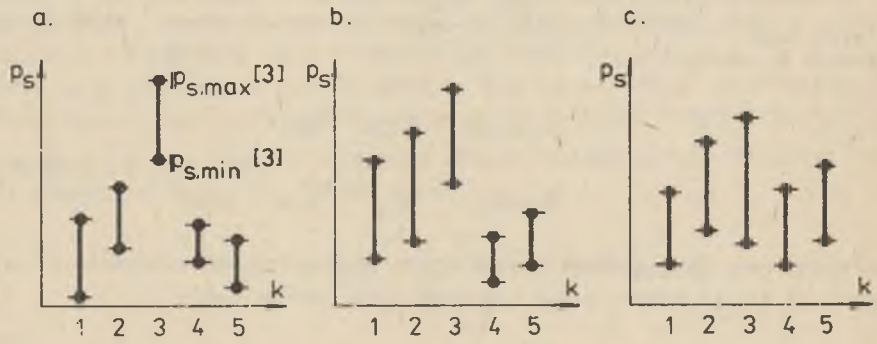
K_s - rodzina rozróżnianych klas stanów.

Wynikiem wnioskowania na podstawie rodziny możliwych rozmytych zbiorów stanu jest podział zbioru K_s rozróżnianych klas stanu na trzy zbiory (por. dodatek D.5, rys. 4.1):

- pusty lub co najwyżej jednoelementowy zbiór K_{s1} taki, że $k \in K_{s1}$ jest klasą stanu badanej maszyny,
- zbiór K_{s2} taki, że klasa stanu badanej maszyny nie jest elementem tego zbioru $(k \notin K_{s2})$,
- zbiór K_{s3} taki, że

$$K_{s3} = K_s \setminus (K_{s1} \cup K_{s2}) \qquad (4.2.2)$$

czyli zbiór klas, dla których na podstawie przeprowadzonych badań nie można orzec czy są one klasami stanu obiektu badań, czy też nie.



Rys. 4.1. Przykłady wyników badań diagnostycznych (ograniczenia obrazów stanów wyznaczone przez układ diagnozujący i wyniki wnioskowania o prawdopodobnej klasie stanów maszyny)

a) jednoznaczny wynik diagnozowania: "stan badanej maszyny należy do klasy 3", $K_{s1} = \{3\}$, $K_{s2} = \{1,2,4,5\}$, $K_{s3} = \phi$, waga diagnozy $\vartheta_o = 1$; b) wieloznaczny wynik diagnozowania: "stan badanej maszyny należy do klasy 1 lub 2 lub 3", $K_{s1} = \phi$, $K_{s2} = \{4,5\}$, $K_{s3} = \{1,2,3\}$, waga diagnozy $\vartheta_o = 0{,}4$; c) trywialny wynik diagnozowania: "stan badanej maszyny należy do klasy 1 lub 2 lub 3 lub 4 lub 5", $K_{s1} = \phi$, $K_{s2} = \phi$, $K_{s3} = \{1,2,3,4,5\}$, waga diagnozy $\vartheta_o = 0$

Zbiory K_{s1}, K_{s2} określa się zgodnie z zależnościami

$$K_{s1} = \left\{k \in K_s : \underset{i \in K_s \setminus \{k\}}{\forall} \left[p_{s,max}[k] > p_{s,min}[k] \geq p_{s,max}[i]\right]\right\} \qquad (4.2.3)$$

$$K_{s2} = \left\{k \in K_s : \underset{i \in K_s}{\exists} \left[p_{s,min}[k] \leq p_{s,max}[k] \leq p_{s,min}[i]\right]\right\} \qquad (4.2.4)$$

Jeżeli zbiór K_{s1} nie jest zbiorem pustym, to określa on jednoznaczny wynik badań diagnostycznych wyznaczając klasę $k \in K_{s1}$, do której należy stan badanej maszyny. Jeżeli zbiór K_{s1} jest zbiorem pustym, to zbiór K_{s3} określa te klasy stanu, wśród których znajduje się klasa stanu obiektu badań, czyli określa wieloznaczny ($\overline{K}_{s3}$ - znaczny) wynik badań diagnostycznych.

Dla porównania wyników działania układu diagnozującego wprowadza się wagę diagnozy (wagę wyniku badań diagnostycznych).

Definicja 4.2.1: Waga diagnozy wyznaczonej przez układ diagnozujący określana jest zgodnie z zależnością

$$\vartheta_0 = \frac{\overline{K}_{s1} + \overline{K}_{s2}}{\overline{K}_s}. \qquad (4.2.5)$$

Wartości tej wagi interpretowane są następująco (por. rys. 4.1):

$\vartheta_0 = 1$ - diagnoza jednoznaczna (jednoznaczne rozpoznanie klasy stanu),

$0 < \vartheta_0 < 1$ - diagnoza wieloznaczna,

$\vartheta_0 = 0$ - diagnoza trywialna (brak rozpoznania klasy stanu).

4.3. Weryfikacja układu diagnozującego

Obiektami badań weryfikacyjnych są maszyny $u_j \in U_1$. Zbiorem U_1 maszyn objętych tymi badaniami może być, w szczególnym przypadku, zbiór U maszyn, dla których wyznaczono "dane uczące". Dla każdej maszyny u_j wyznacza się obraz sygnałów diagnostycznych, a następnie, wykorzystując działanie weryfikowanego układu diagnozującego, określa się ograniczenia obrazów stanów i wnioskuje się o prawdopodobnej klasie stanu maszyny. Wyniki wnioskowania porównuje się z wynikami ocen stanu wyznaczonymi bezpośrednio, określając wagę $\vartheta(j)$ zweryfikowanej diagnozy dla maszyny u_j. Wagi $\vartheta(j)$ wyznaczone dla maszyn $u_j \in U_1$ stanowią podstawę do określania wagi ϑ_1 układu diagnozującego.

Definicja 4.3.1: Waga $\vartheta(j)$ zweryfikowanej diagnozy wyznaczonej dla maszyny u_j określana jest zgodnie z zależnością

$$\vartheta(j) \stackrel{d}{=} \begin{cases} \dfrac{1 + \vartheta_0(j)}{2} & \text{- jeżeli weryfikowany układ diagnozujący prawidłowo wyznaczył zbiory } K_{s1} \text{ i } K_{s2} \text{ klas stanów (wg (4.2.3) i (4.2.4) dla j-tej maszyny,} \\ \dfrac{1 - \vartheta_0(j)}{2} & \text{- jeżeli weryfikowany układ diagnozujący błędnie wyznaczył zbiór } K_{s1} \text{ lub } K_{s2} \text{ klas stanów (wg (4.2.3) i (4.2.4)) dla j-tej maszyny,} \end{cases} \qquad (4.3.1)$$

gdzie: $\vartheta_0(j)$ - waga diagnozy (wg (4.2.5)) dla j-tej maszyny.

Definicja 4.3.2: Waga ϑ_1 układu diagnozującego weryfikowanego na podstawie zbioru maszyn U_1 określana jest zgodnie z zależnością

$$\vartheta_1 \stackrel{d}{=} \frac{1}{\overline{U}_1} \sum_{u_j \in U_1} \vartheta(j); \qquad (4.3.2)$$

gdzie: $\vartheta(j)$ wg (4.3.1).

Wartości wagi modelu układu diagnozującego interpretowane są następująco:

$\vartheta_1 = 1$ - wszystkie wyniki diagnozowania maszyn $u_j \in U_1$ są jednoznaczne i poprawne,

$\vartheta_{gr} \leq \vartheta_1 \leq 1$ - część wyników diagnozowania maszyn $u_j \in U_1$ jest błędnych lub wieloznacznych, lecz dopuszcza się stosowanie weryfikowanego układu diagnozującego do wnioskowania o stanie maszyn,

$\vartheta_1 < \vartheta_{gr}$ - większość wyników diagnozowania maszyn $u_j \in U_1$ jest błędnych lub wieloznacznych i nie dopuszcza się stosowania weryfikowanego układu diagnozującego do wnioskowania o stanie maszyny,

gdzie graniczna wartość wagi ϑ_{gr} przyjmowana jest arbitralnie z zakresu (0,5: 1] - odpowiednio do "ważności" badanej maszyny.

Optymalizacja układu diagnozującego może być przeprowadzana ze względu na kryterium maksymalnej wartości wagi ϑ_1 (wg (4.3.2)) lub kryterium minimalnej wartości oceny rozmycia Ψ działania układu diagnozującego (wg (D.4.6)).

Przyczyną małej wartości wagi ϑ_1 lub dużego stopnia rozmycia Ψ może być niewłaściwie dobrany zbiór uwzględnianych cech sygnałów lub niewłaściwie dobrany sposób wyznaczania obrazów tych cech, lub zbyt małe zróżnicowanie stanów maszyn, których obrazy są elementami zbioru danych uczących.

Dobór zbioru uwzględnianych cech sygnałów i dobór sposobów wyznaczania obrazów może być przedmiotem optymalizacji prowadzonej przy wykorzystaniu metod opisanych w rozdz. 2.5 (przekształcanie cech). Optymalizacja taka jest głównym zabiegiem procesu projektowania (określania istoty działania) układu diagnozującego. Proces projektowania jest procesem twórczym, z czego wynika, że nie może on być w pełni sformalizowany i omawiana optymalizacja wymaga czynnego udziału projektanta układu diagnozującego. Istotnym elementem projektowania układu jest przygotowanie zbioru "danych uczących". Podczas określania programu badań, w wyniku których wyznaczone zostaną dane uczące, należy podjąć decyzję o sposobie uzyskiwania obiektów badań o odpowiednio zróżnicowanych stanach oraz decyzję o sposobie bezpośredniego oceniania tych stanów. Podejmując te decyzje można kierować się ogólnie znanymi relacjami między stanem i sygnałami diagnostycznymi (por. [54], [67], [20]) oraz wnioskami z wcześniejszych badań.

5. PRZYKŁAD ZASTOSOWANIA

Opisana metoda wyznaczania relacji między sygnałami diagnostycznymi i stanem maszyny jest szczególnie przydatna do określania działania układu diagnozującego będącego elementem stanowiska kontroli technicznej obiektów produkowanych seryjnie (silniki, przekładnie, łożyska toczne itp.). Prowadzenie badań, których celem jest wyznaczenie działania układu diagnozującego, wymaga współdziałania producenta lub użytkownika badanych obiektów. Autorowi nie udało się zainteresować takimi badaniami odpowiedniego producenta. W IMiPKM Politechniki Śląskiej podjęto pod kierunkiem autora próbę zastosowania opracowanej metody podczas badań sprężarek promieniowych oraz elementów maszyn włókienniczych. W wyniku przeprowadzonych dotychczas badań nie osiągnięto w pełni zamierzonego celu. Przyczyną nieosiągnięcia celu badań było zbyt małe zróżnicowanie stanów badanych obiektów. Fragmenty opisanej metody postępowania zostały zastosowane z powodzeniem w badaniach elektrycznych pieców łukowych [90].

Dla przeprowadzenia badań, wymagających wyznaczania licznych zbiorów wartości cech sygnałów i ich gromadzenia, konieczny był dostęp do minikomputera wyposażonego w wejście analogowe. Ze względu na brak możliwości zakupu lub ciągłego dostępu do takiego urządzenia zaprojektowano i wykonano zgodnie z założeniami systemu CAMAC układ cyfrowej rejestracji ocen sygnałów [25] i programowany analizator sygnałów PAS4 [24] oraz przygotowano system programów automatycznego wnioskowania SPAW [26].

Dla ilustracji proponowanej metody postępowania opisany zostanie prosty przykład, wykorzystujący wyniki badań prowadzonych przez M. Kurowicza [59]. Dla uproszczenia przykładu, celowo wybrano dane określające przestrzenie wartości cech sygnałów i stanów o małych wymiarach, które analizowane będą tak jakby były one przestrzeniami o dużych wymiarach. Obiektami badań były jednostopniowe walcowe przekładnie zębate [78], charakteryzujące się tym, że przez ułożyskowanie wału wolnoobrotowego w tulejach mimośrodowych umożliwiają one różne względne ustawianie zazębiających się elementów, co pozwala na regulację czynnej szerokości zazębienia (tzw. "śladu dolegania"). Dla badanych przekładni, uwzględniając ich warunki działania, wyznaczono wg metody opisanej szczegółowo w [59], liczby $(\Delta\phi)_{max}$ oceniające ilościowo względne skokowe zmiany obciążenia boku zęba wzdłuż odcinka przyporu. Badania prowadzono w układzie mocy zamkniętej, na stanowisku [45] składającym się z komory bezpogłosowej i układu napędowego wyposażonego między innymi w momentogenerator i tyrystorowy układ regulacji

prędkości obrotowej. Dla badanych zazębień wyznaczono odchyłki ich cech geometrycznych. Podczas badań rejestrowano sygnały przyspieszeń drgań korpusu przekładni, jej efekt akustyczny, prędkość obrotową i moment obciążający zębnik przekładni. Zarejestrowane sygnały drgań poddano analizie częstotliwościowej wykonywanej przy pomocy układu cyfrowej rejestracji częstotliwościowych widm sygnałów [25]. Na podstawie wyników analizy częstotliwościowej sygnałów zarejestrowanych dla czterech prędkości obrotowych zębnika (824; 773; 725; 680 obr/min) wyznaczono ich różnicowe widma reprezentatywne przy pomocy zbioru programów dla EMC ODRA 1305 [25], działających według algorytmu opisanego w [22]. Widma reprezentatywne przedstawiane w funkcji względnej częstotliwości (częstotliwość/częstość zazębiania) można interpretować jako widma umownego zastępczego sygnału, który jest niezależny (w pewnym zakresie) od prędkości obrotowej wałów przekładni. Ze względu na przyjęty [22] sposób określania tych widm są one wyznaczane w postaci widm różnicowych, których składowe stanowią niezmienniki odpowiedniego układu równań. Składowe te definiowane są następująco

$$\Delta S_i = S_i - \frac{1}{2}(S_{i-1} + S_{i+1}), \quad [dB], \qquad (5.0.1)$$

gdzie:

S_i - składowa widma reprezentatywnego w i-tym paśmie częstotliwości (o stałej względnej szerokości),

ΔS_i - składowa różnicowego widma reprezentatywnego w i-tym paśmie częstotliwości.

Wyniki badań zgromadzono w banku danych zorganizowanym w pamięci zewnętrznej EMC ODRA 1305 w sposób opisany w [25].

Szczegółowe dane dotyczące badanych przekładni, stanowiska i aparatury oraz zbiorcze wyniki badań zamieszczono w [59]. Jako "dane uczące" wykorzystano w opisywanym przykładzie wyniki badań zamieszczone w pracy [59] (str. 98), ograniczając się do wyników dotyczących przyspieszeń drgań korpusu przekładni zębatej w okolicach łożyska wału wolnoobrotowego (punkt pomiarowy nr 2 wg [59]). Wyniki te, dotyczące 26 obiektów, zestawiono w tablicy 5.1.

Przyjęto, że przestrzeń stanów wyznaczona jest dwoma cechami: s_1 i s_2 - przyjmującymi wartości równe standaryzowanym wartościom liczby $(\Delta\beta)_{max}$ oceniającej względną skokową zmianę obciążenia wzdłuż odcinka przyporu i momentu M obciążającego zębnik przekładni oraz, że przestrzeń sygnałów diagnostycznych wyznaczona jest sześcioma cechami: w_1 do w_6 - przyjmującymi wartości równe standaryzowanym wartościom składowych ΔS_1 do ΔS_6 różnicowego widma reprezentatywnego przyspieszeń drgań w pasmach częstotliwości od $1f_z$ do $6f_z$, gdzie f_z oznacza częstotliwość zazębiania się współdziałających uzębień. Celem przykładu jest wyznaczenie działania układu

Tablica 5.1

Wyniki badań jednostopniowych, walcowych przekładni zębatych wg [59] (uporządkowane według numerów obiektów)

Lp.	Obiekt	$(\Delta\beta)_{max}$ [-]	M [Nm]	Składowe różnicowego widma reprezentatywnego					
				ΔS_1 [dB]	ΔS_2 [dB]	ΔS_3 [dB]	ΔS_4 [dB]	ΔS_5 [dB]	ΔS_6 [dB]
1	2	3	4	5	6	7	8	9	10
1	110	0,86	100	2,6	4,7	2,9	3,7	-0,9	2,1
2	120	0,66	200	3,7	2,6	3,2	3,4	-0,2	2,9
3	130	0,59	300	2,9	1,1	4,3	4,0	0,8	3,7
4	155	0,53	550	2,1	0,6	5,0	4,1	2,6	2,0
5	210	0,86	100	2,5	4,0	2,2	4,0	-0,8	1,6
6	220	0,66	200	3,2	3,0	3,3	4,4	0,2	2,5
7	230	0,59	300	1,6	-0,8	4,6	3,9	1,3	3,1
8	255	0,53	550	2,4	0,7	5,3	3,5	3,3	3,7
9	410	1,00	100	3,0	3,5	2,5	3,0	0,6	1,5
10	420	0,73	200	2,1	3,2	2,8	4,1	-0,1	2,1
11	430	0,64	300	2,6	2,0	3,8	4,2	1,4	3,1
12	455	0,55	550	2,9	3,6	5,0	2,9	3,8	4,0
13	510	1,00	100	0,5	4,0	2,5	3,4	0,2	2,1
14	520	0,73	200	3,4	2,8	3,4	4,3	0,2	2,6
15	530	0,64	300	2,6	2,1	3,9	3,9	2,3	3,4
16	555	0,55	550	2,9	3,0	5,2	3,3	2,9	4,5
17	1530	0,52	300	2,8	3,1	4,3	3,7	1,6	1,6
18	1630	0,50	300	2,2	2,7	4,4	3,1	1,4	1,7
19	1730	0,54	300	1,6	3,3	4,3	3,2	1,1	3,8
20	1755	0,50	550	2,9	3,9	3,9	2,8	2,4	1,4
21	1830	0,54	300	1,6	3,3	3,9	3,9	2,7	2,8
22	1855	0,49	550	1,9	3,0	4,9	3,8	1,8	1,4
23	1930	0,55	300	0,6	1,5	3,9	3,7	2,5	2,8
24	1955	0,50	550	3,2	3,5	3,9	3,3	2,4	3,2
25	2030	0,52	300	1,1	2,5	3,5	3,4	2,6	2,1
26	2055	0,49	550	2,5	2,0	5,1	4,4	2,4	3,0

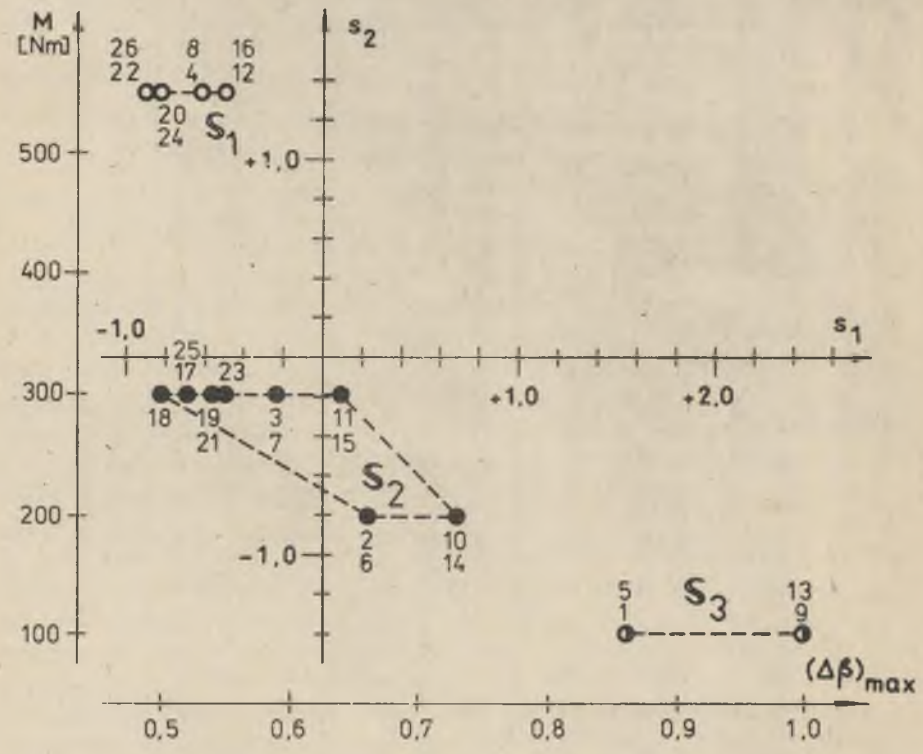
diagnozującego, umożliwiającego wnioskowanie na podstawie ocen sygnałów diagnostycznych o prawdopodobnej klasie tak określanego stanu przekładni.

Przykład obejmuje:

- wyznaczenie klas stanów,
- wyznaczenie obrazów stanów,
- minimalizację liczby wymiarów przestrzeni sygnałów,
- wyznaczenie obrazów sygnałów,
- wyznaczenie rozmytej relacji między obrazami sygnałów i obrazami stanu,
- weryfikację układu diagnozującego.

5.1. Wyznaczenie klas stanów

Przyjęto, że zbiór elementów przestrzeni wartości cech stanu o mocy $\bar{S} = 26$ będzie dzielony na $l_\delta = 3$ klasy. Zgodnie z zaleceniami (C.2.1) i (C.2.2), przyjmując liczbę $\delta = 3$, określono minimalną i maksymalną moc każdej klasy



Rys. 5.1. Elementy przestrzeni wartości cech stanu i wynik ich podziału na 3 klasy (numery elementów według tabl. 5.1 - kolumna 1)

$$\bar{\bar{s}}_{k,min} = \max\left(\frac{\bar{\bar{s}}}{\gamma^{1}_{s}};\ \gamma\right) = \max\left(\frac{26}{3 \cdot 3};\ 3\right) = 3, \qquad (5.1.1)$$

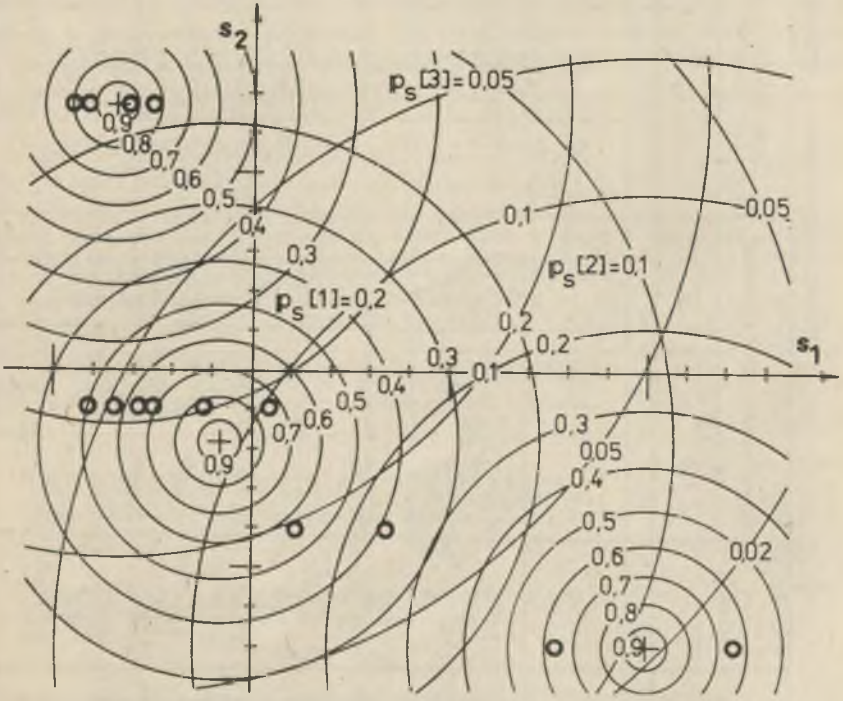
$$\bar{\bar{s}}_{k,max} = \min\left(\frac{\gamma\,\bar{\bar{s}}}{1_s};\ \bar{\bar{s}}\left(1 - \frac{1_s - 1}{\gamma^{1}_{s}}\right)\right) =$$

$$= \min\left(\frac{3 \cdot 26}{3};\ 26\left(1 - \frac{2}{3 \cdot 3}\right)\right) = 20. \qquad (5.1.2)$$

Podział przeprowadzono ze względu na kryterium najdalszego sąsiada (B. 3.1), przyjmując odległości elementów przestrzeni wartości cech stanu jako odległości euklidesowe (A.1.5) dla standaryzowanych wartości cech s_1 i s_2. Wyznaczono podział na klasy o mocach $\bar{\bar{s}}_1 = 8$, $\bar{\bar{s}}_2 = 14$, $\bar{\bar{s}}_3 = 4$. Elementy przestrzeni stanów oraz wyniki podziału pokazano na rys. 5.1.

5.2. Wyznaczenie obrazów stanów

Na podstawie wyników podziału przestrzeni wartości cech stanu wyznaczono rodzinę reprezentantów klas stanów wg kryterium (B.1.1)



Rys. 5.2. Linie stałych wartości podobieństw elementów przestrzeni wartości cech stanu do reprezentantów klas stanów

Tablica 5.2

Obrazy cech stanów i sygnałów diagnostycznych

Lp.	Obiekt	Uzgodniony obraz stanu			Uzgodniony obraz sygnałów diagnostycznych			$\varkappa$ wg (D.3.13)	Ograniczony, uzgodniony obraz stanu		
		$p_s[1]$	$p_s[2]$	$p_s[3]$	$p_w[1]$	$p_w[2]$	$p_w[3]$		$p_s[1]$	$p_s[2]$	$p_s[3]$
1	2	3	4	5	6	7	8	9	10	11	12
1	110	0,05	0,20	0,95	0,08	0,25	0,92	1,07	0,23	0,32	0,77
2	120	0,15	0,85	0,23	0,17	0,54	0,83	1,09	0,29	0,71	0,34
3	130	0,22	0,91	0,09	0,39	0,81	0,19	1,32	0,33	0,75	0,25
4	155	0,98	0,18	0,02	0,93	0,30	0,07	1,12	0,79	0,31	0,21
5	210	0,05	0,20	0,95	0,07	0,22	0,93	1,05	0,23	0,32	0,77
6	220	0,15	0,85	0,23	0,24	0,74	0,76	1,35	0,29	0,71	0,34
7	230	0,22	0,91	0,09	0,81	0,80	0,19	1,32	0,33	0,75	0,25
8	255	0,98	0,18	0,02	0,93	0,30	0,07	1,12	0,79	0,31	0,21
9	410	0,03	0,09	0,97	0,10	0,33	0,90	1,18	0,22	0,25	0,78
10	420	0,17	0,83	0,49	0,11	0,34	0,89	1,00	0,30	0,70	0,50
11	430	0,22	0,88	0,12	0,26	0,81	0,19	1,22	0,33	0,73	0,27
12	455	0,98	0,22	0,02	0,93	0,30	0,07	1,12	0,79	0,33	0,21
13	510	0,03	0,09	0,97	0,09	0,27	0,91	1,15	0,22	0,25	0,78
14	520	0,17	0,83	0,49	0,24	0,76	0,67	1,27	0,30	0,70	0,50
15	530	0,22	0,88	0,12	0,52	0,81	1,19	1,23	0,33	0,73	0,27
16	555	0,98	0,22	0,02	0,93	0,30	0,07	1,12	0,79	0,33	0,21
17	1530	0,34	0,91	0,09	0,62	0,81	0,19	1,32	0,40	0,75	0,25
18	1630	0,39	0,91	0,09	0,64	0,81	0,19	1,32	0,43	0,75	0,25
19	1730	0,31	0,92	0,08	0,45	0,81	0,19	1,36	0,39	0,75	0,25
20	1755	0,98	0,17	0,02	0,55	0,81	0,19	1,55	0,79	0,30	0,21
21	1830	0,31	0,92	0,08	0,66	0,81	0,19	1,36	0,39	0,75	0,25
22	1855	0,98	0,19	0,02	0,90	0,42	0,10	1,20	0,79	0,31	0,21
23	1930	0,32	0,91	0,09	0,59	0,81	0,19	1,32	0,39	0,75	0,25
24	1955	0,98	0,17	0,02	0,55	0,81	0,19	1,55	0,79	0,30	0,21
25	2030	0,34	0,91	0,09	0,34	0,80	0,20	1,37	0,40	0,75	0,25
26	2055	0,98	0,19	0,02	0,93	0,30	0,07	1,12	0,79	0,31	0,21

$$q(S_1) = [-0,691;\ 1,335],$$
$$q(S_2) = [-0,164;\ -0,366], \qquad (5.2.1)$$
$$q(S_3) = [2,007;\ -1,409].$$

Podobieństwa elementów przestrzeni wartości cech stanu wyznaczono wg (A.2.12) (przyjmując liczbę $\alpha = 1$). Na rys. 5.2 pokazano linie stałych wartości podobieństw elementów do klas stanów, wyznaczone zgodnie z (A.3.2). Wartości tych podobieństw są rozpatrywane dalej jako wartości funkcji charakterystycznych wyznaczających zgodnie z (2.6.1) trzy rozmyte klasy stanów. Na ich podstawie określono obrazy stanów badanych przekładni. Wartości współrzędnych punktów $p_{s,j}$ w przestrzeni obrazów, po ich uzgodnieniu wg (2.6.5), zestawiono w tablicy 5.2 (kolumny 3÷5).

5.3. Minimalizacja liczby wymiarów przestrzeni cech sygnałów diagnostycznych

W celu dokonania wyboru uwzględnianych dalej cech sygnałów diagnostycznych wyznaczono, dla sześciowymiarowej przestrzeni wartości cech oraz dla sześciu ograniczonych jednowymiarowych przestrzeni wartości cech (w_1 do w_6), wrażliwości ε rozpoznawania rozmytych klas stanów zgodnie z (2.5.4). Obliczenia przeprowadzono uwzględniając współrzędne uzgodnionych obrazów stanu, zestawione w tablicy 5.2 (kolumny 3÷5). Wyniki obliczeń zestawiono w tablicy 5.3. Pominięto następnie cechy w_1, w_2, w_4, w_6, dla których wy-

Tablica 5.3

Wrażliwość rozpoznawania rozmytych klas stanów w ograniczonych przestrzeniach wartości cech sygnałów diagnostycznych

Lp.	Uwzględniane cechy (ograniczona przestrzeń cech sygnałów diagnostycznych)	tr(B) wg (2.5.5)	tr(T) wg (2.5.7)	ε wg (2.5.4)
1	$\{w_1, w_2, w_3, w_4, w_5, w_6\}$	47.209	60.879	0,78
2	$\{w_1\}$	6.111	9.112	0,67
3	$\{w_2\}$	7.405	9.766	0,76
4	$\{w_3\}$	9.257	10.511	0,88
5	$\{w_4\}$	6.037	8.829	0,68
6	$\{w_5\}$	11.483	13.302	0,86
7	$\{w_6\}$	6.916	9.359	0,74

znaczone wrażliwości są mniejsze od wrażliwości $\varepsilon = 0{,}78$, wyznaczonej dla sześciowymiarowej przestrzeni wartości cech sygnałów diagnostycznych. Rozpatrywana dalej przestrzeń określona jest przez cechy w_3 i w_5, dla których $\varepsilon(w_3, \{s_k\}) = 0{,}88$ i $\varepsilon(w_5, \{s_k\}) = 0{,}86$. Elementy tej przestrzeni pokazano na rys. 5.3. Dla ilustracji znaczenia wrażliwości ε, na rys.5.4 pokazano elementy dwuwymiarowej przestrzeni określonej przez cechy w_1 i w_4, dla których (por. tabl. 5.3) wrażliwości ε przyjmują najmniejsze wartości. Dla sprawdzenia możliwości dalszego zmniejszenia liczby wymiarów przestrzeni wartości cech sygnałów (tj. przejścia do przestrzeni jednowymiarowej) wykorzystano kryterium (2.5.14). Zgodnie z (2.5.5) i (2.5.7) wyznaczono macierz $\mathbf{B}$ rozproszenia między rozmytymi klasami stanów i ogólną macierz $\mathbf{T}$ rozproszenia rozmytych klas stanów:

$$\mathbf{B} = \begin{bmatrix} 9{,}257; & 9{,}190 \\ 9{,}190; & 11{,}483 \end{bmatrix}, \qquad \mathbf{T} = \begin{bmatrix} 10{,}511; & 10{,}152 \\ 10{,}152; & 13{,}302 \end{bmatrix}. \tag{5.3.1}$$

Następnie rozwiązano równanie charakterystyczne (2.5.17), które po rozwinięciu wyznacznika ma postać

$$(9{,}257 - 10{,}511\lambda)(11{,}483 - 13{,}302\lambda) - (9{,}190 - 10{,}152\lambda)^2 = 0 \tag{5.3.2}$$

Wyznaczono wartość własną $\lambda_1 = 0{,}889$ (oraz $\lambda_2 = 0{,}669$) i wykorzystując warunek (2.5.16) rozwiązano równanie macierzowe (2.5.18), otrzymując pierwszy wektor główny

$$g_1 = [0{,}895; \; 0{,}446]. \tag{5.3.3}$$

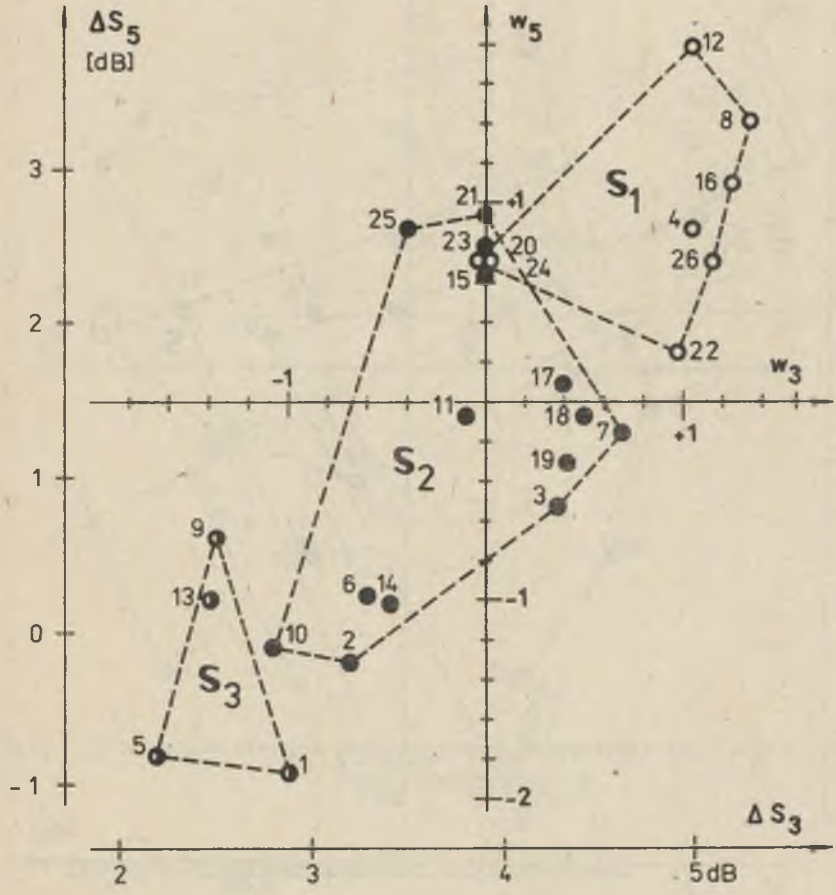
Korzystając z tego, że przekształcana przestrzeń $\{w_3, w_5\}$ jest przestrzenią dwuwymiarową, drugi wektor główny wyznaczono bezpośrednio z warunku ortogonalności (2.5.19), otrzymując macierz przekształcenia

$$\Pi = \begin{bmatrix} 0{,}895; & 0{,}446 \\ -0{,}446; & 0{,}895 \end{bmatrix}. \tag{5.3.4}$$

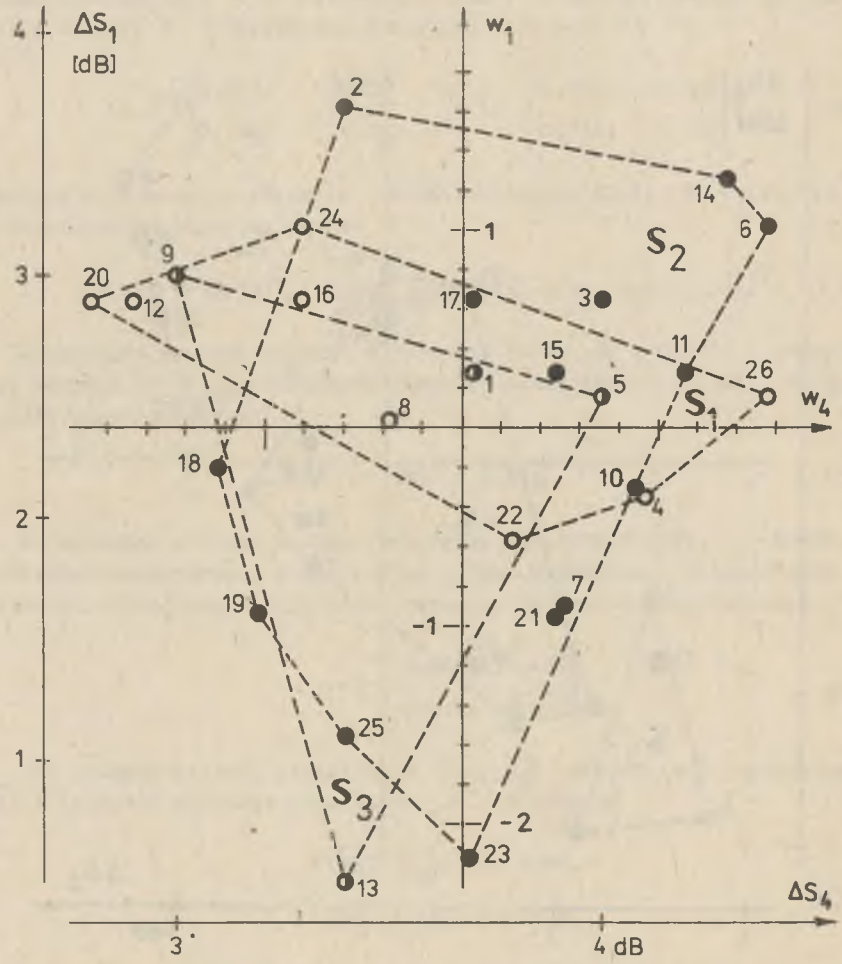
Dla przekształconej przestrzeni $\{w_7, w_8\} = \mathfrak{T}(\{w_3, w_5\})$ wyznaczono zgodnie z (2.5.14) wartości wrażliwości ε, otrzymując

$$\begin{aligned} \varepsilon(w_7, \{s_k\}) &= 0{,}89, \\ \varepsilon(w_8, \{s_k\}) &= 0{,}79. \end{aligned} \tag{5.3.5}$$

Ze względu na "małą" wartość $\varepsilon(w_8)$, zbliżoną do wartości wyznaczonej dla przestrzeni sześciowymiarowej ($\varepsilon = 0{,}78$), pominięto cechę w_8. Roz-



Rys. 5.3. Podział elementów ograniczonej przestrzeni wartości cech sygnałów (o dużej czułości rozpoznawania klas stanów), wyznaczony przez klasy stanów

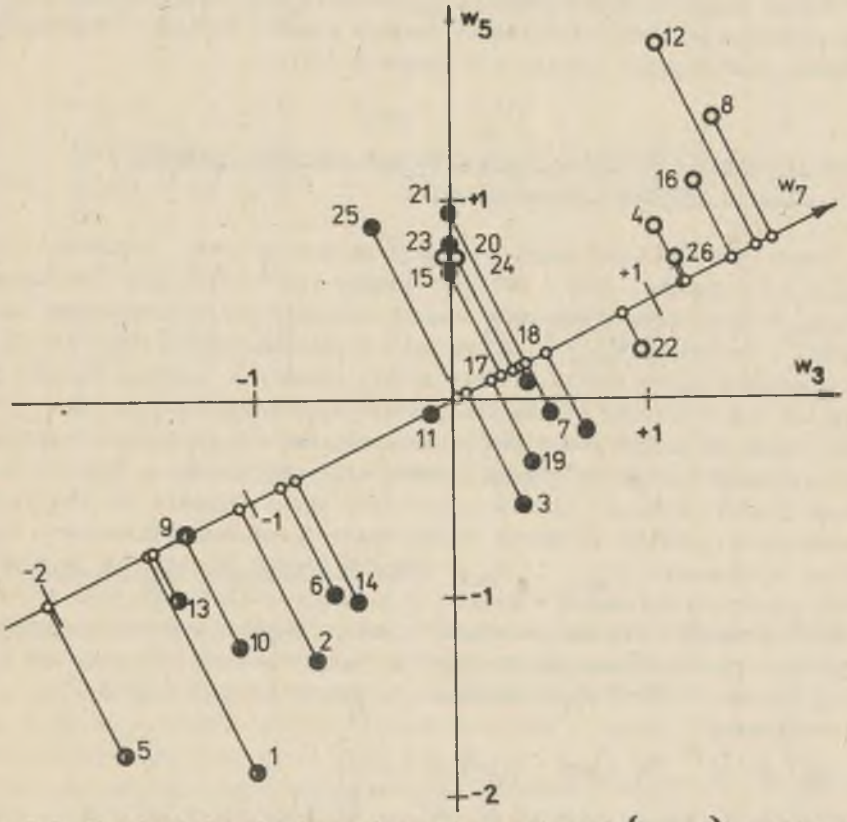


Rys. 5.4. Podział elementów ograniczonej przestrzeni wartości cech sygnałów (o małej czułości rozpoznawania klas stanów), wyznaczony przez klasy stanów

patrywaną dalej cechą sygnału jest wyłącznie cecha w_7, będąca liniową kombinacją cech w_3 i w_5. Standaryzowane wartości cechy w_7 są wyznaczane zgodnie z zależnością

$$w_7 = 0{,}776\, w_3 + 0{,}387\, w_5. \qquad (5.3.6)$$

Przekształcenie dwuwymiarowej przestrzeni $\{w_3, w_5\}$ w jednowymiarową przestrzeń $\{w_7\}$ pokazano na rys. 5.5.



Rys. 5.5. Przekształcenie dwuwymiarowej przestrzeni $\{w_3, w_5\}$ w jednowymiarową przestrzeń $\{w_7\}$

5.4. Wyznaczenie obrazów sygnałów diagnostycznych

Dla wyznaczenia obrazów sygnałów diagnostycznych określono wg (D.5.7) (w jednowymiarowej przestrzeni wartości cech sygnałów) rodzinę reprezentantów trzech rozmytych klas elementów przestrzeni sygnałów, dla których

funkcje charakterystyczne przyjęto na podstawie funkcji charakterystycznych rozmytych klas stanów (wg tabl. 5.2) zgodnie z (D.5.8). Otrzymano

$$\begin{aligned} q(W_1) &= [1,02], \\ q(W_2) &= [-0,11], \\ q(W_3) &= [-1,55]. \end{aligned} \tag{5.4.1}$$

Obrazy sygnałów wyznaczono dla tak określonych reprezentantów, postępując podobnie jak przy wyznaczaniu obrazów stanów. Wartości współrzędnych obrazów zestawiono w tablicy 5.2 (kolumny 6÷8).

5.5. Wyznaczenie rozmytej relacji między obrazami sygnałów diagnostycznych i obrazami stanu

Przed wyznaczeniem rozmytej relacji między obrazami sygnałów (wg tablicy 5.2 - kolumny 6÷8) i obrazami stanów (wg tablicy 5.2 - kolumny 3÷5) przeprowadzono zmianę zakresów obrazów stanów zgodnie z wnioskami wynikającymi z twierdzenia D.3.1. W tym celu wyznaczono zgodnie z (D.3.13) wymagane stopnie ograniczenia zakresu obrazu stanu dla każdego obiektu (por. tablica 5.2 - kolumna 9). Na ich podstawie przyjęto wspólną dla wszystkich obiektów liczbę $\alpha = 1,666$ spełniającą warunek (D.4.3) i wyznaczono ograniczone, uzgodnione obrazy stanów, które zestawiono w tablicy 5.2 (kolumny 10÷12). Dla par obrazów (uzgodniony obraz sygnałów diagnostycznych; ograniczony, uzgodniony obraz stanu) każdej przekładni wyznaczono ograniczone rozwiązania D_{min} i D_{max} układów równań (4.1.18) i (4.1.19) według algorytmu opisanego w dodatku D.4.1, przyjmując w (D.4.9) różne wartości liczby β. Dla ograniczonych rozwiązań układu równań, odpowiadających kolejnym przyjmowanym liczbom β, wyznaczono wagi ϑ zgodnie z (4.3.2) oraz liczby Ψ określające rozmycie rozwiązań zgodnie z (D.4.6).

Otrzymano:

- dla $\beta = 0$: $\vartheta = 0,62$, $_{(26)}\Psi = 0,64$,

$$D_{min} = \begin{bmatrix} 0,33; & 0,30; & 0,21 \\ 0,22; & 0,25; & 0,21 \\ 0,22; & 0,25; & 0,24 \end{bmatrix}, \quad D_{max} = \begin{bmatrix} 0,79; & 0,75; & 0,78 \\ 0,79; & 0,71; & 0,78 \\ 0,79; & 0,75; & 0,27 \end{bmatrix}, \tag{5.5.1}$$

- dla $\beta = 0,2$: $\vartheta = 0,62$, $_{(26)}\Psi = 0,57$,

$$D_{min} = \begin{bmatrix} 0,33; & 0,30; & 0,21 \\ 0,29; & 0,30; & 0,21 \\ 0,22; & 0,25; & 0,34 \end{bmatrix}, \quad D_{max} = \begin{bmatrix} 0,40; & 0,75; & 0,78 \\ 0,79; & 0,70; & 0,78 \\ 0,79; & 0,75; & 0,25 \end{bmatrix}, \tag{5.5.2}$$

- dla $\beta = 0{,}3$: $\vartheta = 0{,}74$, $(26)^{\Psi} = 0{,}49$,

$$D_{min} = \begin{bmatrix} 0{,}39; & 0{,}31; & 0{,}21 \\ 0{,}30; & 0{,}31; & 0{,}21 \\ 0{,}23; & 0{,}32; & 0{,}50 \end{bmatrix}, \quad D_{max} = \begin{bmatrix} 0{,}33; & 0{,}75; & 0{,}77 \\ 0{,}79; & 0{,}33; & 0{,}78 \\ 0{,}79; & 0{,}75; & 0{,}25 \end{bmatrix}, \qquad (5.5.3)$$

- dla $\beta = 0{,}4$: $\vartheta = 0{,}81$, $(26)^{\Psi} = 0{,}31$,

$$D_{min} = \begin{bmatrix} 0{,}39; & 0{,}31; & 0{,}21 \\ 0{,}33; & 0{,}71; & 0{,}25 \\ 0{,}23; & 0{,}32; & 0{,}50 \end{bmatrix}, \quad D_{max} = \begin{bmatrix} 0{,}30; & 0{,}75; & 0{,}77 \\ 0{,}79; & 0{,}33; & 0{,}77 \\ 0{,}79; & 0{,}75; & 0{,}25 \end{bmatrix}, \qquad (5.5.4)$$

- dla $\beta = 0{,}5$: $\vartheta = 0{,}74$, $(26)^{\Psi} = 0{,}23$,

$$D_{min} = \begin{bmatrix} 0{,}43; & 0{,}31; & 0{,}25 \\ 0{,}33; & 0{,}73; & 0{,}25 \\ 0{,}29; & 0{,}32; & 0{,}77 \end{bmatrix}, \quad D_{max} = \begin{bmatrix} 0{,}29; & 0{,}73; & 0{,}50 \\ 0{,}79; & 0{,}32; & 0{,}77 \\ 0{,}40; & 0{,}73; & 0{,}25 \end{bmatrix}, \qquad (5.5.5)$$

- dla $\beta = 0{,}6$: $\vartheta = 0{,}52$, $(26)^{\Psi} = 0{,}47$,

$$D_{min} = \begin{bmatrix} 0{,}79; & 0{,}31; & 0{,}25 \\ 0{,}39; & 0{,}75; & 0{,}25 \\ 0{,}29; & 0{,}70; & 0{,}77 \end{bmatrix}, \quad D_{max} = \begin{bmatrix} 0{,}29; & 0{,}71; & 0{,}50 \\ 0{,}79; & 0{,}32; & 0{,}77 \\ 0{,}40; & 0{,}33; & 0{,}21 \end{bmatrix}. \qquad (5.5.6)$$

Jako optymalne przyjęto rozwiązanie (5.5.4) wyznaczone dla $\beta = 0{,}4$, ponieważ dla tego rozwiązania otrzymano największą wartość wagi ϑ.

5.6. Weryfikacja układu diagnozującego

Rozwiązanie (5.5.4) określa działanie układu diagnozującego umożliwiającego wnioskowanie o prawdopodobnej klasie stanu przekładni na podstawie obrazu jej sygnałów diagnostycznych. Weryfikację przeprowadzono zgodnie z rozdz. 4.3, na podstawie zbioru "danych uczących". Ograniczenia obrazów stanów obliczone zgodnie z (4.1.22) dla D_{min}, D_{max} wg (5.5.4) i p_w wg tablicy 5.2, wnioski o prawdopodobnej klasie stanów przekładni określone zgodnie z rozdz. 4.2 na podstawie ograniczeń obrazów i klasy stanu wyznaczone bezpośrednio zestawiono w tablicy 5.4. Wyznaczono diagnozy:

- jednoznaczne dla 22 przekładni, w tym dla 3 przekładni wyznaczona diagnoza jednoznaczna jest błędna,
- dwuznaczne dla 2 przekładni,
- trywialne (brak diagnozy) dla 2 przekładni.

Tablica 5.4

Weryfikacja działania układu diagnozującego

Lp.	Obiekt	Ograniczenia obrazu stanu						Klasa stanu wg ograniczeń obrazu stanu	Klasa stanu wg p_s (tabl.5.2)	Uwagi o rozpoznaniu klasy stanu
		$p_{s,min}$ [1]	$p_{s,min}$ [2]	$p_{s,min}$ [3]	$p_{s,max}$ [1]	$p_{s,max}$ [2]	$p_{s,max}$ [3]			
1	2	3	4	5	6	7	8	9	10	11
1	110	0,25	0,32	0,50	0,30	0,33	0,77	3	3	diagnoza jednoznaczna poprawna
2	120	0,33	0,54	0,50	0,30	0,54	0,77	1 lub 2 lub 3	2	diagnoza trywialna
3	130	0,39	0,71	0,25	0,39	0,75	0,25	2	2	diagnoza jednoznaczna poprawna
4	155	0,39	0,31	0,25	0,79	0,33	0,25	1	1	diagnoza jednoznaczna poprawna
5	210	0,23	0,32	0,50	0,30	0,33	0,77	3	3	diagnoza jednoznaczna poprawna
6	220	0,33	0,71	0,50	0,30	0,74	0,76	1 lub 2 lub 3	2	diagnoza trywialna
7	230	0,39	0,71	0,25	0,79	0,75	0,25	1 lub 2	2	diagnoza dwuznaczna poprawna
8	255	0,39	0,31	0,25	0,79	0,33	0,25	1	1	diagnoza jednoznacza poprawna
9	410	0,33	0,32	0,50	0,30	0,33	0,77	3	3	diagnoza jednoznaczna poprawna
10	420	0,33	0,34	0,50	0,30	0,34	0,77	3	2	diagnoza jednoznaczna błędna
11	430	0,33	0,71	0,25	0,30	0,75	0,25	2	2	diagnoza jednoznaczna poprawna
12	455	0,39	0,31	0,25	0,79	0,33	0,25	1	1	diagnoza jednoznaczna poprawna
13	510	0,27	0,32	0,50	0,30	0,33	0,77	3	3	diagnoza jednoznaczna poprawna
14	520	0,33	0,71	0,50	0,30	0,75	0,67	2	2	diagnoza jednoznaczna poprawna
15	530	0,39	0,71	0,25	0,52	0,75	0,25	2	2	diagnoza jednoznaczna poprawna
16	555	0,39	0,31	0,25	0,79	0,33	0,25	1	1	diagnoza jednoznaczna poprawna
17	1530	0,39	0,71	0,25	0,62	0,75	0,25	2	2	diagnoza jednoznaczna poprawna
18	1630	0,39	0,71	0,25	0,64	0,75	0,25	2	2	diagnoza jednoznaczna poprawna
19	1730	0,39	0,71	0,25	0,45	0,75	0,25	2	2	diagnoza jednoznaczna poprawna
20	1755	0,39	0,71	0,25	0,55	0,75	0,25	2	1	diagnoza jednoznaczna błędna
21	1830	0,39	0,71	0,25	0,66	0,75	0,25	2	2	diagnoza jednoznaczna poprawna
22	1855	0,39	0,42	0,25	0,79	0,42	0,25	1 lub 2	1	diagnoza dwuznaczna poprawna
23	1930	0,39	0,71	0,25	0,59	0,75	0,25	2	2	diagnoza jednoznaczna poprawna
24	1955	0,39	0,71	0,25	0,55	0,75	0,25	2	1	diagnoza jednoznaczna błędna
25	2030	0,34	0,71	0,25	0,34	0,75	0,25	2	2	diagnoza jednoznaczna poprawna
26	2055	0,39	0,31	0,25	0,79	0,33	0,25	1	1	diagnoza jednoznaczna poprawna

6. UWAGI KOŃCOWE

Praca jest wynikiem poszukiwania metody wnioskowania umożliwiającej identyfikację klas stanów maszyn na podstawie obserwowanych sygnałów diagnostycznych. Podjęto próbę opracowania sposobu postępowania wykorzystującego częściowo metody rozpoznawania obrazów, teorię systemów i teorię zbiorów rozmytych.

Wprowadzono pojęcie przestrzeni obrazów cech sygnałów diagnostycznych (rozdz. 2.6) i stanów maszyny (rozdz.3.3) oraz sformułowano tezę, iż zastosowanie przestrzeni obrazów umożliwia ograniczenie liczby wymiarów przestrzeni analizowanych w badaniach diagnostycznych,upraszczając jednocześnie procedury związane z wyznaczaniem działania układów diagnozujących.

Dla określenia zakresów znaczeniowych pojęć: "przestrzeń cech" i "przestrzeń obrazów cech", dokonano (rozdz. 2) szczegółowej analizy przekształceń oddziaływań otoczenia na maszynę i maszyny na otoczenie, obserwowanych za pośrednictwem sygnałów (rozpatrywanych jako procesy stochastyczne), w zbiory ich cech. Rozpatrując przestrzeń obrazów omówiono szczegółowo zagadnienia związane z przekształcaniem przestrzeni cech w przestrzeń obrazów, obejmujące metody porównywania (dod. A), grupowania i klasyfikacji (dod. B i C) elementów przestrzeni cech oraz uzasadniono (rozdz. 2.6) możliwość interpretowania elementu przestrzeni obrazów jako rozmytego zbioru klas elementów przestrzeni cech.

Model diagnostyczny maszyny będącej obiektem badań określono (rozdz.3) wykorzystując teorię systemów. Uwzględniając (rozdz. 3.2) potrzebę modelowania klasy maszyn wykonanych wg wspólnej konstrukcji sformułowano i następnie uzasadniono tezę (rozdz.3.4),że modelowanie takie może być realizowane za pomocą rozmytych relacji między obrazami cech sygnałów diagnostycznych i obrazami cech stanów maszyn, umożliwiając jednocześnie wyznaczanie działania układu diagnozującego (rozdz.4.1).Ze względu na potrzeby pracy rozwinięto metody wyznaczania relacji rozmytych (dod.D) przez wprowadzenie oryginalnego sposobu zapisu równań (dod.D.3) i układów równań rozmytych (dod.D.4).Zaproponowano algorytmy ich rozwiązywania (dod.D.1.1).

Układy diagnozujące o działaniu określonym relacjami rozmytymi, wyznaczanymi w wyniku rozwiązania odpowiedniego układu równań ((4.1.18) i (4.1.19)) zasługują na uwagę z tego względu, że umożliwiają one (por.rozdział 4.2) dostarczenie użytkownikowi układu informacji o prawdopodobnej klasie stanów maszyny oraz umożliwiają sygnalizowanie przez układ tych przypadków (por. (4.2.5)), w których nie potrafi on rozpoznać klasy stanu. Zaproponowany sformalizowany sposób wnioskowania, mogący stanowić pod-

stawę metodycznego projektowania (wyznaczania działania) układów diagnozujących wymaga znajomości wielu dyscyplin podstawowych (w tym szczególnie wibroakustycznej diagnostyki maszyn) i nie może on zastąpić doraźnych sposobów wnioskowania bazujących głównie na intuicji badacza, które w wielu praktycznych zastosowaniach mogą być bardziej skuteczne.

Cel pracy, którym było sformułowanie metody projektowania układów diagnozujących przeznaczonych do rozpoznawania klas stanów maszyn na podstawie obserwowanych sygnałów diagnostycznych został osiągnięty. Ograniczeniem dla proponowanej metody projektowania jest potrzeba uzyskania wyników "badań uczących", podczas których stan maszyn jest oceniany metodami bezpośrednimi na podstawie opinii ekspertów lub na podstawie szczegółowych badań elementów maszyn.

Oceniając wykonaną pracę autor uważa, że najważniejszymi, oryginalnymi (jego zdaniem) jej elementami są:

- wprowadzenie pojęcia przestrzeni obrazów cech (rozdz. 2.6) umożliwiającego równoczesne stosowanie metod rozpoznawania obrazów i teorii zbiorów rozmytych oraz umożliwiającego prostą interpretację przekształceń przestrzeni wykonywanych w procesie wnioskowania w badaniach diagnostycznych,
- propozycja wyznaczania działania układu diagnozującego za pomocą pary relacji rozmytych ((4.1.20) i (4.1.21)),
- rozwinięcie metody zapisywania i rozwiązywania równań rozmytych (dod. D.3) oraz wprowadzenie relacji podobieństwa zbiorów rozmytych (D.1.14)), umożliwiające rozpatrywanie układów równań rozmytych oraz umożliwiające określanie algorytmów optymalizacji ich rozwiązań (dod. D.4.1),
- określenie sposobu wnioskowania o klasie stanu na podstawie ograniczeń obrazu stanu (rozdz. 4.2) umożliwiającego otrzymanie jednoznacznego,wieloznacznego lub trywialnego wyniku diagnozowania, gdzie wynik trywialny interpretowany jest jako brak możliwości rozpoznania klasy stanu przez układ diagnozujący,
- próba formalizacji (rozdz. 2) podstawowych pojęć związanych z obserwowaniem obiektu w badaniach diagnostycznych.

Przewiduje się, że dalsze (związane z badaniami diagnostycznymi maszyn) prace autora i współpracującej z nim grupy pracowników IMiPKM Politechniki Śląskiej dotyczyć będą:

- modernizacji układu sterowania programowanego analizatora sygnałów PAS4/CAMAC [24] (umożliwiającej przyspieszenie wykonywanych przez analizator działań i zwiększenie objętości realizowanych programów) oraz wynikającej stąd modernizacji podstawowego oprogramowania analizatora,
- modernizacji powstałego podczas wykonywania pracy zbioru programów dla EMC ODRA 1305 [26] oraz jego adaptacji umożliwiającej zastosowanie programowanego analizatora sygnałów,

- opracowania konstrukcji typoszeregu układów pomiarowych umożliwiających diagnozowanie stanu maszyn, których istota działania wyznaczana będzie relacjami rozmytymi,
- prób pełnego zastosowania opracowanej metody wnioskowania do oceniania stanu na stanowisku kontroli technicznej wybranego typu maszyn produkowanych seryjnie.

DODATKI

A. PORÓWNYWANIE ELEMENTÓW PRZESTRZENI WARTOŚCI CECH

W pracy wykorzystuje się pojęcie rozmytej klasy obiektów, definiowanej (2.6.1) jako określony w przestrzeni wartości cech rozmyty zbiór elementów podobnych do reprezentanta klasy. Dla określenia takiego zbioru, tj. dla wyznaczenia jego funkcji charakterystycznej, porównuje się elementy przestrzeni wartości cech. Porównywanie elementów jest również niezbędne podczas rozwiązywania zadań związanych z grupowaniem elementów przestrzeni wartości cech będących podstawą określania ich obrazów (por.rozdz.2.6).

Porównywanie elementów polega na wyznaczaniu ich odległości i/lub podobieństw w przestrzeni wartości cech. Formalizacja metod porównywania elementów wymaga wprowadzenia miar odległości i/lub miar podobieństw.

Przyjęto założenie (por. (2.4.6)), że elementy m-wymiarowej przestrzeni wartości cech V^m określone są przez wartości cech zapisywane w postaci macierzy.

A.1. Odległość elementów przestrzeni wartości cech

Definicja A.1.1: Odległość elementów $x, y \in V$ jest wartością funkcji $d(x, y)$ nazywanej funkcją odległości, takiej że:

$$d : V \times V \longrightarrow \{r \in R^1 : r \geq 0\}, \tag{A.1.1}$$

spełniającą warunki:

$$\underset{x,y \in V}{\forall} \; [d(x,x) \leq d(x,y)], \tag{A.1.2}$$

$$\underset{x,y \in V}{\forall} \; [d(x,y) = d(y,x)], \tag{A.1.3}$$

$$\underset{x,y,z \in V}{\forall} \; [d(x,z) \leq d(x,y) + d(y,z)], \tag{A.1.4}$$

gdzie:

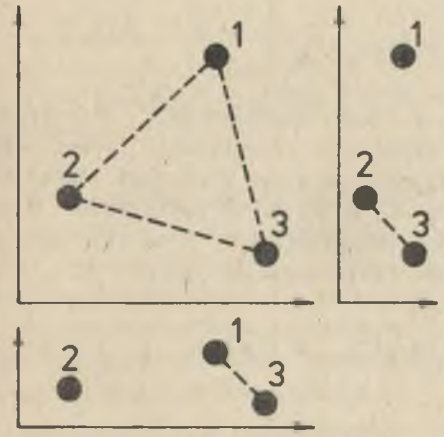
R^1 - zbiór liczb rzeczywistych.

Ustalając postać funkcji odległości uwzględnia się znaczenie rozpatrywanych cech, a zwłaszcza odrębnie rozpatruje się cechy ilościowe i cechy jakościowe (por. rozdz. 2.4).

A.1.1. Odległość w przestrzeni wartości cech ilościowych

Najczęściej stosowaną funkcją odległości jest odległość euklidesowa

$$d_1^2(x,y) \triangleq (x-y)\,(x-y)^T \quad \forall\, x,y \in V. \qquad (A.1.5)$$



Rys. A.1. Przykład wpływu zmian skali współrzędnych dwuwymiarowej przestrzeni wartości cech na wyniki grupowania według odległości euklidesowej

Przyjęcie odległości euklidesowej związane jest z niedogodnościami wynikającymi z dużego wpływu zmian skal współrzędnych na wyniki grupowania elementów przestrzeni cech (por. rys. A.1). Oznacza to, że wprowadzając odległość euklidesową jako podstawę grupowania, należy zakładać izotropowość przestrzeni V. Założenie to jest założeniem bardzo silnym zwłaszcza wtedy, gdy przestrzeń V jest przestrzenią cech o nieporównywalnych wartościach (np. wartość ciśnienia oleju i wartość prędkości obrotowej) i podczas jej określania brak formalnych podstaw uzasadniających przyjęcie skal poszczególnych współrzędnych.

Dla uniknięcia powyższych niedogodności zaleca się normalizować przestrzeń wartości cech tak, aby zbiory wartości kolejnych współrzędnych posiadały zerową wartość średnią i jednostkowe odchylenie standardowe. Ponadto dla uniezależnienia wyników porównywania od wyboru układu współrzędnych można dokonać obrotu układu współrzędnych, ustalając takie jego położenie, które pokrywać się będzie z kierunkami osi głównych dla rozpatrywanego zbioru elementów przestrzeni V. Kierunki te odpowiadać będą kierunkom wektorów własnych macierzy kowariancji współrzędnych elementów v. Otrzymana w ten sposób przestrzeń nazywana jest przestrzenią wartości cech głównych i oznaczana $\mathfrak{G}^m = \mathfrak{R}(V^m)$ (por. (2.5.1)).

Dla zróżnicowania wpływu poszczególnych współrzędnych na odległość elementów przestrzeni V^m można wprowadzić wagi współrzędnych. W przestrzeni wartości cech wagi współrzędnych są wagami kolejnych cech. Prowadzi to do odległości Sebestyena [88]

$$d_2^2(x,y) \triangleq (x-y)\,W(x-y)^T \quad \forall\, x,y \in V, \qquad (A.1.6)$$

gdzie:

W jest diagonalną macierzą wag

$$W = \begin{bmatrix} w_1, & 0, & \dots, & 0 \\ 0, & w_2, & \dots, & 0 \\ \dots & & \dots & \\ 0, & 0, & \dots, & w_m \end{bmatrix}. \qquad (A.1.7)$$

Jeżeli przyjmuje się celowość zmiany układu współrzędnych przestrzeni V i przeprowadza się jego obrót do położenia wyznaczonego przez osie główne, a następnie normalizację współrzędnych w tak obróconym układzie, to dla wyznaczenia odległości euklidesowej w tej nowej przestrzeni G można wykorzystać bezpośrednio współrzędne przestrzeni V i na ich podstawie wyznaczyć tzw. odległość uogólnioną (Mahalanobisa) [63]

$$d_3^2(x,y) \stackrel{d}{=} (x-y)\ C^{-1}(x-y)^T \qquad \forall\ x,y \in V, \qquad (A.1.8)$$

gdzie:

C^{-1} - odwrotna macierz kowariancji

$$C = \frac{1}{n} \sum_{j=1}^{n} (v_j - \bar{v})^T (v_j - \bar{v}), \qquad v_j \in V, \qquad (A.1.9)$$

n - liczba elementów rozpatrywanego zbioru w przestrzeni V,

$\bar{v}$ - średni element rozpatrywanego zbioru

$$\bar{v} = \frac{1}{n} \sum_{j=1}^{n} v_j. \qquad (A.1.10)$$

Odległość uogólniona uwzględnia korelację współrzędnych. Opisane odległości d_1, d_2, d_3 posiadają niedogodność polegającą na dużej liczbie działań niezbędnych do ich wyznaczania, co wpływa na czas wykonywania obliczeń podczas realizacji algorytmów grupowania wykorzystujących te odległości. Dla zmniejszenia czasu obliczeń zaleca się stosowanie uśrednionej odległości Hamminga (odległości Manhattan [100])

$$d_4(x,y) \stackrel{d}{=} \frac{1}{m} \sum_{i=1}^{m} |x[i] - y[i]| \qquad \forall\ x,y \in V^m. \qquad (A.1.11)$$

Pewną odmianą odległości Hamminga jest odległość Canbera [100]

$$d_5(x,y) \stackrel{d}{=} \sum_{i=1}^{m} \frac{|x[i] - y[i]|}{x[i] + y[i]} \qquad \forall\, x,y \in V^m. \qquad (A.1.12)$$

Stosowanie odległości d_5 wymaga założenia, że współrzędne $x[i]$, $y[i]$ są dodatnie. Szczególną zaletą tej odległości jest brak wpływu niezależnego powiększania współrzędnych na wielkość d_5, istnieje jednak wpływ przesuwania tego układu.

A.1.2. Odległość w przestrzeni wartości cech jakościowych

Dla cech jakościowych przyjmuje się kody umożliwiające zapisywanie wartości tych cech w postaci liczb, co powoduje, że w większości przypadków bezpośrednie zastosowanie funkcji odległości podobnych do odległości euklidesowej nie posiada uzasadnienia. Dla porównywania wielowartościowych cech jakościowych, dla których nie określono uporządkowania wartości, wprowadza się pojęcie cechy binarnej.

Definicja A.1.2: Cecha binarna jest cechą o dwuelementowym zbiorze wartości. Przyjmuje się zbiór wartości cech binarnych $\{0; 1\}$.

Twierdzenie A.1.1: Niech V^m będzie m-wymiarową przestrzenią wartości cech jakościowych, takich, że moce zbiorów wartości $\mathrm{Cod}(v_i)$ kolejnych cech v_i są skończone. Istnieje wzajemnie jednoznaczne odwzorowanie

$$V^m \rightarrowtail V_b^n \subset \bigtimes_{j=1}^{n} \{0, 1\} \qquad (A.1.13)$$

przestrzeni V^m w przestrzeń cech binarnych V_b^n taką, że

$$n \leqslant \prod_{i=1}^{m} \overline{\overline{\mathrm{Cod}(v_i)}}. \qquad (A.1.14)$$

Z twierdzenia wynika, że wielowartościowe cechy jakościowe o skończonej mocy zbioru wartości mogą być zastępowane ciągami cech binarnych.

Przykład: Jednowymiarowa przestrzeń wartości cech jakościowych wyznaczona cechą "stan uzębienia koła zębatego" o wartościach "brak pittingu", "pitting przemijający", "pitting postępujący", może być odwzorowana w 3-wymiarową przestrzeń wartości cech binarnych o elementach definiowanych następująco

$$v = [v_1, v_2, v_3], \qquad (A.1.15)$$

gdzie:

$$v_1 = \begin{cases} 1 & \text{jeżeli nie wystąpił pitting,} \\ 0 & \text{jeżeli wystąpił pitting,} \end{cases}$$

$$v_2 = \begin{cases} 1 & \text{jeżeli wystąpił pitting przemijający,} \\ 0 & \text{jeżeli nie wystąpił pitting przemijający,} \end{cases}$$

$$v_3 = \begin{cases} 1 & \text{jeżeli wystąpił pitting postępujący,} \\ 0 & \text{jeżeli nie wystąpił pitting postępujący} \end{cases}$$

lub w dwuwymiarową przestrzeń wartości cech binarnych o elementach

$$v = [v_2, v_3]. \qquad (A.1.16)$$

Przyjmuje się, że dla określenia odległości w przestrzeni wartości cech jakościowych uzasadnione jest wyróżnianie dwóch wartości tej odległości: "cechy równe", "cechy różne" [82]. Odległość "cechy równe" odpowiada odległości zerowej dla cech ilościowych, co można zapisać

$$d_6(x,y) \stackrel{d}{=} \begin{cases} 0 & \text{jeżeli } x = y \\ \infty & \text{jeżeli } x \neq y \end{cases} \qquad \forall x,y \in V_b. \qquad (A.1.17)$$

Dla określenia odległości w przestrzeni wartości cech binarnych można również wykorzystać funkcję odległości Hamminga d_4 według (A.1.11).

Szczególnym przypadkiem cech jakościowych są cechy, których wartości posiadają określone uporządkowanie. Przykładem takiej cechy może być cecha o wartościach "mały", "średni", "duży". Dla takich cech zaleca się przyjmować ich wartości liczbowe równe rangom w ciągu uporządkowanych wartości jakościowych. Funkcja odległości może być przyjmowana jako uproszczona wersja funkcji odległości Kendalla [15]

$$d_7(x,y) \stackrel{d}{=} \sum_{i=1}^{m} (r(x[i]) - r(y[i]))^2 \qquad \forall x,y \in V^m, \qquad (A.1.18)$$

gdzie:

r(v) - ranga wartości cechy w uporządkowanym ciągu wartości tej cechy.

Odległość d_7 może być również wyznaczona w przestrzeni cech ilościowych, gdzie szczególną jej zaletą jest niezależność od przekształceń układu współrzędnych nie zmieniających porządku elementów, a więc na przykład niezależność od zmian skali.

A.2. Podobieństwo elementów przestrzeni wartości cech

Definicja A.2.1: Podobieństwo elementów $x,y \in V$ jest wartością funkcji $h(x,y)$ nazywanej funkcją podobieństwa, takiej, że:

$$h : V \times V \rightarrow \left\{ r \in R^1 : 0 \leqslant r \leqslant 1 \right\}, \tag{A.2.1}$$

spełniającej warunki:

$$\underset{x,y \in V}{\forall} \left[h(x,y) \leqslant h(x,x) \right], \tag{A.2.2}$$

$$\underset{x,y \in V}{\forall} \left[h(x,y) = h(y,x) \right]. \tag{A.2.3}$$

W większości przypadków zakłada się, że

$$\underset{x \in V}{\forall} \left[h(x,x) = 1 \right]. \tag{A.2.4}$$

Założenie (A.2.4) nie dotyczy funkcji podobieństwa wykorzystywanych do porównywania zbiorów elementów ze zbiorami elementów (por.(A.4.3),(A.4.4)).

Funkcja podobieństwa może być określana bezpośrednio lub do jej określenia można wykorzystać funkcję odległości, tzn. można definiować $h(d)$ gdzie d jest wartością funkcji odległości elementów przestrzeni wartości cech. Tak definiowana funkcja podobieństwa powinna spełniać warunki

$$h(0) = 1, \tag{A.2.5}$$

$$\underset{d_1,d_2}{\forall} \left[d_2 > d_1 \Longrightarrow h(d_2) < h(d_1) \right]. \tag{A.2.6}$$

Z porównania różnych postaci funkcji odległości wynika, że w większości przypadków wartości tych funkcji rosną ze wzrostem liczby wymiarów m przestrzeni V^m, w której te odległości są wyznaczane. Dla zmniejszenia wpływu liczby wymiarów przestrzeni wartości cech na wartość funkcji podobieństwa elementów przeprowadza się skalowanie funkcji odległości przed wykorzystaniem jej do określania funkcji podobieństwa:

- dla d_1 wg (A.1.5), d_2 wg (A.1,6), d_3 wg (A.1.8), d_7 wg (A.1.18)

$$d(x,y) := d(x,y)/\sqrt{m}, \tag{A.2.7}$$

- dla d_5 wg (A.1.12)

$$d(x,y) := d(x,y)/m, \tag{A.2.8}$$

- dla d_4 wg (A.1.11), d_6 wg (A.1.17)

$$d(x,y) := d(x,y). \tag{A.2.9}$$

A.2.1. Podobieństwo elementów w przestrzeni wartości cech ilościowych

Dla określenia podobieństwa na podstawie funkcji odległości można wykorzystywać różne funkcje podobieństwa. Przykładami takich funkcji są następujące zależności

$$h_1(x,y) \stackrel{d}{=} \frac{1}{1 + \alpha d(x,y)^{\beta}}, \tag{A.2.10}$$

$$h_2(x,y) \stackrel{d}{=} 1 - \left(\frac{d(x,y)}{\max\limits_{(x,y) \in V \times V} (d(x,y))}\right), \tag{A.2.11}$$

$$h_3(x,y) \stackrel{d}{=} \exp(-\alpha d(x,y)), \tag{A.2.12}$$

gdzie:

α, β - stałe warunkujące własności funkcji h.

Niezależnie od wprowadzonych wcześniej funkcji odległości d_1 do d_7 można wyznaczać funkcję podobieństwa elementów x, y bezpośrednio w postaci cosinusa kąta między wektorami x, y w przestrzeni V

$$h_4(x,y) \stackrel{d}{=} x\, y^T / \sqrt{x\, x^T \; y \; y^T}. \tag{A.2.13}$$

Funkcja podobieństwa wg (A.2.13) jest niezależna od obrotów układu współrzędnych i jego powiększania, lecz zależy od ewentualnych przesunięć tego układu.

A.2.2. Podobieństwo elementów w przestrzeni wartości cech jakościowych

Wielu autorów, np. [82], proponuje następujące definiowanie funkcji podobieństwa w przestrzeni cech jakościowych

$$h_5(x,y) \stackrel{d}{=} \begin{cases} 1 & \text{jeżeli } x = y, \\ 0 & \text{jeżeli } x \neq y. \end{cases} \tag{A.2.14}$$

Definicja taka jest intuicyjnie uzasadniona dla jednowymiarowej przestrzeni cech. Nie umożliwia ona jednak skutecznego porównywania elementów wielowymiarowych przestrzeni cech jakościowych.

Niech V_b^m jest przestrzenią cech binarnych. Współrzędne elementów tej przestrzeni przyjmują wartości liczbowe 0 lub 1, co pozwala na interpretację geometryczną przedstawiającą elementy tej przestrzeni jako wierzchołki m-wymiarowego sześcianu o jednakowej długości krawędzi. Z wielu określeń funkcji podobieństwa dla takiej przestrzeni, podawanych przez różnych autorów, zwraca się uwagę na następujące

$$h_6(x,y) \triangleq \sum_{i=1}^{m} \min(x[i], y[i]), \tag{A.2.15}$$

$$h_7(x,y) \triangleq \frac{\sum_{i=1}^{m} \min(x[i], y[i])}{\sum_{i=1}^{m} \max(x[i], y[i])}, \tag{A.2.16}$$

$$h_8(x,y) \triangleq \frac{\sum_{i=1}^{m} \min(x[i], y[i])}{\min(\sum_{i=1}^{m} x[i], \sum_{i=1}^{m} y[i])}. \tag{A.2.17}$$

A.2.3. Podobieństwo elementów w ogólnej przestrzeni wartości cech

Pod pojęciem ogólnej przestrzeni wartości cech rozumie się przestrzeń będącą iloczynem kartezjańskim przestrzeni wartości cech ilościowych oraz przestrzeni wartości cech jakościowych. Problem jednoczesnego rozpatrywania cech ilościowych i jakościowych jest trudny. Niektórzy autorzy, np. [29], proponują określanie funkcji podobieństwa elementów w ogólnej przestrzeni wartości cech jako kombinację funkcji podobieństwa h_i, h_j elementów w podprzestrzeni cech ilościowych V_i i podprzestrzeni cech jakościowych V_j

$$h_9(x,y) \triangleq \alpha\, h_i(x_i,y_i) + \beta\, h_j(x_j,y_j) \tag{A.2.18}$$

lub

$$h_{10}(x,y) \triangleq (h_i(x_i,y_i))^{\alpha} \quad (h_j(x_j,y_j))^{\beta}, \tag{A.2.19}$$

gdzie:

$$x,y \in V; \quad x_i, y_i \in V_i; \quad x_j, y_j \in V_j,$$

$$V = V_i \times V_j, \tag{A.2.20}$$

$$\alpha + \beta = 1. \tag{A.2.21}$$

Przyjmuje się, że liczby α i β są proporcjonalne do liczby wymiarów podprzestrzeni wartości cech ilościowych i jakościowych.

Wyznaczając funkcję podobieństwa w ogólnej przestrzeni wartości cech należy podjąć decyzję o postaci funkcji h_i, h_j. Brak podstaw do określenia ogólnych zaleceń w tym zakresie. Przyjęta postać tych funkcji wpływa na przebieg grupowania elementów przestrzeni wartości cech, jednak w większości zastosowań praktycznych wpływ ten nie powoduje znaczących zmian wyników grupowania, zwłaszcza wtedy, gdy rozpatrywany zbiór elementów dzieli się "wyraźnie" na rozłączne podzbiory.

A.2.4.Układy pomiarowe

Podczas konstruowania specjalistycznej aparatury pomiarowej wykorzystującej koncepcję zbiorów rozmytych, ustala się sposób wyznaczania wartości funkcji podobieństwa elementów przestrzeni wartości cech. Jeżeli analiza sygnałów realizowana jest metodami cyfrowymi, to brak jest szczególnych ograniczeń w zakresie stosowanych sposobów wyznaczania podobieństw.Jeżeli stosuje się aparaturę analogową, to przyjmuje się, że współrzędne elementów v przestrzeni cech reprezentowane są sygnałami napięciowymi zmieniającymi się w zakresie od U_{min} do U_{max} lub sygnałami prądowymi w zakresie od I_{min} do I_{max}. Poszukując funkcji podobieństwa wymagającej prostych analogowych układów elektronicznych dla jej wyznaczania należy zalecić funkcję

$$h_{11}(x^m,y^m) = U_{max} - \frac{1}{m} \sum_{i=1}^{m} |x[i] - y[i]| \tag{A.2.22}$$

przyjmującą wartości w zakresie $[U_{min} : U_{max}]$ interpretowanym jako zakres $[0 : 1]$ w rozumieniu definicji (A.2.1). Funkcję taką można wyznaczać za pomocą układu analogowego złożonego ze wzmacniaczy operacyjnych działających jako wzmacniacze różnicowe, prostowniki liniowe oraz sumatory.

A.3. Porównywanie elementów ze zbiorami elementów

Niech w przestrzeni V określony będzie zbiór X elementów $x \in X \subset V$ oraz niech dany będzie element $y \in V$. Funkcja podobieństwa między elementem y oraz zbiorem X może być określona w różny sposób. Do najprostszych należy podobieństwo średnie

$$h_{12}(y,X) \stackrel{d}{=} \frac{1}{X} \sum_{x \in X} h(y,x). \qquad (A.3.1)$$

Uwzględniając potrzeby wynikające z konieczności minimalizacji czasu obliczeń, wygodnie jest przyjąć, że podobieństwo między elementem y i zbiorem X jest równe podobieństwu między elementem y i reprezentantem $q(X)$ zbioru X

$$h_{13}(y,X) \stackrel{d}{=} h(y,q(X)). \qquad (A.3.2)$$

Dla wykorzystania zależności (A.3.2) wymagane jest określenie reprezentanta zbioru. Z wielu sposobów wyznaczania reprezentantów zbiorów zaleca się postępowanie, według którego reprezentantem zbioru jest element średni

$$q(X) = \bar{x} = \frac{1}{X} \sum_{x \in X} x, \qquad (A.3.3)$$

jeżeli X jest zbiorem wartości cech ilościowych lub postępowanie, według którego reprezentantem ogólnego zbioru wartości cech $X \subset V$ jest element $q(X) \in X$, dla którego osiąga się maksimum podobieństwa średniego

$$h_{14}(q(X),X) \stackrel{d}{=} \max_{y \in V} \left[\sum_{x \in X} h(y,x) \right] \qquad (A.3.4)$$

lub maksimum podobieństwa minimalnego

$$h_{15}(q(X),X) \stackrel{d}{=} \max_{y \in V} \left[\min_{x \in X} (h(y,x)) \right]. \qquad (A.3.5)$$

W pewnych przypadkach rozpatrywanie podobieństwa elementu i zbioru elementów jako podobieństwa elementu i jednego reprezentanta tego zbioru jest trudne. Dla rozwiązywania takich przypadków przyjmuje się, że zbiór elementów jest reprezentowany przez zbiór reprezentantów $\{q(X)\}$. Dla zbioru reprezentantów funkcja podobieństwa między elementem y i tym zbiorem może być wyznaczana

$$h_{16}(y,\{q(X)\}) \stackrel{d}{=} \max_{x \in \{q(X)\}} (h(y,x)). \qquad (A.3.6)$$

Inne sposoby porównywania elementów ze zbiorami elementów, uwzględniające postać zbioru elementów (dokładniej: uwzględniające rozmieszczenie elementów tych zbiorów w przestrzeni) wykorzystują funkcję odległości elementu $y \in V$ od zbioru $X \subset V$ wyznaczaną według Mahalanobisa

$$d_8^2(y,X) \stackrel{d}{=} (y - q(X))\, C_X^{-1} (y - q(X))^T, \qquad (A.3.7)$$

gdzie:

C_X^{-1} - odwrotna macierz kowariancji elementów zbioru X.

Wykorzystanie odległości d_8 jest celowe wtedy, gdy moc zbioru X jest odpowiednio duża. Dla zbiorów X o małej mocy macierz kowariancji C_X powinna być zastąpiona macierzą kowariancji C_V, wyznaczoną dla wszystkich rozpatrywanych elementów przestrzeni V.

A.4. Porównywanie zbiorów ze zbiorami

W przestrzeni V dane są dwa zbiory X i Y. Znana jest funkcja h podobieństwa elementów przestrzeni V. Podobieństwo zbiorów X i Y można oceniać:

- na podstawie podobieństwa reprezentantów zbiorów X i Y

$$h_{17}(X,Y) \stackrel{d}{=} h(q(X), q(Y)), \qquad (A.4.1)$$

- na podstawie podobieństwa tzw. "najbliższych sąsiadów" tych zbiorów

$$h_{18}(X,Y) \stackrel{d}{=} \max_{(x,y) \in X \times Y} (h(x,y)), \qquad (A.4.2)$$

- na podstawie podobieństwa tzw. "najdalszych sąsiadów" w tych zbiorach

$$h_{19}(X,Y) \stackrel{d}{=} \min_{(x,y) \in X \times Y} (h(x,y)), \qquad (A.4.3)$$

- na podstawie średniego podobieństwa elementów tych zbiorów

$$h_{20}(X,Y) = \frac{1}{\overline{X}\,\overline{Y}} \sum_{(x,y) \in X \times Y} h(x,y). \qquad (A.4.4)$$

Dla porównywania zbiorów $X,Y \subset V$ można również wykorzystać odległość Mahalanobisa

$$d_9^2(X,Y) \stackrel{d}{=} (q(X) - q(Y))\, C_V^{-1} (q(X) - q(Y))^T \qquad (A.4.5)$$

lub

$$d_{10}^2(X,Y) \stackrel{d}{=} (q(X) - q(Y))(C_X^{-1} + C_Y^{-1})(q(X) - q(Y))^T. \qquad (A.4.6)$$

Podczas wyznaczania odległości d_9 uwzględniane jest rozmieszczenie wszystkich elementów w przestrzeni V, a podczas wyznaczania odległości d_{10} uwzględnia się rozmieszczenie elementów w zbiorze X i w zbiorze Y.

B. KRYTERIA GRUPOWANIA ELEMENTÓW PRZESTRZENI CECH

Ogólnie pod pojęciem grupowania elementów rozumie się wyznaczanie podziału pewnego zbioru elementów na podzbiory w świetle jakiegoś kryterium. Rozpatrywane będzie grupowanie elementów przestrzeni cech, którego potrzeba wynika z konieczności wyznaczania w tej przestrzeni zbiorów elementów podobnych po to, aby można było wnioskować o relacjach między tymi elementami.

Formalnie zadanie grupowania można sformułować następująco: "Niech V (przestrzeń wartości cech według rozdz. 2.4) będzie skończonym zbiorem elementów v. Należy wyznaczyć podział zbioru V na l podzbiorów V_k elementów podobnych".

Podział ten można zapisać w postaci rodziny $Q(V)$ podzbiorów V_k zbioru V

$$Q(V) = \left\{ V_k \subset V : k \in [1 : l] \right\}. \qquad (B.0.1)$$

Zbiór wszystkich możliwych (lub wszystkich dopuszczalnych) podziałów $Q(V)$ zbioru V na l podzbiorów oznaczany jest $Q^l(V)$. Podział Q będzie uznany za podział optymalny Q_{opt}, jeżeli funkcja kryterialna $e(Q)$ wynikająca z przyjętego kryterium osiąga dla tego podziału ekstremum. Opisane dalej funkcje kryterialne zostały zapisane tak, aby dla podziału optymalnego osiągały one wartości minimalne

$$Q = Q_{opt} \Longleftrightarrow e(Q) = \min_{Q_j \in Q^l} (e(Q_j)). \qquad (B.0.2)$$

Rozpatrywane kryteria umożliwiają optymalizację podziału zbioru V na zadaną liczbę podzbiorów i/lub optymalizację wyboru reprezentantów pod-

zbiorów. Są one mało skuteczne w zakresie optymalizacji liczby podzbiorów, na które dzielony jest zbiór V.

Istnieje duża liczba opracowań omawiających metody grupowania. Ogólnie można wyróżnić metody grupowania wynikające bezpośrednio lub pośrednio z metod statystycznej weryfikacji hipotez oraz heurystyczne i arbitralne metody grupowania realizowane ze względu na nie w pełni sformalizowane lub nawet nieznane kryteria. Większość metod znanych autorowi dotyczy wyznaczenia podziału zbioru na rozłączne podzbiory. Problemy grupowania przy założeniu, że wyznaczane podzbiory nie są rozłączne, rozpatrywane są przez niewielu autorów, np. [15]. Ogólnie przyjmuje się, że liczba wyznaczanych podzbiorów jest znana a priori (przed grupowaniem). Autorowi nie są znane skuteczne metody grupowania przy założeniu nieznanej liczby podzbiorów. Przyczyną braku takich metod jest to, że większość funkcji kryterialnych, których minimum ma wyznaczać optymalny podział, przyjmuje wartości malejące monotonicznie ze wzrostem liczby podzbiorów nie osiągając ekstremum dla liczby podzbiorów mniejszej od liczby elementów. Należy przypuszczać, że problem ten może znaleźć rozwiązanie przy wykorzystaniu algorytmów konwersacyjnych realizowanych w układzie "człowiek - komputer".

Większość autorów uznaje metody statystyczne za uprzywilejowane. Należy stwierdzić, że stanowisko takie, w dużej liczbie praktycznych zastosowań, jest nieuzasadnione. Stosując bowiem takie metody należy przyjmować dużą liczbę założeń - np. o postaci funkcji gęstości rozkładu elementów podzbiorów itp. Wykorzystując następnie złożony i formalnie w pełni poprawny aparat statystyki matematycznej otrzymuje się wynik, który może być uznany za poprawny, jeżeli przyjęto odpowiednie założenia, co najczęściej prowadzi do warunku, aby moce wyznaczanych podzbiorów były duże. Na podstawie badań własnych oraz badań prowadzonych w innych ośrodkach (badań diagnostycznych maszyn) można stwierdzić, że uwzględniana w procesie wnioskowania liczba elementów przestrzeni wartości cech nie jest duża i zawiera się najczęściej w zakresie od 10 do 50.

Dla rozwiązania postawionego zadania konieczna jest pełna formalizacja kryteriów oceny jakości grupowania. Z obserwacji sposobów rozwiązywania zadań grupowania przez człowieka należy stwierdzić, że stosowane są różne, subiektywnie wybierane kryteria, które ulegają zmianom nawet podczas rozwiązywania jednego zadania. Wielu autorów proponuje różne kryteria. Brak jest uzasadnionych ogólnych wniosków co do zalecanych zakresów ich zastosowań.

B.1. Kryteria wykorzystujące funkcje odległości

Najprostszą funkcją kryterialną [40] jest

$$e_1(Q) \triangleq \frac{1}{\bar{V}} \sum_{k=1}^{1} (\bar{V}_k \, d_k), \qquad (B.1.1)$$

gdzie:

Q - rodzina podzbiorów V_k wg (B.0.1),

d_k - średnia odległość elementów k-tego podzbioru V_k od reprezentanta tego podzbioru.

Bardziej złożone funkcje kryterialne określa się, wykorzystując metody analizy wariancji. Niech V będzie przestrzénią, w której określono odległość euklidesową. Wprowadza się:

- średni element $\bar{v}_k$ podzbioru V_k

$$\bar{v}_k = \frac{1}{\bar{V}_k} \sum_{v \in V_k} v, \tag{B.1.2}$$

- średni element $\bar{v}$ zbioru V

$$\bar{v} = \frac{1}{\bar{V}} \sum_{v \in V} v, \tag{B.1.3}$$

- macierz rozproszenia w podzbiorze V_k

$$R_k = \sum_{v \in V_k} (v - \bar{v}_k)^T (v - \bar{v}_k), \tag{B.1.4}$$

- sumaryczną macierz rozproszenia w rozłącznych podzbiorach zbioru V takich, że $V = \bigcup_k V_k$

$$R = \sum_{k=1}^{l} R_k, \tag{B.1.5}$$

- macierz rozproszenia między podzbiorami V_k zbioru V

$$B = \sum_{k=1}^{l} [\bar{V}_k (\bar{v}_k - \bar{v})^T (\bar{v}_k - \bar{v})], \tag{B.1.6}$$

- ogólną macierz rozproszenia dla zbioru V

$$T = \sum_{v \in V} (v - \bar{v})^T (v - \bar{v}). \tag{B.1.7}$$

Można wykazać [35], że

$$\mathbf{T} = \mathbf{R} + \mathbf{B}. \tag{B.1.8}$$

Ogólna macierz rozproszenia $\mathbf{T}$ nie zależy od podziału zbioru V na podzbiory. Macierze $\mathbf{R}$ i $\mathbf{B}$ zależą od tego podziału. Podział minimalizujący sumaryczną macierz $\mathbf{R}$ rozproszeń w podzbiorach maksymalizuje jednocześnie macierz rozproszeń $\mathbf{B}$ między podzbiorami. Jako miara "wielkości" macierzy podczas poszukiwania optymalnego podziału zbioru V może być wykorzystywana dowolna jej norma. Szczególnie zalecane jest zastosowanie w tym celu śladu $\mathrm{tr}(\mathbf{R})$ lub wyznacznika $\det(\mathbf{R})$ macierzy rozproszeń. Otrzymuje się funkcję kryterialną

$$e_2(Q) \triangleq \mathrm{tr}(\mathbf{R}) = \sum_{k=1}^{l} \mathrm{tr}(\mathbf{R}_k) = \sum_{k=1}^{l} \sum_{v \in V} (v - \bar{v}_k)(v - \bar{v}_k)^T. \tag{B.1.9}$$

Ślad $\mathrm{tr}(\mathbf{R})$ macierzy $\mathbf{R}$ jest proporcjonalny do sumy wariancji elementów zbioru V określanych w kierunkach odpowiadających kierunkom osi układu współrzędnych. Wyniki grupowania minimalizującego ten ślad są zależne od zmian skali współrzędnych. Inną postacią funkcji kryterialnej jest

$$e_3(Q) \triangleq \det(\mathbf{R}). \tag{B.1.10}$$

Warunkiem koniecznym dla stosowania tego kryterium jest nieosobliwość macierzy $\mathbf{R}$. Jeżeli istnieje potrzeba wykorzystania funkcji e_3 wtedy, gdy $\mathbf{R}$ jest macierzą osobliwą, to grupowanie należy poprzedzić zmniejszeniem liczby wymiarów przestrzeni V lub zmianą liczby klas. Funkcję e_3 można zastąpić przez równoważną jej funkcję kryterialną [39]

$$e_4(Q) \triangleq -\det(\mathbf{R}^{-1}\mathbf{B}). \tag{B.1.11}$$

Wyznaczenie wartości własnych $\lambda_1, \lambda_2, \ldots, \lambda_j$ macierzy $\mathbf{R}^{-1}\mathbf{B}$, będących rozwiązaniami równania

$$\det(\mathbf{B} - \lambda\mathbf{R}) = 0 \tag{B.1.12}$$

umożliwia zapisanie funkcji e_4 w postaci

$$e_4(Q) = -\prod_{i=1}^{m} \lambda_i, \tag{B.1.13}$$

gdzie:

$m = \dim(V)$.

Funkcja e_4 posiada niedogodność polegającą na tym, że osobliwość macierzy $R^{-1} B$ prowadzi do zerowej wartości tej funkcji. Niedogodności tej nie posiada funkcja kryterialna Wilksa [100]

$$e_5(Q) \triangleq - \prod_{i=1}^{m} (1 + \lambda_i). \qquad (B.1.14)$$

Zastępując w (B.1.11) wyznacznik macierzy $R^{-1} B$ jej śladem, otrzymuje się

$$e_6(Q) \triangleq -\mathrm{tr}(R^{-1} B) = - \sum_{i=1}^{m} \lambda_i. \qquad (B.1.15)$$

Fukcja kryterialna e_6 ocenia średnią odległość Mahalanobisa [63] średnich elementów $\bar{v}_k$ podzbiorów V_k od średniego elementu $\bar{v}$ zbioru V. Niedogodnością tej funkcji jest duży wpływ największej wartości własnej λ_{max} na wynik grupowania. Dla uzależnienia wartości funkcji kryterialnej od względnych zmian, nawet małych wartości własnych, wprowadza się [41]

$$e_7(Q) \triangleq \mathrm{tr}(T^{-1} R) = \sum_{i=1}^{m} \frac{1}{1 + \lambda_i}. \qquad (B.1.16)$$

Minimalizacja śladu macierzy $T^{-1} R$ jest równoważna minimalizacji śladu macierzy R dla zbioru V przekształconego do takiej postaci, że macierz T jest macierzą jednostkową. Wspólną zaletą kryteriów wykorzystujących wartości własne jest niezależność wyników grupowania od liniowych transformacji przestrzeni V. Praktyczna niedogodność tych kryteriów polega na dużej liczbie koniecznych do wykonania działań.

B.2. Kryteria wykorzystujące funkcję podobieństwa

Funkcja podobieństwa (według dodatku A.2) umożliwia określenie kryteriów maksymalizujących "spójność" podzbiorów. Jedną z najprostszych funkcji kryterialnych jest [15]

$$e_8(Q) \triangleq - \sum_{k=1}^{l} \sum_{(x,y) \in V_k \times V_k} h(x,y). \qquad (B.2.1)$$

Fritsche [40] proponuje

$$e_9(Q) \triangleq \sum_{k=1}^{l} s(V_k, V_k'), \tag{B.2.2}$$

gdzie:

V_k' - dopełnienie podzbioru V_k

$$V_k' = V \setminus V_k \tag{B.2.3}$$

oraz $s(\ )$ - funkcja spójności

$$s(V_k, V_k') = \frac{h(V_k, V_k')}{h(V_k, V_k) + h(V_k', V_k')} \tag{B.2.4}$$

Inne bardziej złożone zależności dla wyznaczania funkcji spójności podzbioru podano w [29].

Określono dużą liczbę kryteriów stosowanych dla oceny jakości klasyfikacji w przypadkach, gdy znane są rozkłady prawdopodobieństwa przynależności elementów $v \in V$ do podzbiorów V_k, V_k' [91]. Zastępując prawdopodobieństwa funkcjami podobieństwa elementów v do podzbiorów V_k, V_k' otrzymuje się wygodne do praktycznych zastosowań funkcje kryterialne będące:

- analogonem uogólnionej miary Kołmogorowa [91]

$$e_{10}(Q) \triangleq - \sum_{k=1}^{l} \sum_{v \in V_k} |h(v,V_k) - h(v,V_k')|^{\alpha}, \tag{B.2.5}$$

dla $\alpha > 0$ (gdzie najczęściej $\alpha = 1$);

- analogonem wskaźnika Bhattaharyya [91]

$$e_{11}(Q) \triangleq - \sum_{k=1}^{l} \sum_{v \in V_k} \sqrt{h(v,V_k)\ h(v,V_k')}, \tag{B.2.6}$$

- analogonem miary Sammona [84]

$$e_{12}(Q) \triangleq \sum_{k=1}^{l} \min_{v \in V_k} [h(v,V_k),\ h(v,V_k')]. \tag{B.2.7}$$

B.3. Kryteria progowe

Szczególną grupę kryteriów stanowią kryteria progowe. Kryteria te mogą być określane przy pomocy funkcji odległości lub funkcji podobieństwa elementów. Z wielu kryteriów tego typu wyróżnić można dwa:

- kryterium "najdalszego sąsiada"

$$\forall_{V_k \in Q_{opt}} \quad \forall_{x,y \in V_k} \quad [d(x,y) \leqslant d_{max}] \qquad (B.3.1)$$

- kryterium "najbliższego sąsiada"

$$\forall_{V_k \in Q_{opt}} \quad \forall_{x \in V_k} \quad \exists_{\substack{y \neq x \\ y \in V_k}} \quad [d(x,y) \leqslant d_{max}] \qquad (B.3.2)$$

lub podobnie dla $h(x,y) \geqslant h_{min}$.

Zaletą tych kryteriów jest ich prostota oraz duża zgodność wyników grupowania optymalnego z wynikami grupowania realizowanego arbitralnie przez człowieka. Kryterium "najbliższego sąsiada" uwypukla podobieństwa par elementów. Kryterium "najdalszego sąsiada" uwypukla wpływ tzw. "izolowanych elementów" przestrzeni V. Stanowi to pewną niedogodność ze względu na istnienie odchyleń elementów przestrzeni V wynikających z określania położenia tych elementów w przestrzeni V na podstawie wielkości wyznaczanych doświadczalnie. Dla częściowego uniezależnienia kryterium "najbliższego sąsiada" od tych lokalnych odchyleń stosuje się kryterium "n najbliższych sąsiadów"

$$\forall_{V_k \in Q_{opt}} \quad \forall_{x \in V_k} \quad \exists_{S_n(x) \subset V_k} \quad \forall_{y \in S_n(x)} \quad [d(x,y) \leqslant d_{max}] \qquad (B.3.3)$$

lub podobnie dla $h(x,y) \geqslant h_{min}$,

gdzie:

$S_n(x)$ - n-elementowe otoczenie elementu x.

Liczbę uwzględnianych n najbliższych sąsiadów przyjmuje się w zależności od mocy przestrzeni V oraz w zależności od liczby l poszukiwanych podzbiorów V_k.

Główną trudnością związaną z praktycznym wykorzystaniem opisanych kryteriów progowych jest konieczność przyjmowania wartości progowej h_{min} lub d_{max}. Przyjęcie tych wartości jest równoważne z przyjęciem liczby zbiorów, na które dzielona jest przestrzeń V.

C. ALGORYTMY GRUPOWANIA ELEMENTÓW PRZESTRZENI CECH

Należy przeprowadzić grupowanie elementów przestrzeni V, w celu wyznaczenia podziału tej przestrzeni na określoną liczbę zbiorów elementów podobnych. Duża liczba stosowanych algorytmów grupowania nie polega na rozwiązywaniu zadania w sposób wynikający bezpośrednio z przyjętego kryterium grupowania (por. dod. B). Algorytmy grupowania przyjmowane są często jako wynik pewnych działań heurystycznych.

Stwierdza się [102], że dla zbioru elementów dzielących się "wyraźnie" na rozłączne podzbiory, prawidłowo opracowane algorytmy prowadzą do jednoznacznych wyników podziału. W przypadku równomiernego rozkładu tych elementów w analizowanej przestrzeni różne algorytmy mogą prowadzić do różnych wyników. Większość zadań rozwiązywanych w praktyce zawiera się pomiędzy dwoma wyżej wymienionymi przypadkami. Skłania to do zalecenia stosowania równolegle kilku istotnie różnych procedur grupowania i podziału, w celu stwierdzenia czy wynik podziału uzależniony jest od zastosowanej metody, czy od dzielonego zbioru elementów.

Stosowane algorytmy można podzielić na kilka grup ze względu na istotę ich działania. Można wyróżnić:

- algorytmy poszukiwania ogólnego ekstremum funkcji kryterialnej,
- algorytmy hierarchicznego podziału i grupowania,
- algorytmy iteracyjne.

Zaletą algorytmów poszukiwania ogólnego ekstremum funkcji kryterialnej jest to, że prowadzą one do rozwiązań optymalnych w świetle przyjętego kryterium. Ich wadą jest konieczność wykonania dużej liczby działań, co zwiększa czas ich realizacji. Algorytmy hierarchicznego podziału i grupowania wymagają mniejszej liczby działań od algorytmów poszukiwania ogólnego ekstremum funkcji kryterialnej. Prowadzą one w ogólnym przypadku do rozwiązań, które mogą nie być rozwiązaniami optymalnymi w świetle przyjętego kryterium. Algorytmy iteracyjne wymagają określenia pewnego podziału zbioru V jako rozwiązania początkowego, który to podział może być wyznaczony losowo lub może być wynikiem działania innych algorytmów. W ogólnym przypadku dla większości algorytmów tego typu ostateczne rozwiązanie zależy od rozwiązania początkowego.

Pewne algorytmy wykorzystują metody funkcji potencjalnych [5]. Duża liczba znanych algorytmów polega na wyznaczaniu tzw. funkcji rozdzielających, określających analitycznie granice między podzbiorami, np.: [102]. Cechą charakterystyczną wielu opisanych w literaturze algorytmów jest jednoznaczne przyporządkowywanie każdego elementu dzielonego zbioru do jednego z wyznaczanych podzbiorów tego zbioru. Duża liczba algorytmów została opracowana dla przypadku, gdy grupowane są elementy przestrzeni cech jakościowych, które można opisywać binarną macierzą podobieństw. Przegląd tych

algorytmów można znaleźć w [29]. Znane są algorytmy grupowania rozmytego [12, 11, 19].

Można porównywać wyniki grupowania elementów zbioru z wynikami klasycznych metod weryfikacji hipotez wykorzystujących kryterium Bayesa lub podobne [15]. Przewagą metod grupowania i podziału nad metodami weryfikacji hipotez jest brak konieczności przyjmowania a priori pewnych założeń o rozwiązaniu. Skutkiem tego jest brak możliwości analitycznego oceniania prawdopodobieństwa błędu [94, 102]. Należy jednak zauważyć, że przyjmując jawne lub domyślne założenie, np. o postaci funkcji gęstości prawdopodobieństwa, można "obliczyć" prawdopodobieństwa błędów klasyfikacji, lecz ich wartość może nie mieć praktycznego znaczenia.

C.1. Algorytmy poszukiwania ogólnego ekstremum funkcji kryterialnej

Jedną z możliwych metod postępowania jest metoda polegająca na określeniu zbioru $Q^l(V)$ wszystkich możliwych podziałów przestrzeni V na rodziny $Q(V)$ podzbiorów V_k. Następnie dla tak wyznaczonych podziałów określa się wartości funkcji kryterialnej i za optymalny uznaje się podział, dla którego osiągnięto ekstremum tej funkcji. Realizacja algorytmu jest prosta. W praktyce algorytm ten nie jest stosowany ze względu na konieczność wykonania dużej liczby działań wynikającej z faktu, że dla przestrzeni V o mocy n można wyznaczyć

$$k = \frac{(n-1)!}{(l-1)!\,(n-l)!} \qquad (C.1.1)$$

podziałów na l zbiorów.

C.2. Algorytmy hierarchicznego podziału i grupowania

Istotą algorytmów hierarchicznego podziału jest poszukiwanie podziału przestrzeni V na dwa rozłączne zbiory V_1, V_2 elementów podobnych, który warunkuje osiągnięcie ekstremum przyjętej funkcji kryterialnej. Następnie z otrzymanych zbiorów wybiera się zbiór o mniejszej spójności, który podlega analogicznemu podziałowi. Postępując tak dalej dochodzi się do wymaganej liczby podzbiorów.

Istotą algorytmów hierarchicznego grupowania jest przyjęcie, że przestrzeń V elementów v o mocy n podzielona jest na n rozłącznych jednoelementowych zbiorów takich, że każdy element tej przestrzeni należy do innego zbioru. W rodzinie n zbiorów poszukuje się dwóch "najbardziej podobnych" zbiorów. Zbiory takie łączy się, w wyniku czego otrzymuje się rodzinę złożoną z $n-1$ zbiorów. W rodzinie $n-1$ zbiorów poszukuje się dwóch zbiorów najbardziej podobnych, które łączy się, w wyniku czego otrzymuje się rodzinę $n-2$ zbiorów. Postępując tak dalej dochodzi się do wymaganej liczby zbiorów.

Podczas realizowania algorytmów polegających na hierarchicznym dzieleniu zbiorów przyjmuje się ograniczenia w postaci minimalnej mocy zbiorów wyznaczanych w wyniku podziału. Skutkiem takiego ograniczenia będzie niewykonywanie podziału zbioru, jeżeli w wyniku podziału otrzyma się zbiory o mocach mniejszych od założonej mocy minimalnej

$$\overline{V}_k \geq \overline{V}_{k,min} = \max\left(\frac{V}{\gamma l}; \gamma\right), \qquad (C.2.1)$$

gdzie:

l - jest liczbą zbiorów, na które należy podzielić przestrzeń V,

γ - jest stałą (przyjmowaną np. w zakresie od 2 do 4).

Podobnie podczas realizowania algorytmów hierarchicznego grupowania przyjmować można ograniczenia w postaci maksymalnej mocy wyznaczanych zbiorów

$$\overline{V}_k \leq \overline{V}_{k,max} = \min\left(\frac{\gamma \overline{V}}{l}; \overline{V}\left(1 - \frac{l-1}{\gamma l}\right)\right). \qquad (C.2.2)$$

Metody hierarchicznego podziału posiadają wadę wynikającą z faktu, że zbiory dzielone są na dwie części. Na przykład, jeżeli poszukiwany jest podział zbioru liczb $\{1, 2, 3, 4, 5, 6\}$ na trzy podzbiory, to rozwiązaniem optymalnym w świetle kryterium najbliższego sąsiada jest rodzina $Q = \{\{1,2\}, \{3,4\}, \{5,6\}\}$. Stosując metodę hierarchicznego podziału otrzymuje się po pierwszym podziale dwa zbiory $\{1, 2, 3\}$, $\{4, 5, 6\}$, co wyklucza możliwość otrzymania rozwiązania optymalnego. Dla uniknięcia tej niedogodności stosuje się algorytmy kombinowane dokonujące podziału i grupowania.

Stwierdzenia wielu autorów sugerują brak istotnych różnic w zaletach algorytmów hierarchicznego podziału i algorytmów hierarchicznego grupowania. W wyniku badań własnych stwierdzono, że za wystarczająco skuteczne dla uzyskania optymalnego podziału przestrzeni V na l zbiorów elementów podobnych należy uznać postępowanie polegające na hierarchicznym podziale przestrzeni V na około 2 l zbiorów, a następnie hierarchiczne grupowanie (łączenie) tak otrzymanych zbiorów w celu wyznaczenia wymaganych l zbiorów.

Dla opisania szczegółowych algorytmów hierarchicznego podziału wystarczy określić sposób dzielenia zbioru V na dwa podzbiory V_1, V_2 elementów podobnych, który powtarzany odpowiednią liczbę razy prowadzi do rozwiązania. Typowy algorytm hierarchicznego podziału zbioru V na dwa podzbiory składa się z następujących pięciu faz:

1° - zakłada się $V_1 = \emptyset$, $V_2 = V$,

2° - poszukuje się elementu $v_m \in V_2$ minimalizującego wartość funkcji podobieństwa $h(v_m, V_2 \setminus \{v_m\})$,

3° - element v_m przenosi ze zbioru V_2 do V_1

$$V_1 := V_1 \cup \{v_m\}, \quad V_2 := V_2 \setminus \{v_m\}, \tag{C.2.3}$$

4° - poszukuje się elementu $v_m \in V_2$ maksymalizującego funkcję

$$e(v_m) = h(v_m, V_1) - h(v_m, V_2 \setminus \{v_m\}), \tag{C.2.4}$$

5° - jeżeli $e(v_m) > 0$ to przechodzi się do punktu 3°; jeżeli $e(v_m) \leq 0$ to V_1, V_2 wyznaczają poszukiwany podział zbioru V na dwa podzbiory.

Ciekawy algorytm hierarchicznego podziału opisany został w [105]. Według tego algorytmu:

1° - losuje się element $v_1 \in V$,

2° - poszukuje się elementu $v_2 \in V \setminus \{v_1\}$, który minimalizuje wartość funkcji podobieństwa $h(v_1, v_2)$,

3° - wyznacza się przybliżenia V_1^*, V_2^*

$$V_1^* = \{v \in V : h(v_1, v) > h(v_2, v)\}, \quad V_2^* = V \setminus V_1^*, \tag{C.2.5}$$

4° - dla wyznaczonych przybliżeń wyznacza się reprezentantów $q(V_1^*)$, $q(V_2^*)$ maksymalizujących wartości funkcji podobieństwa

$$h(q(V_1^*), V_1^* \setminus \{q(V_1^*)\}), \quad h(q(V_2^*), V_2^* \setminus \{q(V_2^*)\}),$$

5° - jeżeli $q(V_1^*) \neq v_1$ lub $q(V_2^*) \neq v_2$, to przyjmuje się $v_1 = q(V_1^*)$ i $v_2 = q(V_2^*)$ oraz przechodzi się do punktu 3°,

- jeżeli $q(V_1^*) = v_1$ i $q(V_2^*) = v_2$, to $V_1 = V_1^*$ i $V_2 = V_2^*$ są optymalnymi rozwiązaniami zadania podziału zbioru V na dwa podzbiory.

Algorytm ten jest przykładem wykorzystania metod iteracyjnych w hierarchicznym wyznaczaniu podziału zbioru.

C.3. Algorytmy iteracyjne

Istotę algorytmów iteracyjnych można pokazać na przykładzie algorytmu będącego uproszczeniem algorytmu opisanego w [8].

1° - wybiera się elementy, które uznaje się za przybliżenie reprezentantów $q(V_1), \ldots, q(V_1)$,

2° - przeprowadza się klasyfikację elementów przestrzeni V polegającą na przyporządkowaniu każdego elementu do tego zbioru wyznaczonego przez reprezentanta, dla którego osiąga się maksimum funkcji podobieństwa,

3° - dla tak otrzymanych zbiorów wyznacza się nowych reprezentantów,

4° - jeżeli wyznaczeni nowi reprezentanci zbiorów różnią się od reprezentantów poprzednich, to powraca się do punktu 2°, który realizuje się, uwzględniając nowych reprezentantów.

Podstawową niedogodnością przedstawionego algorytmu jest konieczność wykonywania dużej liczby działań związanych z wyznaczaniem kolejnych reprezentantów zbiorów oraz konieczność porównywania z tymi reprezentantami wszystkich elementów dzielonej przestrzeni.

Dla zmniejszenia liczby działań można zastosować algorytm opisany w[35]. Za optymalny przyjęto podział przestrzeni V na rodzinę $Q(V)$ zbiorów V_k, minimalizujący funkcję e_2 wg (B.1.9)

$$e_2(Q) = \sum_{k=1}^{l} \sum_{v \in V_k} (v - \bar{v}_k)(v - \bar{v}_k)^T, \qquad (C.3.1)$$

gdzie:

$\bar{v}_k$ - średni element zbioru V_k wg (B.1.2).

Niech dany będzie pewien początkowy podział przestrzeni V. Poszukując podziału optymalnego przenosi się element $v_m \in V_k$ ze zbioru o mocy $\bar{V}_k > 1$ do zbioru V_j w wyniku czego powstają dwa nowe zbiory

$$V_k^* = V_k \setminus \{v_m\}, \quad V_j^* = V_j \cup \{v_m\}, \qquad (C.3.2)$$

takie, że

$$\bar{V}_k^* = \bar{V}_k - 1, \qquad \bar{V}_j^* = \bar{V}_j + 1, \qquad (C.3.3)$$

$$\bar{v}_k^* = \bar{v}_k - \frac{1}{\bar{V}_k^*}(v_m - \bar{v}_k),$$

$$\bar{v}_j^* = \bar{v}_j + \frac{1}{\bar{V}_j^*}(v_m - \bar{v}_j). \qquad (C.3.4)$$

Z (C.3.1) wynika

$$e(Q^*) - e(Q) = \frac{\bar{V}_j}{\bar{V}_j^*}(v_m - \bar{v}_j)(v_m - \bar{v}_j)^T -$$

$$- \frac{\bar{V}_k}{\bar{V}_k^*}(v_m - \bar{v}_k)(v_m - \bar{v}_k)^T. \qquad (C.3.5)$$

Istotą algorytmu wg [35] jest badanie znaku różnicy (C.3.5) i przenoszenie ze zbioru V_k do zbioru V_j tych elementów v_m, dla których różnica ta jest ujemna, co zapewnia minimalizację funkcji kryterialnej.

Znane są różne wersje opisanego wyżej algorytmu nazywanego przez niektórych autorów algorytmem adaptacyjnym. Wersje te różnią się postacią przyjmowanej funkcji kryterialnej.

Wadą większości algorytmów iteracyjnych jest duże uzależnienie wyników klasyfikacji od lokalnych ekstremów funkcji kryterialnej czyli duże uzależnienie wyników klasyfikacji od przyjętego podziału początkowego i kolejności obliczeń. Dla uniknięcia tych wad można kilkakrotnie powtarzać realizację algorytmu dla różnych, losowo wybieranych podziałów początkowych oraz dla losowo ustalonej kolejności porównywania zbiorów V_k, V_j. Brak ogólnych zaleceń co do sposobu ustalania podziału początkowego.

Do najpoważniejszych zalet algorytmów iteracyjnych należy możliwość określenia prostego algorytmu, którego celem będzie adaptacja podziału przestrzeni V uwzględniająca dołączenie nowych elementów do podzielonej wcześniej przestrzeni.

Odrębną grupę algorytmów iteracyjnych stanowią algorytmy wykorzystujące bezpośrednio gradient funkcji kryterialnej. Przykładem może być algorytm wykorzystujący gradient funkcji gęstości prawdopodobieństwa warunkowego [42]. Istotą algorytmu jest iteracyjne przemieszczanie grupowanych elementów w kierunku największego gradientu tej funkcji. Algorytm ten prowadzi do wyznaczenia reprezentantów zbiorów bez konieczności (co jest szczególną zaletą algorytmu) zakładania liczby zbiorów przed grupowaniem. W odróżnieniu od metod hierarchicznego grupowania elementów, elementy w tej metodzie postępowania nie są łączone, lecz przemieszczane w kierunku reprezentantów zbiorów. Przyjmuje się następujący algorytm iteracyjny [42]

$$ {}^{(0)}v_n := v_n', \qquad n = 1,2,\ldots,\bar{V}, \tag{C.3.6} $$

$$ {}^{(i+1)}v_n := {}^{(i)}v_n + a\nabla_v(\ln p({}^{(i)}v_n)), \tag{C.3.7} $$

gdzie:

$ {}^{(i)}v_n $ - jest n-tym elementem przestrzeni V określonym w wyniku i-tego przemieszczania (i-tego kroku algorytmu).

Na zwrócenie uwagi zasługuje tutaj zastosowanie normalizowanego gradientu

$$ \nabla_v(\ln p(v)) = \frac{\nabla_v p(v)}{p(v)}. \tag{C.3.8} $$

Zaletą tak określonego gradientu jest to, że dla lokalnych ekstremów p(v), dla których $p(v) \ll \max(p(v))$

$$\nabla_v(\ln p(v)) > \nabla_v(p(v)), \tag{C.3.9}$$

co umożliwia uwzględnianie podczas grupowania tych właśnie ekstremów.

Gdy postać analityczna p(v) nie jest znana, autorzy algorytmu proponują zastąpienie (C.3.7) przez

$$^{(i+1)}v_n := \frac{1}{I} \sum_{x \in S_1(^{(i)}v_n)} x, \tag{C.3.10}$$

co powoduje, że w i-tym kroku działania algorytmu każdy element $^{(i)}v_n$ przekształcany jest w średnią elementów z pewnego jego 1-elementowego otoczenia S_1. Ważnym problemem jest wybór mocy tego otoczenia. Przyjęcie zbyt małego otoczenia prowadzi do niewłaściwych wniosków, ponieważ uwzględniane są wyłącznie podobieństwa elementów sąsiadujących ze sobą. Przyjęcie zbyt dużego otoczenia prowadzi do niewłaściwych wniosków, ponieważ uwzględniane mogą być jednocześnie wszystkie elementy zbioru, co prowadzi do ogólnej średniej.

Można rozpatrywać różne modyfikacje pokazanego algorytmu. Na podstawie badań własnych stwierdzono celowość wyznaczania średniej (C.3.10) jako średniej ważonej, gdzie wagami są wartości funkcji podobieństwa

$$^{(i+1)}v_n = \sum_{x \in {}^{(i)}V} (h(x, {}^{(i)}v_n)x) \Big/ \sum_{x \in {}^{(i)}V} h(x, {}^{(i)}v_n). \tag{C.3.11}$$

Wadą takiej modyfikacji algorytmu jest zwiększenie liczby wymaganych obliczeń.

C.4. Algorytmy wykorzystujące zbiory rozmyte

Uogólniając algorytm opisany w [19] można określić iteracyjny algorytm optymalizacji podziału przestrzeni V, którego wynikiem jest rodzina pseudorozłącznych zbiorów rozmytych.

Definicja C.1: Rodzina zbiorów rozmytych

$$\{\tilde{A}\} = \{\tilde{A}_k = \{(v, a_k(v)) : v \in V\} : k \in [1:1]\}$$

w przestrzeni V jest rodziną pseudorozłącznych zbiorów rozmytych wtedy i tylko wtedy, gdy

$$\bigwedge_{v \in V} [\sum_{k=1}^{1} a_k(v) \leq 1]. \tag{C.4.1}$$

Podział przestrzeni V uznany będzie za optymalny, jeżeli osiągnięte zostanie minimum funkcji kryterialnej. Niech ${}^{(t)}\tilde{A}_k$ oznacza poszukiwany k-ty zbiór rozmyty po t-tym kroku iteracji. Określa się przekształcenie wartości funkcji charakterystycznej zbiorów rozmytych

$$\omega_m : {}^{(t)}a_k(v) \longmapsto {}^{(t+1)}a_k(v) = \begin{cases} \alpha_i + (1-\alpha_i)\,{}^{(t)}a_k(v) & \text{dla } k = m \\ (1-\alpha_i)\,{}^{(t)}a_k(v) & \text{dla } k \neq m \end{cases} \qquad (C.4.2)$$

gdzie:

$0 \leqslant \alpha_i \leqslant 1$ - jest przyjmowanym parametrem.

Opisywany algorytm składa się z następujących faz:

1° - przyjmuje się rozwiązanie początkowe

$${}^{(0)}a_k(v) = 1/l, \qquad (C.4.3)$$

- oblicza się wartość funkcji kryterialnej $e({}^{(0)}\tilde{A})$,

2° - kolejno dla wszystkich elementów $v_j \in V$ i kolejno dla wszystkich wyznaczanych zbiorów rozmytych $\tilde{A}_k$:

- przekształca się funkcje charakterystyczne zbiorów rozmytych

$$a_k^*(v_j) = \omega_m({}^{(t)}a_k(v_j)), \qquad (C.4.4)$$

- oblicza się wartość funkcji kryterialnej $e(\{\tilde{A}^*\})$,

- jeżeli $e(\{\tilde{A}^*\}) < e(\{{}^{(t)}\tilde{A}\})$, to ${}^{(t+1)}a_k := a_k^*$ dla $k = 1,\ldots,l$ i powraca się do początku punktu 2°,

3° - jeżeli $e(\{\tilde{A}^*\}) \geqslant e(\{{}^{(t)}\tilde{A}\})$ dla wszystkich $m = 1,\ldots,l$, to rodzina $\{{}^{(t)}\tilde{A}\}$ wyznacza optymalny podział przestrzeni V.

Podczas praktycznej realizacji opisanego algorytmu występują trudności związane z koniecznością minimalizacji czasu wykonywania dużej liczby działań.

Przyjmowane są różne postacie przekształcenia (C.4.2) (por. np. [12]) uwzględniające stosowaną funkcję kryterialną, co pozwala na ograniczenie liczby działań związanych z wyznaczaniem wartości tej funkcji.

D. ZBIORY I RELACJE ROZMYTE

Pojęcie zbioru rozmytego wprowadził w latach sześćdziesiątych L.A. Zadeh [103]. Koncepcja L.A. Zadeha była rozwijana. Obecnie literatura dotycząca zbiorów rozmytych jest bardzo liczna; ciekawsze pozycje to [47, 99, 48,49,55,86]. Z opracowań krajowych na ten temat można wymienić [28]. Znane są zastosowania zbiorów i relacji rozmytych w diagnostyce medycznej (np. [37, 87, 86] oraz w diagnostyce technicznej (np. [17, 70]).

D.1. Pojęcie zbioru i relacji rozmytej

Definicja D.1.1: Zbiór rozmyty $\tilde{A}$ w przestrzeni V jest zbiorem par uporządkowanych

$$\tilde{A} \triangleq \{(v, a(v)) : v \in V\}, \qquad (D.1.1)$$

gdzie:

a(v) - funkcja charakterystyczna zbioru rozmytego $\tilde{A}$ (nazywana też funkcją przynależności do zbioru $\tilde{A}$) taka, że

$$a : V \rightarrow [0 ; 1] \subset R^1. \qquad (D.1.2)$$

Wartość funkcji charakterystycznej a(v) dla elementu v określa "stopień przynależności" tego elementu do zbioru rozmytego $\tilde{A}$. Przyjmuje się następującą interpretację wartości funkcji charakterystycznej:

a(v) = 0 - oznacza, że element v nie należy do $\tilde{A}$,

a(v) = 1 - oznacza, że element v należy do $\tilde{A}$,

$0 < a(v) < 1$ - oznacza, że element v należy "częściowo" do zbioru $\tilde{A}$.

Szczególnym przypadkiem zbioru rozmytego jest zbiór w znaczeniu klasycznym. Odpowiada on zbiorowi rozmytemu, którego funkcja charakterystyczna przyjmuje wyłącznie dwie wartości: 0 i 1.

Definicja D.1.2: Zbiór rozmyty $\tilde{A} = \{(v,a(v)) : v \in V\}$ jest zbiorem skończonym, jeżeli V jest zbiorem o skończonej liczbie elementów.

Wartości funkcji charakterystycznej skończonego zbioru rozmytego A można zapisywać w postaci macierzy, nazywanej macierzą charakterystyczną zbioru rozmytego

$$a^n = [a(v_1), a(v_2), \ldots, a(v_n)]. \qquad (D.1.3)$$

Definicja D.1.3: Dopełnieniem zbioru rozmytego $\tilde{A} = \{(v, a(v)) : v \in V\}$ jest zbiór rozmyty $\tilde{A}'$ taki, że

$$\tilde{A}' \triangleq \{(v\ a'(v)) : v \in V;\ a'(v) = 1 - a(v)\}. \qquad (D.1.4)$$

Definicja D.1.4: Moc $\bar{\tilde{A}}$ zbioru rozmytego $\tilde{A} = \{(v, a(v)) : v \in V\}$ określana jest zgodnie z zależnością

$$\bar{\tilde{A}} \triangleq \sum_{v \in V} a(v) < \bar{\bar{V}}. \qquad (D.1.5)$$

Definicja D.1.5: Zakres $\delta(\tilde{A})$ zbioru rozmytego $\tilde{A} = \{(v, a(v)) : v \in V\}$ jest odcinkiem uporządkowanego zbioru liczb rzeczywistych

$$\delta(\tilde{A}) = [\min_{v \in V} (a(v)) : \max_{v \in V} (a(v))] \subset [0:1]. \qquad (D.1.6)$$

Definicja D.1.6: Zbiór rozmyty $\tilde{A} = \{(v, a(v)) : v \in V\}$ jest zawarty w zbiorze rozmytym $\tilde{B} = \{(v, b(v) : v \in V\}$, co zapisujemy $\tilde{A} \subset \tilde{B}$, wtedy i tylko wtedy, gdy

$$\forall_{v \in V} [a(v) \leq b(v)]. \qquad (D.1.7)$$

Definicja D.1.7: Zbiór rozmyty $\tilde{A}$ jest równy zbiorowi rozmytemu $\tilde{B}$, co zapisujemy $\tilde{A} = \tilde{B}$, wtedy i tylko wtedy, gdy $\tilde{A} \subset \tilde{B}$ i $\tilde{B} \subset \tilde{A}$ tzn. wtedy, gdy

$$\forall_{v \in V} [a(v) = b(v)]. \qquad (D.1.8)$$

Definicja D.1.8: α- obcięcie zbioru rozmytego $\tilde{A} = \{(v, a(v)) : v \in V\}$, oznaczane $A_{>\alpha}$, jest zbiorem

$$A_{>\alpha} \triangleq \{v \in V : a(v) \geq \alpha\} \qquad (D.1.9)$$

Definicja D.1.9: Relacja rozmyta $\tilde{R}$ między dwoma niepustymi zbiorami X i Y, zapisywana (X) $\tilde{R}$ (Y), jest zbiorem rozmytym określonym na iloczynie kartezjańskim $X \times Y$

$$\tilde{R} \triangleq \{((x,y), r(x,y)) : (x,y) \in X \times Y\}. \qquad (D.1.10)$$

Wartości funkcji charakterystycznej relacji rozmytej określonej między skończonymi zbiorami X i Y, można zapisywać w postaci macierzy, nazywanej macierzą charakterystyczną

$$R^{n,m} = \begin{bmatrix} r(x_1,y_1),\ldots,r(x_1,y_m) \\ \ldots\ldots\ldots,\ldots,\ldots\ldots\ldots \\ r(x_n,y_1),\ldots,r(x_n,y_m) \end{bmatrix} . \qquad (D.1.11)$$

Definicja D.1.10: Relacja $\tilde{R}^T$ przestawiona (transponowana) względem relacji $\tilde{R}$ wg (D.1.10), zapisywana (Y) $\tilde{R}^T$ (X), jest zbiorem rozmytym

$$R^T \triangleq \{((y,x),\ r(x,y)) : (y,x) \in Y \times X\}. \qquad (D.1.12)$$

Definicja D.1.11: Zbiór rozmyty $\tilde{A} = \{(v,a(v)) : v \in V\}$ jest zbiorem "dobrze określonym" w przestrzeni V wtedy, gdy spełniony jest warunek

$$\min_{v \in V}(a(v)) + \max_{v \in V}(a(v)) = 1. \qquad (D.1.13)$$

Definicja D.1.12: Dwa zbiory rozmyte $\tilde{A} = \{(v,\ a(v)) : v \in V\}$ i $\tilde{B} = \{(v,\ b(v)) : v \in V\}$ określone w przestrzeni V są zbiorami podobnymi $\tilde{A} \sim \tilde{B}$ wtedy, gdy

$$\forall_{(x,y) \in V \times V} \ [a(x) > a(y) \iff b(x) > b(y)] . \qquad (D.1.14)$$

D.2. Działania

Rozpatrywane będą następujące zbiory rozmyte

$$\begin{aligned} \tilde{A} &= \{(v,\ a(v)) : v \in V\}, \\ \tilde{B} &= \{(v,\ b(v)) : v \in V\}, \\ \tilde{C} &= \{(v,\ c(v)) : v \in V\}, \end{aligned} \qquad (D.2.1)$$

Definicja D.2.1: Dla wartości funkcji charakterystycznych wprowadza się następujące działania dwuargumentowe

- suma

$$c(v) = a(v) \vee b(v) \triangleq \begin{cases} a(v) & \text{dla} \quad a(v) \geqslant b(v), \\ b(v) & \text{dla} \quad b(v) > a(v), \end{cases} \qquad (D.2.2)$$

- iloczyn

$$c(v) = a(v) \wedge b(v) \stackrel{d}{=} \begin{cases} a(v) & \text{dla } a(v) \leq b(v), \\ b(v) & \text{dla } b(v) < a(v), \end{cases} \qquad \text{(D.2.3)}$$

- działanie α

$$c(v) = a(v) \alpha b(v) \stackrel{d}{=} \begin{cases} 1 & \text{dla } a(v) \leq b(v), \\ b(v) & \text{dla } a(v) > b(v), \end{cases} \qquad \text{(D.2.4)}$$

- działanie ε

$$c(v) = a(v) \varepsilon b(v) \stackrel{d}{=} \begin{cases} b(v) & \text{dla } a(v) < b(v), \\ 0 & \text{dla } a(v) \geq b(v). \end{cases} \qquad \text{(D.2.5)}$$

Działania wprowadzone dla wartości funkcji charakterystycznych są podstawą określenia działań na zbiorach rozmytych, zapisywanych ogólnie

$$\tilde{C} = \tilde{A} \square \tilde{B} \Rightarrow \tilde{C} \stackrel{d}{=} \{(v, c(v)) : v \in V; \quad c(v) = a(v) \square b(v)\} \qquad \text{(D.2.6)}$$

lub stosując zapis wartości funkcji charakterystycznych w postaci macierzy

$$\mathbf{c} = \mathbf{a} \square \mathbf{b}, \qquad \text{(D.2.7)}$$

gdzie:

$\square$ - oznacza jeden z operatorów $\vee$, $\wedge$, α, ε.

Ilustrację wprowadzonych działań pokazano na rys. D.1. Zbiory $\tilde{A}$ i $\tilde{B}$ rozpatrywane dotychczas jako argumenty działań były określone w tej samej przestrzeni V. Dla zbiorów rozmytych

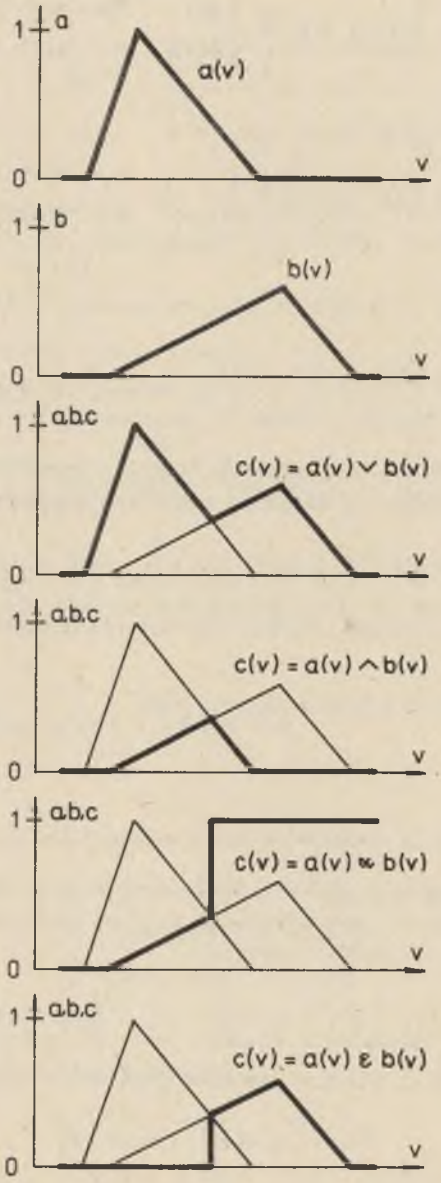
$$\tilde{A} = \{(x, a(x)) : x \in X\}, \qquad \tilde{B} = \{(y, b(y)) : y \in Y\} \qquad \text{(D.2.8)}$$

określa się działania, których wynikiem jest zbiór rozmyty

$$\tilde{C} = \{(z, c(z)) : z \in X \times Y\} \qquad \text{(D.2.9)}$$

w przestrzeni będącej iloczynem kartezjańskim przestrzeni argumentów. Działania te zapisywane są, przy zastosowaniu oznaczenia wprowadzonego dla relacji transponowanej, następująco

$$\tilde{C} = \tilde{A}^T \square \tilde{B} \Rightarrow \tilde{C} \stackrel{d}{=} \{(z, c(z)) : z \in X \times Y, \; c(z) = a(x) \square b(y)\} \qquad \text{(D.2.10)}$$



Rys. D.1. Ilustracja działań $\vee$, $\wedge$, $\propto$, ε. Funkcje charakterystyczne a(v), b(v) będące argumentami oraz funkcje charakterystyczne c(v) będące wynikami działań

lub dla zapisu przy pomocy macierzy

$$\mathbf{C} = \mathbf{a}^{T} \square \mathbf{b}, \tag{D.2.11}$$

gdzie:

$\square$ - oznacza jeden z operatorów $\vee$, $\wedge$, α, ε.

Niezależnie od działań określonych dla zbiorów rozmytych wprowadza się działania określone dla relacji rozmytych (składanie relacji). Niech dane będą trzy niepuste zbiory X, Y, Z oraz dwie relacje rozmyte

$$\begin{aligned} \tilde{Q} &= \{((x,y),\ q(x,y)) : (x,y) \in X \times Y\}, \\ \tilde{R} &= \{((y,z),\ r(y,z)) : (y,z) \in Y \times Z\} \end{aligned} \tag{D.2.12}$$

i relacja rozmyta

$$\tilde{T} = \{((x,z),\ t(x,z)) : (x,z) \in X \times Z\} \tag{D.2.13}$$

będąca złożeniem relacji $\tilde{Q}$ i $\tilde{R}$.

Definicja D.2.2: Wprowadza się cztery sposoby składania relacji, oznaczane operatorami $\circ$, $\triangle$, $\circledcirc$, $\circledcirc$, takie, że

$$\tilde{T} = \tilde{Q} \circ \tilde{R} \Rightarrow$$
$$\Rightarrow t(x,z) = (q \circ r)(x,z) \overset{d}{=} \bigvee_{y \in Y} [q(x,y) \wedge r(y,z)], \tag{D.2.14}$$

$$\tilde{T} = \tilde{Q} \triangle \tilde{R} \Rightarrow$$
$$\Rightarrow t(x,z) = (q \triangle r)(x,z) \overset{d}{=} \bigwedge_{y \in Y} [q(x,y) \vee r(y,z)], \tag{D.2.15}$$

$$\tilde{T} = \tilde{Q} \circledcirc \tilde{R} \Rightarrow$$
$$\Rightarrow t(x,z) = (q \circledcirc r)(x,z) \overset{d}{=} \bigwedge_{y \in Y} [q(x,y) \alpha r(y,z)], \tag{D.2.16}$$

$$\tilde{T} = \tilde{Q} \circledcirc \tilde{R} \Rightarrow$$
$$\Rightarrow t(x,z) = (q \circledcirc r)(x,z) \overset{d}{=} \bigvee_{y \in Y} [q(x,y) \varepsilon r(y,z)]. \tag{D.2.17}$$

Zbiór rozmyty można traktować jako szczególny przypadek relacji rozmytej. Pozwala to na wykonywanie działań wprowadzonych jako złożenia relacji, na argumentach, z których jeden jest zbiorem rozmytym. Niech

$$\tilde{A} = \{(x, a(x)) : x \in X\},$$
$$\tilde{B} = \{(y, b(y)) : y \in Y\}, \qquad (D.2.18)$$
$$\tilde{Q} = \{((x,y), q(x,y)) : (x,y) \in X \times Y\}.$$

Definicja D.2.3: Wprowadza się następujące przekształcenie zbioru rozmytego $\tilde{A}$ przez relację rozmytą $\tilde{Q}$ w zbiór rozmyty $\tilde{B}$:

$$\tilde{B} = \tilde{A} \circ \tilde{Q} \Rightarrow b(y) \triangleq \bigvee_{x \in X} [a(x) \wedge q(x,y)], \qquad (D.2.19)$$
$$\Rightarrow b = a \circ Q,$$

$$\tilde{B} = \tilde{A} \triangle \tilde{Q} \Rightarrow b(y) \triangleq \bigwedge_{x \in X} [a(x) \vee q(x,y)], \qquad (D.2.20)$$
$$\Rightarrow b = a \triangle Q,$$

$$\tilde{B} = \tilde{A} \textcircled{\scriptsize @} \tilde{Q} \Rightarrow b(y) \triangleq \bigwedge_{x \in X} [a(x) \propto q(x,y)], \qquad (D.2.21)$$
$$\Rightarrow b = a \textcircled{\scriptsize @} Q,$$

$$\tilde{B} = \tilde{A} \textcircled{\scriptsize o} \tilde{Q} \Rightarrow b(y) \triangleq \bigvee_{x \in X} [a(x) \delta q(x,y)], \qquad (D.2.22)$$
$$\Rightarrow b = a \textcircled{\scriptsize o} Q.$$

D.3. Relacyjne równania rozmyte

Dla określania związków między zbiorami rozmytymi można wykorzystać tzw. relacyjne równania rozmyte. Najczęściej równania takie zapisywane są w postaci (por. np. [28])

$$\tilde{A} \circ \tilde{X} = \tilde{B}, \qquad (D.3.1)$$

gdzie:

$\tilde{A}$, $\tilde{B}$ - znane zbiory rozmyte,

$\tilde{X}$ - poszukiwana relacja.

Rozważania związane z wyznaczaniem rozwiązań równania (D.3.1) wymagają jawnego lub domyślnego przyjęcia założenia, że rozwiązanie takie istnieje. Dla uniknięcia wynikającej stąd niedogodności oraz dla umożliwienia rozwiązywania układów równań przyjęto w tej pracy bardziej ogólną postać podstawowego relacyjnego równania rozmytego

$$\tilde{A} \circ \tilde{X} \subset \tilde{B} \subset \tilde{A} \triangle \tilde{Y}, \qquad (D.3.2)$$

gdzie:

$\tilde{A}$, $\tilde{B}$ - znane (porównywane) zbiory rozmyte,

$\tilde{X}$, $\tilde{Y}$ - poszukiwane relacje.

Wymaga się, aby poszukiwane relacje $\tilde{X}$ i $\tilde{Y}$ (będące rozwiązaniami równania (D.3.2)) spełniały następujące kryteria:

- relacja $\tilde{X}$ jest relacją maksymalną, tzn. taką, że każda relacja $\tilde{X}_i$ spełniająca równanie (D.3.2) jest zawarta w poszukiwanej relacji maksymalnej

$$\tilde{A} \circ \tilde{X}_i \subset \tilde{B} \Rightarrow \tilde{X}_i \subset \tilde{X}, \tag{D.3.3}$$

- relacja $\tilde{Y}$ jest relacją minimalną, tzn. taką, która jest zawarta w każdej relacji $\tilde{Y}_i$ spełniającej równanie (D.3.2)

$$\tilde{B} \subset \tilde{A} \triangle \tilde{Y}_i \Rightarrow \tilde{Y} \subset \tilde{Y}_i. \tag{D.3.4}$$

Wykorzystując schemat dowodu wg [85] można wykazać, że rozwiązaniami równania (D.3.2) są relacje

$$\tilde{X} = \tilde{A}^T \alpha \tilde{B} \quad \text{oraz} \quad \tilde{Y} = \tilde{A}^T \varepsilon \tilde{B} \tag{D.3.5}$$

lub w zapisie przy pomocy macierzy charakterystycznych

$$X = a^t \alpha b \quad \text{oraz} \quad Y = a^T \varepsilon b. \tag{D.3.6}$$

Rozwiązując (D.3.2) wyznacza się relacje $\tilde{X}$ i $\tilde{Y}$ zgodnie z (D.3.5) na podstawie znanych zbiorów rozmytych $\tilde{A}$ i $\tilde{B}$, po to, by następnie można było określać zbiór rozmyty $\tilde{B}$ na podstawie zbioru rozmytego $\tilde{A}$ oraz relacji $\tilde{X}$ i $\tilde{Y}$, przez wyznaczanie jego dolnego $\tilde{B}_{min}$ i górnego $\tilde{B}_{max}$ ograniczenia

$$\tilde{B}_{min} \subset \tilde{B} \subset \tilde{B}_{max}, \tag{D.3.7}$$

gdzie:

$$\tilde{B}_{min} = \tilde{A} \circ \tilde{X} \quad \text{oraz} \quad \tilde{B}_{max} = \tilde{A} \circ \tilde{Y}. \tag{D.3.8}$$

Oceną "rozmycia" rozwiązania równania (D.3.2) może być względna "odległość" ograniczeń $\tilde{B}_{min}$ i $\tilde{B}_{max}$ zbioru $\tilde{B}$. Uwzględniając potrzeby występujące podczas rozwiązywania układów równań rozmytych wprowadzono ocenę rozmycia rozwiązania.

Definicja D.3.1: Ocena rozmycia Ψ rozwiązania równania (D.3.2) określana jest zgodnie z zależnością

$$\Psi \triangleq \frac{\tilde{\tilde{B}}_1 - \tilde{\tilde{B}}_2}{\tilde{\tilde{B}}_1}, \qquad (D.3.9)$$

gdzie:

$\tilde{\tilde{B}}_1$, $\tilde{\tilde{B}}_2$ - moce zbiorów rozmytych

$$\tilde{B}_1 = \tilde{B}_{max} \vee \tilde{B}_{min}; \quad \tilde{B}_2 = \tilde{B}_{max} \wedge \tilde{B}_{min} \qquad (D.3.10)$$

dla $\tilde{B}_{min}$, $\tilde{B}_{max}$ wg (D.3.8).

Dla otrzymania minimalnego "rozmycia" zbioru $\tilde{B}$ określonego przez jego ograniczenia $\tilde{B}_{min}$, $\tilde{B}_{max}$ (zgodnie z (D.3.7)) należy dążyć do minimalizacji "odległości" tych ograniczeń. Minimalna "odległość" tych ograniczeń wystąpi wtedy, gdy ograniczenia te będą równe.

Twierdzenie D.3.1: Warunkiem koniecznym równości zbiorów rozmytych $\tilde{B}_{min}$ i $\tilde{B}_{max}$ wyznaczanych wg (D.3.8) dla relacji rozmytych $\tilde{X}$ i $\tilde{Y}$ będących rozwiązaniem równania (D.3.2) określonym wg (D.3.5) jest to, aby zakres $\delta(\tilde{B})$ zbioru rozmytego $\tilde{B}$ był zawarty w zakresie $\delta(\tilde{A})$ zbioru rozmytego $\tilde{A}$

$$\delta(\tilde{B}) \subset \delta(\tilde{A}), \qquad (D.3.11)$$

gdzie:

$\delta(\tilde{A})$, $\delta(\tilde{B})$ wg (D.1.6).

Z twierdzenia D.3.1 wynika, że przed przystąpieniem do rozwiązywania równania (D.3.2) należy rozpatrzyć możliwość takiej zmiany zakresów zbiorów rozmytych $\tilde{A}$ i $\tilde{B}$, aby spełniony był warunek (D.3.11). Proponuje się ograniczanie zakresu zbioru rozmytego $\tilde{B}$.

Definicja D.3.2: Stopień $\varkappa$ ograniczenia zakresu $\delta(\tilde{B})$ dobrze określonego (por. def. D.1.11) zbioru rozmytego $\tilde{B} = \{(v, b(v)) : v \in V\}$ jest parametrem przekształcenia tego zbioru w podobny zbiór rozmyty $\tilde{C} \sim \tilde{B}$ taki, że:

$$\tilde{C} = \left\{(v, c(v)) : v \in V;\ c(v) = 0{,}5 + \frac{b(v) - 0.5}{\varkappa}\right\} \qquad (D.3.12)$$

gdzie:

$\varkappa \geq 1$.

Jeżeli znane są zakresy dobrze określonych (por. def. D.1.11) zbiorów $\tilde{A}$ i $\tilde{B}$:

$$\delta(\tilde{A}) = [\min(a) ; \max(a)], \quad \delta(\tilde{B}) = [\min(b) ; \max(b)],$$

to minimalny wymagany stopień ograniczenia zakresu zbioru rozmytego $\tilde{B}$ (dla spełnienia warunku (D.3.11)) wyznacza się

$$\varkappa = \max\left(\frac{\min(b) - 0,5}{\min(a) - 0,5}, \frac{\max(b) - 0,5}{\max(a) - 0,5}, 1\right). \qquad (D.3.13)$$

Przykład: Dla zbiorów rozmytych $\tilde{A}$ i $\tilde{B}$ (dobrze określonych) takich, że

$$a = [0,7; 0,4; 0,3],$$
$$b = [0,1; 0,2; 0,9],$$

wyznaczyć rozwiązanie równania (D.3.2):

$$\tilde{A} \circ \tilde{X} \subset \tilde{B} \subset \tilde{A} \triangle \tilde{Y}.$$

A. Rozwiązanie wyznaczane zgodnie z (D.3.6) bez uwzględniania wniosków wynikających z twierdzenia D.3.1

$$X = \begin{bmatrix} 0,7 \\ 0,4 \\ 0,3 \end{bmatrix} \alpha [0,1; 0,2; 0,9] = \begin{bmatrix} 0,1; & 0,2; & 1 \\ 0,1; & 0,2; & 1 \\ 0,1; & 0,2; & 1 \end{bmatrix},$$

$$Y = \begin{bmatrix} 0,7 \\ 0,4 \\ 0,3 \end{bmatrix} \varepsilon [0,1; 0,2; 0,9] = \begin{bmatrix} 0; & 0; & 0,9 \\ 0; & 0; & 0,9 \\ 0; & 0; & 0,9 \end{bmatrix}.$$

Wyznaczone relacje $\tilde{X}$ i $\tilde{Y}$ umożliwiają określenie zbioru $\tilde{B}$ jako przekształconego zbioru $\tilde{A}$. Zbiór $\tilde{B}$ wyznaczany jest przez jego ograniczenia $\tilde{B}_{min}$, $\tilde{B}_{max}$ (wg (D.3.7), (D.3.8))

$$b_{min} = [0,7; 0,4; 0,3] \circ \begin{bmatrix} 0,1; & 0,2; & 1 \\ 0,1; & 0,2; & 1 \\ 0,1; & 0,2; & 1 \end{bmatrix} = [0,1; 0,2; 0,7],$$

$$b_{max} = [0,7; 0,4; 0,3] \triangle \begin{bmatrix} 0; & 0; & 0,9 \\ 0; & 0; & 0,9 \\ 0; & 0; & 0,9 \end{bmatrix} = [0,3; 0,3; 0,9].$$

Ocena rozmycia rozwiązania (wg (D.3.9)):

$$\psi = \frac{(0,3 + 0,3 + 0,9) - (0,1 + 0,2 + 0,7)}{(0,3 + 0,3 + 0,9)} = 0,33.$$

B. Rozwiązanie wyznaczane po uwzględnieniu wniosków wynikających z twierdzenia D.3.1

Zakresy zbiorów rozmytych $\tilde{A}$ i $\tilde{B}$ (wg (D.1.6))

$$\delta(\tilde{A}) = [0,3 : 0,7]; \quad \delta(\tilde{B}) = [0,1 : 0,9]$$

nie spełniają warunku (D.3.11). Wymagany stopień ograniczenia zakresu zbioru rozmytego $\tilde{B}$ (wg (D.3.13))

$$\varkappa = \max\left(\frac{0,1 - 0,5}{0,3 - 0,5}, \frac{0,9 - 0,5}{0,7 - 0,5}, 1\right) = 2.$$

Zbiór rozmyty $\tilde{B}$ zastępowany jest zbiorem rozmytym $\tilde{C}$ określonym wg (D.3.12) dla $\varkappa = 2$:

$$c = [0,3;\ 0,35;\ 0,7].$$

którego zakres $\delta(\tilde{C}) = [0,3 : 0,7]$ spełnia warunek (D.3.11).
Relacja $\tilde{X}$, $\tilde{Y}$ wg (D.3.6):

$$X = \begin{bmatrix} 0,7 \\ 0,4 \\ 0,3 \end{bmatrix} \alpha_\varphi [0,3;\ 0,35;\ 0,7] = \begin{bmatrix} 0,3; & 0,35; & 1 \\ 0,3; & 0,35; & 1 \\ 1; & 1; & 1 \end{bmatrix},$$

$$Y = \begin{bmatrix} 0,7 \\ 0,4 \\ 0,3 \end{bmatrix} \varepsilon [0,3;\ 0,35;\ 0,7] = \begin{bmatrix} 0; & 0; & 0 \\ 0; & 0; & 0,7 \\ 0; & 0,35; & 0,7 \end{bmatrix}.$$

Ograniczenia zbioru $\tilde{C}$ (wg (D.3.8)):

$$c_{min} = [0,7;\ 0,4;\ 0,3] \circ \begin{bmatrix} 0,3; & 0,35; & 1 \\ 0,3; & 0,35; & 1 \\ 1; & 1; & 1 \end{bmatrix} = [0,3;\ 0,35;\ 0,7],$$

$$c_{max} = [0,7;\ 0,4;\ 0,3] \vartriangle \begin{bmatrix} 0; & 0; & 0 \\ 0; & 0; & 0,7 \\ 0; & 0,35; & 0,7 \end{bmatrix} = [0,3;\ 0,35;\ 0,7].$$

Ocena rozmycia rozwiązania $\Psi = 0$ ($c_{min} = c_{max}$).

D.4. Układy równań rozmytych

Rozpatruje się następujące zadanie:

- dany jest ciąg par zbiorów rozmytych $(\tilde{A}_1, \tilde{B}_1), (\tilde{A}_2, \tilde{B}_2), \ldots, (\tilde{A}_n, \tilde{B}_n)$ opisujących obiekt badań,
- należy wyznaczyć relacyjny model $(\tilde{X}, \tilde{Y})$ obiektu badań taki, że dla każdej danej pary zbiorów rozmytych $(\tilde{A}_j, \tilde{B}_j)$

$$\tilde{A}_j \circ \tilde{X} \subset \tilde{B}_j \subset \tilde{A}_j \triangle \tilde{Y}, \qquad (D.4.1)$$

- wyznaczony model $(\tilde{X}, \tilde{Y})$ ma umożliwić oszacowanie zbiorów rozmytych $\tilde{B}_j$ za pomocą ich ograniczeń, wyznaczonych wg (D.3.8) na podstawie zbiorów rozmytych $\tilde{A}_j$, bez konieczności bezpośredniego określania zbiorów $\tilde{B}_j$.

W celu rozwiązania powyższego zadania wyznacza się rozwiązanie $(\tilde{X}_j, \tilde{Y}_j)$ równania relacyjnego (D.4.1) dla każdej pary $(\tilde{A}_j, \tilde{B}_j)$, otrzymując ciąg rozwiązań

$$\tilde{X}_j = \tilde{A}_j^T \alpha \tilde{B}_j, \qquad \tilde{Y}_j = \tilde{A}_j^T \varepsilon \tilde{B}_j. \qquad (D.4.2)$$

Zaleca się, aby pary zbiorów rozmytych będące elementami rozpatrywanego ciągu $\{(\tilde{A}_j, \tilde{B}_j)\}$ spełniały warunek (D.3.11). Wymaga to wyznaczenia wg (D.3.13) stopni $\mathcal{H}_j$ ograniczenia zakresów $\delta(B_j)$ i przyjęcia na ich podstawie wspólnego dla wszystkich n równań (D.4.1) stopnia $\mathcal{H}$ ograniczenia zakresu, spełniającego warunek

$$\bigwedge_{j \in [1:n]} \mathcal{H} \geq \alpha \mathcal{H}_j \qquad \text{dla } \alpha \geq 1. \qquad (D.4.3)$$

Po wyznaczeniu liczby $\mathcal{H}$ ogranicza się zakresy wszystkich zbiorów rozmytych $\tilde{B}_j$ wg (D.3.12).

Rozwiązaniem ogólnym, tzn. rozwiązaniem spełniającym wszystkie równania (D.4.1) rozpatrywanego n-elementowego ciągu i optymalnym w świetle warunków (D.3.3) i (D.3.4) jest para relacji rozmytych $({}_{(n)}\tilde{X}, {}_{(n)}\tilde{Y})$:

$${}_{(n)}\tilde{X} = \bigwedge_{j \in [1:n]} \tilde{X}_j; \quad {}_{(n)}\tilde{Y} = \bigvee_{j \in [1:n]} \tilde{Y}_j. \qquad (D.4.4)$$

Rozwiązanie ogólne układu (ciągu) równań relacyjnych można wyznaczać metodą rekurencyjną, co nie wymaga jednoczesnego zapamiętywania (np.w pamięci EMC) wartości funkcji charakterystycznych wszystkich par zbiorów rozmytych $(\tilde{A}_j, \tilde{B}_j)$. Możliwość ta jest szczególnie korzystna w analizowaniu dużych zbiorów danych. Oznaczając $({}_{(n)}\tilde{X}, {}_{(n)}\tilde{Y})$ rozwiązanie uwzględniające ciąg danych $\{(\tilde{A}_j, \tilde{B}_j) : j \in [1:n]\}$ otrzymuje się

$$
\begin{aligned}
{}_{(n)}\tilde{X} &= {}_{(n-1)}\tilde{X} \wedge \tilde{X}_n = {}_{(n-1)}\tilde{X} \wedge (\tilde{A}_n^T \subset \tilde{B}_n), \\
{}_{(n)}\tilde{Y} &= {}_{(n-1)}\tilde{Y} \vee \tilde{Y}_n = {}_{(n-1)}\tilde{Y} \vee (\tilde{A}_n^T \in \tilde{B}_n).
\end{aligned}
\tag{D.4.5}
$$

D.4.1. Ograniczone rozwiązanie układu równań rozmytych

Wyznaczone rozwiązanie $({}_{(n)}\tilde{X},\ {}_{(n)}\tilde{Y})$ jest rozmytym modelem danych określonych ciągiem $\{(\tilde{A}_j,\ \tilde{B}_j)\}$. Oceną rozmycia rozwiązania układu n równań jest liczba

$$
{}_{(n)}\psi = \max_{j \in [1:n]} {}_{(n)}\psi_j, \tag{D.4.6}
$$

gdzie:

${}_{(n)}\psi_j$ - jest określone dla zbioru $\tilde{A}_j$

zgodnie z (D.3.9) na podstawie ogólnego rozwiązania $({}_{(n)}\tilde{X},\ {}_{(n)}\tilde{Y})$ układu równań. Dla zmniejszenia rozmycia rozwiązania poszukuje się pary $(\tilde{A}_k, \tilde{B}_k)$, która jest przyczyną dużej wartości liczby ${}_{(n)}\psi$. Po wskazaniu pary $(\tilde{A}_k, \tilde{B}_k)$ może być podjęta decyzja o jej pominięciu podczas wyznaczania rozwiązania (co wpłynęłoby na zmniejszenie rozmycia rozwiązania). Tak otrzymane rozwiązanie nazywane będzie "ograniczonym rozwiązaniem" układu równań rozmytych $({}_{(n\backslash k)}\tilde{X},\ {}_{(n\backslash k)}\tilde{Y})$. Dla wskazania pary $(\tilde{A}_k,\ \tilde{B}_k)$ wyznacza się ciąg rozwiązań $\{({}_{(n\backslash 1)}\tilde{X},\ {}_{(n\backslash 1)}\tilde{Y}) : 1 \in [1:n]\}$, którego elementami są ogólne rozwiązania układów równań uwzględniających ciągi danych $\{(\tilde{A}_j, \tilde{B}_j) : j \in [1:n];\ j \neq 1\}$. Dla wyznaczonego ciągu rozwiązań wyznacza się następnie ciąg ich ocen rozmycia

$$
{}_{(n\backslash 1)}\psi = \max_{j \in [1:n]} {}_{(n\backslash 1)}\psi_j \tag{D.4.7}
$$

Poszukiwaną parą $(\tilde{A}_k,\ \tilde{B}_k)$ jest para, dla której

$$
\bigwedge_{1 \in [1:n]} \left[{}_{(n\backslash k)}\psi \leq {}_{(n\backslash 1)}\psi \right]. \tag{D.4.8}
$$

Praktyczna realizacja opisanego sposobu wyznaczania ograniczonego rozwiązania układu równań wymaga dużej liczby działań i prowadzić może do pominięcia nadmiernej liczby równań.

Jako drugi, wygodniejszy sposób, proponuje się wyznaczanie ograniczonego rozwiązania układu równań rozmytych wg algorytmu:

- dla każdego j-tego równania (D.4.1), gdzie $j \in [1:n]$ wyznacza się rozwiązanie $(\tilde{X}_j,\ \tilde{Y}_j)$ wg (D.4.2),

- zestawia się n-elementowe ciągi $\{x_j[k,l]\}$, $\{y_j[k,l]\}$ elementów macierzy X i Y,
- elementy ciągów porządkuje się dla otrzymania niemalejących ciągów $\{x_j\}_{k,l}$ i nierosnących ciągów $\{y_j\}_{k,l}$,
- dla każdego ciągu $\{x_j\}_{k,l}$ określa się liczbę $r_{k,l}$ równą numerowi porządkowemu ostatniego elementu < 1 $(0 \leq r_{k,l} \leq n)$,
- dla każdego ciągu $\{y_j\}_{k,l}$ wyznacza się liczbę $r_{k,l}$ równą numerowi porządkowemu ostatniego elementu > 0 $(0 \leq r_{k,l} \leq n)$,
- przyjmuje się liczbę β $(0 < \beta < 1)$,
- macierze X i Y będące poszukiwanymi ograniczonymi rozwiązaniami układu równań zestawia się z elementów (elementów odpowiednich uporządkowanych ciągów) o numerach porządkowych

$$n_{k,l} = \min(\text{int}(\beta r_{k,l} + 1), n), \qquad \text{(D.4.9)}$$

gdzie:

int( ) - część całkowita argumentu.

Rozwiązanie wyznaczone dla $\beta = 0$ jest pełnym rozwiązaniem układu równań rozmytych. Wartość liczby β powinna być ustalona w wyniku optymalizacji.

D.5. Zbiór reprezentantów

Wprowadzone wcześniej (D.1.9) pojęcie α-obcięcia zbioru rozmytego zostanie rozszerzone.

Zakłada się, że w przestrzeni V są elementy posiadające pewną własność. Elementy te tworzą zbiór rozmyty $\tilde{A}$ o znanej funkcji charakterystycznej. Istnieje potrzeba określenia w przestrzeni V trzech zbiorów elementów (zbiorów w sensie klasycznym):

- zbioru V_1 elementów, które mogą być uznane jako elementy posiadające tę własność,
- zbioru V_2 elementów, które mogą być uznane jako elementy nie posiadające tej własności,
- zbioru V_3 elementów, co do których brak podstaw do stwierdzenia czy posiadają one tę własność, czy też nie.

Formalne określenie takich zbiorów jest możliwe po przyjęciu wartości α i β (gdzie $\alpha \geq \beta$) określających dwa poziomy obcięcia zbioru rozmytego $\tilde{A}$:

$$V_1 = A_{\geq \alpha} \tag{D.5.1}$$

$$V_2 = A_{\leq \beta} = V \setminus A_{>\beta} \tag{D.5.2}$$

$$V_3 = A_{>\beta, <\alpha} = V \setminus (A_{\geq \alpha} \cup A_{\leq \beta}). \tag{D.5.3}$$

Trójka zbiorów (V_1, V_2, V_3) określa podział przestrzeni V wyznaczony przez zbiór rozmyty $\tilde{A}$. Liczby α i β można interpretować podobnie jak interpretuje się pojęcia "poziom istotności testu" i "moc testu" stosowane w zakresie weryfikacji hipotez statystycznych.

Szczególnym przypadkiem obcięcia skończonego zbioru rozmytego $\tilde{A}$ jest obcięcie $A_{\geq \alpha}$ dla takiej największej wartości poziomu α, która wyznacza jeszcze niepusty zbiór $A_{\geq \alpha}$. Tak wyznaczone obcięcie zbioru rozmytego, oznaczane $q(\tilde{A})$ nazywane jest zbiorem reprezentantów zbioru $\tilde{A}$.

Definicja D.5.1: Zbiór reprezentantów skończonego zbioru rozmytego $\tilde{A} = \{(v, a(v)) : v \in V\}$ jest zbiorem $q(\tilde{A})$ takim, że

$$q(\tilde{A}) \stackrel{d}{=} A_{\geq \alpha} \quad \text{dla} \quad \alpha = \max_{v \in V} [a(v)]. \tag{D.5.4}$$

Wyznaczanie zbioru reprezentantów napotyka na pewne trudności wtedy, gdy zbiór rozmyty $\tilde{A}$ określony jest na podstawie rozwiązania $(\tilde{X}, \tilde{Y})$ relacyjnego równania rozmytego, wyłącznie za pomocą jego dolnego i górnego ograniczenia

$$\tilde{A}_{min} = \{(v, a_{min}(v)) : v \in V\}, \quad \tilde{A}_{max} = \{(v, a_{max}(v)) : v \in V\}.$$

Zbiór reprezentantów określa się w takim przypadku jako zbiór

$$q(\tilde{A}) \stackrel{d}{=} \left\{ v \in V : \forall_{x \in V} \; [a_{min}(x) \leq a_{max}(v)] \right\}. \tag{D.5.5}$$

Przekształcając zależność (D.5.5), można definiować zbiór reprezentantów jako obcięcie górnego ograniczenia zbioru rozmytego, na poziomie wyznaczonym przez dolne ograniczenie tego zbioru

$$q(\tilde{A}) = A_{max, \geq \alpha} \quad \text{gdzie} \quad \alpha = \max_{v \in V} [a_{min}(v)]. \tag{D.5.6}$$

Zbiory reprezentantów zbiorów rozmytych wyznaczane są między innymi dla umożliwienia określania obrazów elementów przestrzeni cech (por. rozdz. 2.6). W celu uproszczenia sposobu wyznaczania obrazów proponuje się wyznaczanie w przestrzeni wartości cech, współrzędnych reprezentanta $q(\tilde{A}_k)$ zbioru rozmytego $\tilde{A}_k = \{(v, a_k(v)) : v \in V\}$, zgodnie z zależnością

$$q(\tilde{A}_k) = \frac{\sum_{v \in V} (a_k(v)\ v)}{\sum_{v \in V} a_k(v)}. \qquad \text{(D.5.7)}$$

Przed stosowaniem zależności (D.5.7) do wyznaczania reprezentantów rodziny zbiorów rozmytych $\{\tilde{A}_k\}$, dla potrzeb związanych z wyznaczaniem obrazów, zaleca się przekształcenie funkcji charakterystycznych zbiorów $\tilde{A}_k$ będących elementami tej rodziny

$$a_k(v) \mapsto b_k(v) = \begin{cases} a_k(v) & \text{jeżeli} \quad \underset{j \neq k}{\forall} \ [a_j(v) \leqslant a_k(v)], \\ 0 & \text{jeżeli} \quad \underset{j \neq k}{\exists} \ [a_j(v) > a_k(v)]. \end{cases} \qquad \text{(D.5.8)}$$

LITERATURA

[1] Ackoff R.L.: Decyzje optymalne w badaniach stosowanych. PWN, Warszawa 1969.

[2] Adamczyk J., Jaźwiński J., Sypnik R.: System techniczny jako obiekt badań diagnostycznych. W: [30] s. 29-41.

[3] Adamczyk J.: Wybrane zagadnienia wibroakustycznej diagnostyki maszyn. Zeszyty Naukowe AGH nr 679. Kraków 1979.

[4] Advances in Fuzzy Set Theory and Applications. Red.: Gupta M.M. North-Holland, Amsterdam 1979.

[5] Ajzerman M.A., Brawerman E.M., Rozonoer L.I.: Rozpoznawanie obrazów. Metoda funkcji potencjalnych. WNT, Warszawa 1976.

[6] Anderson J., Forsythe J.M.: Information processing of medical records. North Holland, Amsterdam 1970.

[7] Ashby W.R.: Wstęp do cybernetyki. PWN, Warszawa 1963.

[8] Ball G.H., Hall D.J.: A clustering technique for summarizing multivariate data. "Behavioral Science" R. 12 1967, nr 3, s.153-155.

[9] Batko W.: Wybrane testy statystyczne i ich zastosowanie w diagnostyce. W: [66] s. 269-292.

[10] Bendat J.S., Piersol A.G.: Analiza i pomiary sygnałów losowych. PWN, Warszawa 1975.

[11] Bezdek J.: Prototype Classification and Feature Selection with Fuzzy Sets. "IEEE Trans. on Syst., Man and Cyb." 1977 R. 7, nr 2, s.87-92.

[12] Bezdek J.C.: Numerical Taxonomy with Fuzzy Sets. "J. Math. Biol.", 1974, nr 1, s. 57-71.

[13] Birger I.A.: Techničeskaja diagnostika. Mašinostroenie,Moskva 1978.

[14] Birger I.A.: Osnovy teorii techničeskoj diagnostiki. W: Birger I.A. i in.: "Rasčet na pročnost detalej mašin". Moskva: Mašinostroenie 1979, s. 656-670.

[15] Bock H.H.: Automatische Klassifikation. Theoretische und praktische Methoden zur Gruppierung und Strukturierung von Daten. Vandenhoeck und Ruprecht, Gottingen 1974.

[16] Bogomolov A.M., Tverdochlebov V.A.: Diagnostika složnych sistem.Naukova Dumka, Kiev 1974.

[17] Borisov A.N., Vul'f G.N., Osis Ja.Ja.: Prognozirovanie sostojanija složnych sistem s ispol'zovaniem teorii razmytych množestv. W: [57] s. 79-84.

[18] Borko H.: Maszyny cyfrowe w badaniach naukowych. WNT,Warszawa 1969.

[19] Borodkin L.I.: Algoritm approksimacii empiričeskich grafov. "Avtomatika i Telemechanika", 1978, nr 2, s. 158-168.

[20] Cempel C.: Podstawy wibroakustycznej diagnostyki maszyn. WNT, Warszawa 1982.

[21] Chen C.H.: Statistical pattern recognition. Spartan, Washington 1973.

[22] Cholewa W.: Różnicowe widmo reprezentatywne umownego zastępczego źródła sygnału w badaniach maszyn. Archiwum Akustyki 1976, r. 11, nr 3, s. 275-289.

[23] Cholewa W., Jaskóła Z.: Doświadczalne badania konstrukcyjne. Pojęcia podstawowe. Ergonomiczne problemy hałasu i drgań. Politechnika Śląska, Gliwice 1980. Skr. Uczel. nr 942.

[24] Cholewa W., Moczulski W.: Programowany analizator sygnałów PAS4. Instrukcja obsługi - wersja 2. 1982 maszyn. Raport IMiPKM, Politechnika Śląska, Gliwice.

[25] Cholewa W., Moczulski W., Kurowicz M.: Cyfrowa rejestracja ocen sygnałów w badaniach maszyn metodami akustycznymi. Część I - Układ aparatury, Część II - Opis programów. 1978 maszyn. Raport IPKM,Politechnika Śląska, Gliwice.

[26] Cholewa W., Moczulski W., Słomski M., Solipiwko A.: System programów automatycznego wnioskowania w badaniach diagnostycznych SPAW.Instrukcja obsługi. 1980 maszyn. Raport IPKM. Politechnika Śląska, Gliwice.

[27] Cypkin Z.J.: Podstawy teorii układów uczących się. WNT,Warszawa 1973.

[28] Czogała E., Pedrycz W.: Elementy i metody teorii zbiorów rozmytych. Politechnika Śląska, Gliwice 1980. Skr. uczel. nr 989.

[29] Dąbrowski M., Laus-Mączyńska K.: Metody wyszukiwania i klasyfikacji informacji. WNT, Warszawa 1978.

[30] Diagnostyka urządzeń mechanicznych. Praca zbiorowa. Ossolineum, Wrocław 1977.

[31] Diagnostyka łożysk tocznych. Praca zbiorowa. Ossolineum,Wrocław 1980.

[32] Diagnostyka pojazdów. Praca zbiorowa. Ossolineum, Wrocław 1981.

[33] Dietrych J.: System i konstrukcja. WNT, Warszawa 1978.

[34] Draper N.R., Smith H.: Analiza regresji stosowana.PWN,Warszawa 1973.

[35] Duda R.O., Hart P.E.: Pattern classification and scene analysis. J.Wiley and Sons, New York 1973.

[36] Fomin V.N.: Matematičeskaja teorija obyčaemych opoznajuščich sistem. Leningrad: Izd. Leningradskogo universiteta 1976.

[37] Fordon W.A., Bezdek J.C.: The Application of Fuzzy Set Theory to Medical Diagnosis. W: [4], s. 445-461.

[38] Franks L.E.: Teoria sygnałów. PWN, Warszawa 1975.

[39] Friedman H.P., Rubin J.: On some invariant criteria for grouping data. "J. Am. Stat. Assos." 1967, R. 67, s. 1159-1178.

[40] Fritsche M.: Automatic Clustering Techniques in Information Retrieval. Diplomarbeit.Institut für Informatik der Universität Stuttgart 1973.

[41] Fukunaga K., Ando S.: The Optimum Nonlinear Features for a Scatter Criterion in Discriminant Analysis. "IEEE Trans. on Inf.Theory" 1977 R. 23, nr 4, s. 453-459.

[42] Fukunaga K., Hostetler L.D.: The Estimation of the Gradient of a Density Function, with Applications in Pattern Recognition."IEEE Trans. on Inf. Theory", 1975, R. 21, nr 1, s. 32-40.

[43] Genkin M.D. i in.: Voprosy akustičeskoj diagnostiki. "Metody vibroizoljacii mašin i prisoedinennych konstrukcij". Nauka, Moskva 1975, s. 67-90.

[44] Genkin M.D. i in.: Vibroakustičeskaja diagnostika mašin i mechanizmov. "Vibracii v technike. Spravočnik". Mašinostroenie, Moskva 1981, T. 5, s. 380-419.

[45] Giemza H.: Opis stanowiska do badań przekładni zębatych w układzie mocy zamkniętej. Zesz. Nauk. Pol. Śl. nr 280, Energetyka z. 35, Gliwice 1970, s. 67-77.

[46] Ginsberg A.S., Offensend F.L.: An application of decision theory to a medical diagnosis. "IEEE Trans. System Sci.", 1968, R. 4, nr 3, s. 355-362.

[47] Goguen J.A.: L-fuzzy sets. "J. Math. Anal. Appl.", 1967, R.18,s.145-174.

[48] Goguen J.A. (jr): Concept Representation in Natural and Artificial Languages: Axioms, Extensions and Applications for Fuzzy Sets. "Int. J. Man-Machine Studies" 1974, R. 6, s. 513-561.

[49] Gottwald S.: Set theory for fuzzy sets of higher level. "Fuzzy Sets and Systems", 1979, R. 2, s. 125-151.

[50] Grundspen'kis Ja.A., Markovič Z.P., Osis Ja.Ja.: Postroenie topologičeskoj modeli ob"ekta. W: [57], s. 19-35.

[51] Harmuth H.F.: Transmission of Information by Orthogonal Functions. Springer, Berlin 1969.

[52] Hotelling H.: Relations between two sets of variables. "Biometrika", 1936, R. 28, s. 321-377.

[53] Hunt E.B.: Artificial Intelligence. Academic Press, New York 1975.

[54] IRD Mechanalysis: Machinery Troubleshooting Vibration and Noise Identification. Application Report No 106, Columbus, Ohio 1974.

[55] Kandel A., Byatt W.J.: Fuzzy Sets, Fuzzy Algebra and Fuzzy Statistics. "Proc. of the IEEE" 1978, R. 66, nr 12, s. 1619-1639.

[56] Kibernetika i diagnostika. Praca zbiorowa. R.2. Zinatne, Riga 1968.

[57] Kibernetika i diagnostika. Praca zbiorowa. R.5. Zinatne, Riga 1972.

[58] Krzyworzeka P.: Uwagi o charakterystykach odwzorowania istotnych własności obiektu diagnostyki. W: [66], s. 293-323.

[59] Kurowicz M.: Metoda badania wpływu cech geometrycznych kół zębatych na rozkład obciążenia wzdłuż odcinka przyporu. 1979 maszyn. Rozprawa doktorska (6.07.1979). Politechnika Śląska, Gliwice.

[60] Kurowski W.: Systemy diagnostyczne urządzeń mechanicznych. W: [79] s. 5-85.

[61] Kurowski W., Ziemba S.: Wybrane zagadnienia z cybernetyki technicznej. W: [30], s. 43-63.

[62] Linnik J.W.: Metoda najmniejszych kwadratów i teoria opracowywania obserwacji. PWN, Warszawa 1962.

[63] Mahalanobis P.C.: On the generalized distance in statistics. Proceedings of the National Institute of Science of India, 1936, nr 2, s. 49-55.

[64] Markovič Z.P.: Ispol'zovanie graf-modeli dlja rešenija zadač techničeskoj diagnostiki. W: [56], s. 49-62.

[65] Mesarovič M.D., Takahara Y.: General Systems Theory-Mathematicall Foundation. Acad. Press, New York 1975.

[66] Metody cyfrowej analizy sygnałów wibroakustycznych. Praca zbiorowa. Ossilineum, Wrocław 1979.

[67] Mitchell J.S.: Machinery Health Monitoring. ENDEVCO Tech.Paper TP26%. San Juan Capistrano: Endevco 1974.

[68] Mozgalevskij A.V., Gaskarov D.V.: Techničeskaja diagnostika. Vyšsaja Škola, Moskva 1975.

[69] Müller L.: Nowy model dynamiczny przekładni walcowej o zębach prostych. "Przegląd Mechaniczny" 1974, nr 21, s. 725-728.

[70] Osis Ja.Ja.: Raspoznavanie neispravnostej složnych ob"ektov diagnostiki s ispol'zovaniem teorii rozmytych množestv. W: [56], s.13-17.

[71] Otnes R.K., Enochson L.: Analiza numeryczna szeregów czasowych. WNT, Warszawa 1978.

[72] Papoulis A.: Prawdopodobieństwo, zmienne losowe i procesy stochastyczne. WNT, Warszawa 1972.

[73] Pau L.F.: Diagnosis of Equipment Failures by Pattern Recognition. "IEEE Trans. on Reliability", 1974, R. 23, nr 3, s. 202-208.

[74] Pau L.F.: Diagnostic des pannes dans les systemes. Toulouse: CEPADUS. 1975.

[75] Pawłow B.W.: Badania diagnostyczne w technice. WNT, Warszawa 1967.

[76] Pavlov B.V.: Akustičeskaja diagnostika mechanizmov. Mašinostroenie, Moskva 1971.

[77] Percepton - sistema raspoznavanija obrazov. Red.: Ivachnenko A.G. Naukova Dumka, Kiev 1975.

[78] Ploch M.: Metoda badania wpływu nierównomiernego dolegania zębów na nośność przekładni zębatej. 1978 maszyn. Rozprawa doktorska (8.12. 1978) Politechnika Śląska, Gliwice.

[79] Podstawy diagnostyki urządzeń mechanicznych. Praca zbiorowa. Ossolineum, Wrocław 1978.

[80] Popkov V.I.: Vibroakustičeskaja diagnostika i sniženie vibroaktivnosti sudovych mechanizmov. Sudostroenie, Leningrad 1974.

[81] Quantitative Sociology. International Perspectives on Mathematical and Statistical Modeling. Red.: Blalock H.M. Academic Press, New York 1975.

[82] Reznikov A.P.: Obrabotka nakoplennoj informacii v zatrudnennych uslovijach. Nauka, Moskva 1976.

[83] Salenieks Ja.K.: Process diagnostiki, složnych ob"ektov. "Diagnostika i identifikacija". Zinatne, Riga 1974.

[84] Sammon J.: Interactive pattern analysis and classification. "IEEE Trans. on Comp." 1970, R.7, s. 594-616.

[85] Sanchez E.: Resolution of Composite Fuzzy Relation Equations. "Information and Control", 1976, R.30, s. 38-48.

[86] Sanchez E.: Inverses of Fuzzy Relations. Application to Possibility Distributions and Medical Diagnosis. "Fuzzy Sets and Systems", 1979 nr 2, s. 75-86.

[87] Sanchez E.: Medical Diagnosis and Composite Fuzzy Relations. W:[4] s. 437-444.

[88] Sebestyen G.: Decision - Making Processes in Pattern Recognition, "ACM Monograph". New York 1962.

[89] Silvey S.D.: Wnioskowanie statystyczne. PWN, Warszawa 1978.

[90] Słomski M.: Akustyczna metoda badań konstrukcyjnych elektrycznych pieców łukowych. 1982 maszyn. Rozprawa doktorska (24.09.1982). Politechnika Śląska, Gliwice.

[91] Sobczak W., Malina W.: Metody selekcji informacji. WNT, Warszawa 1978.

[92] Sobkowski J.: Częstotliwościowa analiza sygnałów. Wyd. MON, Warszawa 1975.

[93] Stanicki T., Szlachcic E.: Wybrane zagadnienia teorii systemów. W: [30], s. 66-102.

[94] Steinhagen H.E., Fuchs S.: Objekterkennung. Einführung in die mathematische Methoden der Zeichenerkennung. VEB Verlag Technik, Berlin 1976.

[95] Szabatin J.: Podstawy teorii sygnałów. WKiŁ, Warszawa 1982.

[96] Tadeusiewicz R.: Przykłady wykorzystania metod cybernetycznych. W: [66], s. 225-268.

[97] Thomas D.W.: Vehicle Sounds and Recognition. "Pattern Recognition. Ideas in Practice" Red.: Batchelor B.G. Plenum Press, New York 1978.

[98] Tou J.T., Gonzalez R.C.: Pattern Recognition Principles. Addison-Wesley, London 1974.

[99] Watanabe S.: Modified concepts of logic, probability and information based on generalized continuous characteristic function. "Information and Control", 1969, R. 15, nr 1.

[100] Webster R.: Quantitative and numerical methods in soil classification and survey. Oxford: Clavendon Press 1977.

[101] Wojnarowski J.: Zastosowanie grafów w analizie drgań układów mechanicznych. PWN, Warszawa 1981.

[102] Young T.Y., Calvert T.W.: Classification, estimation and pattern recognition. Elsevier Publ. Comp. New York 1974.

[103] Zadeh L.A.: Fuzzy Sets. "Information and Control". 1965, R.8, s.338-353.

[104] Ziemba S.: Rola diagnostyki technicznej. W: [30] s. 5-8.

[105] Žitkov G.N.: O klassifikacii strukturnych elementov pri analize složnych izobraženij. "Naučno Techničeskaja Informacija" 1970, R.2, s. 14-18.

METODA DIAGNOZOWANIA MASZYN Z ZASTOSOWANIEM ZBIORÓW ROZMYTYCH

S t r e s z c z e n i e

Praca opisuje sposób postępowania podczas projektowania (wyznaczania działania) układu diagnozującego przeznaczonego do wyznaczania klas stanów badanych maszyn na podstawie obserwowanych sygnałów diagnostycznych. Omówiono zagadnienia związane z obserwowaniem maszyny będącej obiektem badań, wprowadzając pojęcia cech sygnałów, cech oddziaływań i cech stanu. Zaproponowano metodę minimalizacji liczby uwzględnianych cech, polegającą między innymi na przekształcaniu przestrzeni cech w przestrzeń obrazów cech, w której współrzędne punktu są wartościami funkcji przynależności tego punktu do wyróżnionych rozmytych klas punktów podobnych. Szczegółowo omówiono problemy związane z wyznaczaniem obrazów. Omówiono sposób modelowania maszyny oraz sposób modelowania klasy maszyn wykonanych wg wspólnej dla nich konstrukcji. Opisano sposób projektowania układu diagnozującego, polegający na określeniu jego działania, wyznaczanego na podstawie rozmytych relacji między obrazami cech sygnałów diagnostycznych i obrazami cech stanu maszyn. Przyjęto, że dla wyznaczenia działania układu diagnozującego możliwe jest przeprowadzenie "badań uczących", podczas których znany jest stan badanych obiektów. Omówiono pojęcia dotyczące zbiorów i relacji rozmytych oraz problemy formułowania i rozwiązywania układów równań rozmytych.

Słowa kluczowe:
badania diagnostyczne, sygnał, obraz sygnału, stan maszyny, klasyfikacja, grupowanie, rozpoznawanie obrazów, modelowanie, model systemowy, relacja rozmyta, układ równań rozmytych, projektowanie układów diagnozujących.

Dziedziny:
podstawy konstrukcji maszyn, diagnostyka wibroakustyczna maszyn, eksploatacja maszyn.

МЕТОД ДИАГНОЗИРОВАНИЯ МАШИН С ПРИМЕНЕНИЕМ РАЗМЫТЫХ МНОЖЕСТВ

Р е з ю м е

В работе представлен метод действования во время проектирования диагностической системы, которая предназначена для определения классов состояний испытываемых машин на основе наблюдаемых диагностических сигналов.Обсуждены проблемы связанные с наблюданием машины (объекта исследования),введены понятия признаков сигналов, признаков влияний и признаков состояния.Предположен метод минимализации числа признаков, принимаемых во внимание. Этот метод связан между прочем с преобразованиями пространства признаков в пространство образов признаков. В этом пространстве координаты элементов являются величинами функций принадлежности этих элементов к отмеченным размытым классам похожих элементов. Подробно обсуждены проблемы связанные с определением образов. Обсуждён метод моделирования машины и метод моделирования класса машин произведённых на основе общей конструкции. Разработанный метод проектирования диагностической системы, касается определения сущности действия этой системы на основе размытых реляций между образами признаков диагностических сигналов и образами признаков состояния машин. Принято,что для определения такого действия диагностической системы можно реализовать "обучаемые испытания", во время которых известно состояние испытываемых объектов. Обсуждены понятия размытых множеств и реляций вместе с задачами решения систем размытых уравнений.

THE METHOD OF DIAGNOSTIC OF MACHINES WITH THE USE OF FUZZY SETS

S u m m a r y

The paper deals with a method of projecting the diagnostic system designed for state classes recognition on the basis of observed diagnostic signals of machines. The problem connected with tested machine observation has been described. The terms of signals features as well as interactions and state features have been introduced. The method of minimizing of number of features that have to be taken into account has been proposed. Generally the method is based on the special transformation of feature space into feature image space. Co-ordinates of a point in this image space are the same as the values of compatibility functions expressing the memberships of that point in the fuzzy classes of similar points. The problems dealing with calculating of images have been discussed in detail. In the paper the method of the single machine modelling as well as the whole machines class modelling have been developed. The method of machinery diagnostic system projecting has been suggested. That is based on the fuzzy relations between images of diagnostic signals features and machines state features. The proposed method needs "selftraining data" obtained from learning tests of machines. The problems dealing with fuzzy sets, fuzzy relations and the methods of double fuzzy equations sets solution have been described too.

BIBLIOTEKA GŁÓWNA
Politechniki Śląskiej
3361/83/79

WYDAWNICTWA NAUKOWE I DYDAKTYCZNE POLITECHNIKI ŚLĄSKIEJ MOŻNA NABYĆ W NASTĘPUJĄCYCH PLACÓWKACH:

44-100 Gliwice — Księgarnia nr 096, ul. Konstytucji 14 b
44-100 Gliwice — Spółdzielnia Studencka, ul. Wrocławska 4 a
40-950 Katowice — Księgarnia nr 015, ul. Żwirki i Wigury 33
40-096 Katowice — Księgarnia nr 005, ul. 3 Maja 12
41-900 Bytom — Księgarnia nr 048, Pl. Kościuszki 10
41-500 Chorzów — Księgarnia nr 063, ul. Wolności 22
41-300 Dąbrowa Górnicza — Księgarnia nr 081, ul. ZBoWiD-u 2
47-400 Raciborz — Księgarnia nr 148, ul. Odrzańska 1
44-200 Rybnik — Księgarnia nr 162, Rynek 1
41-200 Sosnowiec — Księgarnia nr 181, ul. Zwycięstwa 7
41-800 Zabrze — Księgarnia nr 230, ul. Wolności 288
00-901 Warszawa — Ośrodek Rozpowszechniania Wydawnictw Naukowych PAN — Pałac Kultury i Nauki

Wszystkie wydawnictwa naukowe i dydaktyczne zamawiać można poprzez Składnicę Księgarską w Warszawie, ul. Mazowiecka 9.