

Ekstrakcja cech obrazów płomienia współpalania węgla i biomasy z wykorzystaniem wizyjnego systemu diagnostycznego

Streszczenie. W pracy przedstawiono porównanie wybranych metod ekstrakcji cech obrazów dla współpalania pyłu węglowego i biomasy. Zbiór cech dyskryminacyjnych poddano transformacji w celu redukcji wymiarowości oraz grupowania przypadków stabilnego i niestabilnego spalania. Porównano metody: analizę składowych głównych, skalowanie wielowymiarowe oraz analizę czynnikową. Dla każdej z zastosowanych metod redukcji wymiarowości stwierdzono, że analizowane zbiory danych obrazu płomienia trudno jednoznacznie pogrupować na przypadki należące do klasy stabilny i niestabilny. Ponadto za pomocą analizy PCA określono istotność poszczególnych cech obrazu płomienia w reprezentowanym modelu.

Abstract. This paper presents comparison image feature extraction methods of co-firing pulverized coal and biomass. In order to dimension reduction and classify state of process (stable and unstable) discriminant features were transformed. Compared method are: principal components analysis, multidimensional scaling and multivariate analysis. Experimental results show that presented methods for flame image analysis is difficult to be grouped into cases belonging to the class of stable and unstable. Furthermore, significance of the individual characteristics of the flame image in the represented model were obtained by PCA analysis. (**Flame image feature extraction of co-firing coal and biomass with vision diagnostic system**).

Słowa kluczowe: ekstrakcja cech, płomień, współpalanie, analiza danych.

Keywords: feature extraction, flame, co-firing, data analysis.

Wstęp

Wykorzystanie płomienia jako źródła informacji o procesie spalania jest jednym ze sposobów diagnozowania tego procesu. Analiza płomienia może być wykonana za pomocą obrazowania [1] jak też również za pomocą spektrometru [2]. Wykorzystując obraz płomienia można uzyskać informację o stanie procesu praktycznie bez żadnych opóźnień. Jest to szczególnie istotne w przypadku spalania paliw charakteryzujących się dużą zmiennością właściwości fizyko-chemicznych. Do tej grupy należą m.in. mieszaniny biomasy i węgla, których spalanie jest najbardziej rozpowszechnionym w Polsce sposobem wykorzystania paliw odnawialnych [1].

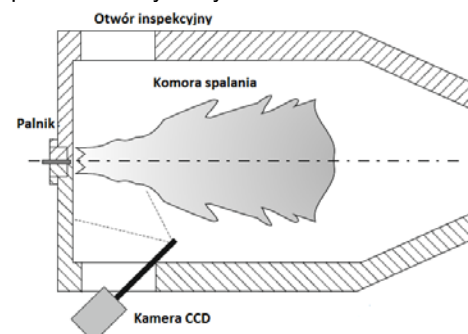
Płomień towarzyszy egzotermicznym reakcjom utleniania. Obecność płomienia związana jest więc z miejscem w przestrzeni, gdzie taki rodzaj reakcji występuje. Trudno jest jednak w takim przypadku wskazać granicę oddzielającą przestrzeń, w której spalanie występuje, ponieważ stężenia reagentów nie zmieniają się w sposób skokowy. Głównym źródłem promieniowania w płomieniu pyłowym jest obłok rozgrzanych do wysokiej temperatury cząstek stałych (pyłu węglowego, sadzy, popiołu itp.). Ponieważ koncentracja świecących cząstek nie zmienia się skokowo, nie można wyróżnić krawędzi płomienia, jak np. w przypadku ciała stałego. Występujące w płomieniu zjawiska pochłaniania i rozpraszania promieniowania wprowadzają dodatkowe niejednoznaczności w określeniu granicy płomienia.

Sygnał pochodzący z systemu monitorowania płomienia, zawiera bardzo szeroką informację o procesie spalania, stąd jego analiza może przebiegać na wiele różnych sposobów. Jednym z nich jest analiza czasowo-częstotliwościowa [3,4], analiza temperatury [5], analiza obrazu [6] oraz analiza w oparciu o wyodrębnienie cech geometrycznych obrazu [1]. Nie przeprowadzono badań dotyczących klasyfikacji stanu procesu spalania z jednoczesną redukcją wymiarowości danych. Do analizy wybrano 20 cech geometrycznych obrazu i dokonano próby redukcji liczby cech ze wskazaniem do klasyfikacji stanu procesu. W tym celu porównano metody ekstrakcji cech: analizę składowych głównych, analizę czynnikową oraz

skalowanie wielowymiarowe dla obrazów zarejestrowanych w czasie testów spalania mieszanin pyłu węglowego i biomasy w warunkach laboratoryjnych przy różnych wydatkach powietrza i paliwa. Analizę przeprowadzono z wykorzystaniem pakietu Statistica.

Stanowisko pomiarowe

Testy spalania pyłu węglowego i biomasy zostały przeprowadzone na stanowisku laboratoryjnym, którego zasadniczym elementem jest cylindryczna komora spalania o długości 2,5m i średnicy ok. 0,7m. Wewnątrz komory znajduje się wykonany w skali 1:10 model wirowego palnika pyłowego oraz palnik olejowy. Obydwa palniki umieszczone są w jednej osi i mogą pracować jednocześnie. Mieszanka paliwowa wraz z powietrzem pierwotnym dostarczana jest do palnika pyłoprzewodem. Regulację wydatku węgla i biomasy dokonywana jest poprzez zmianę prędkości obrotowej jego podajnika. Na przebieg procesu spalania istotny wpływ ma ponadto wartość przepływu powietrza wtórnego. Komora spalania zaopatrzona jest w dwa podłużne otwory rozmieszczone po obu jej stronach, które umożliwiają obserwację płomienia. Widok komory spalania wraz z zaznaczonym umiejscowieniem kamery został przedstawiony na rysunku 1.



Rys.1. Widok ogólny stanowiska pomiarowego

Badania procesu spalania wykonano dla czterech wariantów z różnymi ustawieniami mocy cieplnej (P_{th}), współczynnika wydatku powietrza (λ) dla stałej zawartości

biomasy (20%). Podczas testu spalania, obrazy płomienia były rejestrowane przez system rejestracji w skład którego wchodziła kamera CCD współpracująca z obrazowodem oraz pamięć masowa. Obrazy płomienia w skali 24-bitowej RGB ograniczone do rozmiarów 1280x1024 pikseli, rejestrowane były szybkością do 150klatek/s (bez kompresji) i poddawane późniejszej analizie. Dokładne wartości mocy cieplnej i współczynnika wydatku powietrza zestawiono w tabeli 1.

Tabela 1. Warianty procesu spalania

Wariant	Moc [kW]	Współczynnik nadmiaru powietrza λ
1	250	0,75
2	250	0,65
3	250	0,85
4	300	0,75
5	300	0,65
6	300	0,85
7	400	0,75
8	400	0,65
9	400	0,85

Cechy obrazu

W celu analizy wyodrębniono obszar płomienia za pomocą metody aktywnego konturu. Do dalszej analizy wykorzystano dwadzieścia cech geometrycznych opisujących kształt obszaru zainteresowania (obszar płomienia) [7]. Pole powierzchni obszaru A zdefiniowano jako miara wielkości obiektu wyliczaną na podstawie sumy liczby pikseli wewnątrz i na brzegu danego obszaru. Długość konturu L obliczana jest jako suma odległości pomiędzy pikselami rozpatrywanego brzegu obszaru. Kolistość (sferyczność) C_{kl} , obszaru jest miarą jego podobieństwa do okręgu. Momenty centralne obszaru rzędu od 1 do 4 – I_1, I_2, I_3, I_4 . Współrzędne środka ciężkości obszaru S_x, S_y definiowane są jako znormalizowane momenty $M_{0,1}$ i $M_{1,0}$. Zawartość C_{zw} , będąca wielkością zależną od długości konturu płomienia L i pola powierzchni obszaru płomienia A. Prostokątność C_{prost} zdefiniowana jako różnicę pomiędzy powierzchnią prostokąta o identycznych jak rozpatrywany obszar momentach centralnych I_1, I_2 , a tym obszarem. Promienie R_A, R_B oraz kąt nachylenia elipsy φ wyznaczonych dla obszaru płomienia. Współrzędne dwóch najbardziej wysuniętych punktów obszaru płomienia $D_{1x}, D_{1y}, D_{2x}, D_{2y}$ i odległość między nimi D.

Analiza składowych głównych

Analiza składowych głównych jest to wielowymiarowa technika statystyczna, pozwalająca na odwzorowanie danych o dużej liczbie zmiennych w nowe dane, o mniejszej liczbie zmiennych określane jako składowe główne. Celem analizy składowych głównych jest zmniejszenie liczby wymiarów badanego zbioru danych przy zachowaniu maksymalnej zmienności w sensie wariancji/kowariancji. Jest to powszechnie stosowana technika nienadzorowanej ekstrakcji cech w przypadku posiadania próbek o nieznanym przynależności do klas. Dodatkowo, analiza PCA pozwala na wskazanie, która z pierwotnych zmiennych jest najbardziej reprezentowana w tworzonym modelu. Wymiar nowej przestrzeni zależy od tego, jak dużą część wariancji chcemy zachować. Wartości cech zostały wstępnie poddane standaryzacji. Główne składowe zostały obliczone na podstawie macierzy korelacji. Natomiast wariancję obliczono w oparciu o wzór bazujący na danych dla całej populacji, ponieważ analizowane przypadki stanowiły pewną populację generalną, a nie próbę pobraną z tej populacji.

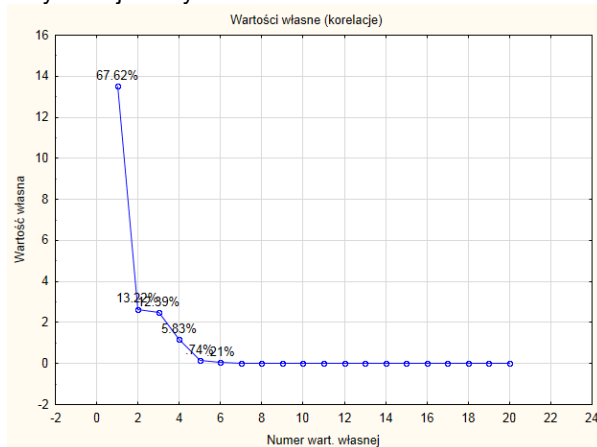
Tabela 2 przedstawia skumulowane wartości i wartości własne oraz skumulowane udziały procentowe i udziały procentowe sumy wariancji dla sześciu najistotniejszych składowych głównych.

Tabela 2. Statystyki analizy składowych głównych.

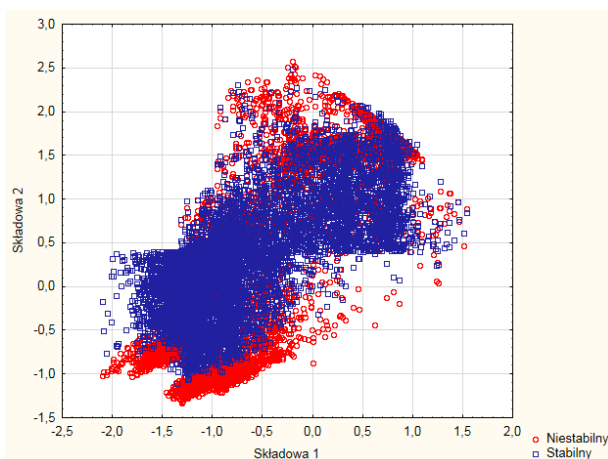
Numer wartości	Wartość własna	Proc. udział sumy wariancji	Skumulowana wartość własna	Skumulowany proc. udział sumy wariancji
1	62,8308	73,9185	62,8308	73,918
2	7,9217	9,3196	70,7525	83,238
3	5,4691	6,4343	76,2216	89,6726
4	2,2165	2,6077	78,4382	92,2803
5	1,5926	1,8736	80,0308	94,1540
6	1,1679	1,3740	81,1988	95,5280

Wartości własne przedstawionych składowych głównych, prezentują ich hierarchię w wyjaśnianiu zasobów informacyjnych zmiennych wejściowych. Procentowy udział sumy wariancji zmiennych wejściowych pokazuje, że pierwsza składowa główna przenosi ponad 73% informacji o analizowanych przypadkach zawartych w zmiennych wejściowych. Skumulowane wartości własne i skumulowane udziały procentowe wariancji wskazują, aby w dalszej analizie wykorzystać sześć pierwszych składowych głównych, ponieważ wartość własna jest większa od 1 oraz przenoszą one ponad 95% informacji zawartych w zmiennych wejściowych.

W celu określenia liczby składowych głównych przeznaczonych do dalszej analizy, skorzystano także z kryterium osypiska. Z rys. 2 wynika, że osypisko rozpoczyna się od piątej składowej głównej. Na tej podstawie do dalszej analizy uwzględniono cztery pierwsze składowe przenoszące ponad 92% informacji zawartych w zmiennych wejściowych.



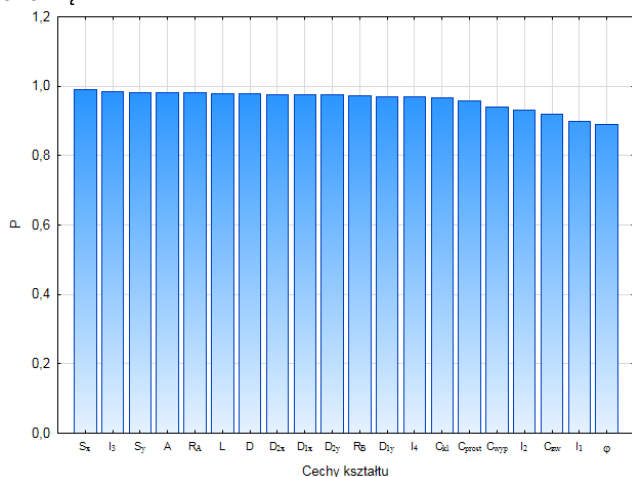
Rys.2. Wykres wartości własnych macierzy korelacji



Rys.3. Rzut przypadków dla dwóch składowych głównych

Na rys. 3 przedstawiono rzutowanie przypadków dla pierwszych dwóch składowych głównych, na podstawie wartości współrzędnych czynnikowych poszczególnych obserwacji. Nie uzyskano jednoznacznej separacji na przypadki stabilnego i niestabilnego procesu. Określono natomiast dwa obszary dla których niestabilność procesu jest jednoznaczna. Pierwsza składowa przyjmuje wtedy wartość mniejszą niż -0,5 oraz większą niż 2. Natomiast druga składowa waha się w zakresie -2 do 0,5.

Z wykonanej analizy składowych głównych wynika, że cztery pierwsze składowe główne są wystarczające do tego, by wyjaśniać większość zasobów informacyjnych zmiennych wejściowych. Skumulowany udział procentowy sumy wariancji dla tych składowych przekracza 95%. Analiza składowych głównych daje możliwość wyznaczenia wagi (istotności zmiennej P) każdej cechy w tworzonym modelu składowych głównych. Wartość wskaźnika P zbliżona do 1 oznacza istotny wpływ danej cechy na analizę.



Rys.4. Wartości wskaźnika P dla poszczególnych parametrów geometrycznych

Z rysunku 4 wynika, że badane cechy geometryczne płomienia są ze sobą skorelowane i mają duży wpływ na stan procesu spalania (wartość zbliżona do 1). Można wyodrębnić grupę cech mających mniejszy wpływ na stan procesu (C_{wyp} , I_2 , C_{zw} , I_1 , φ).

Analiza czynnikowa

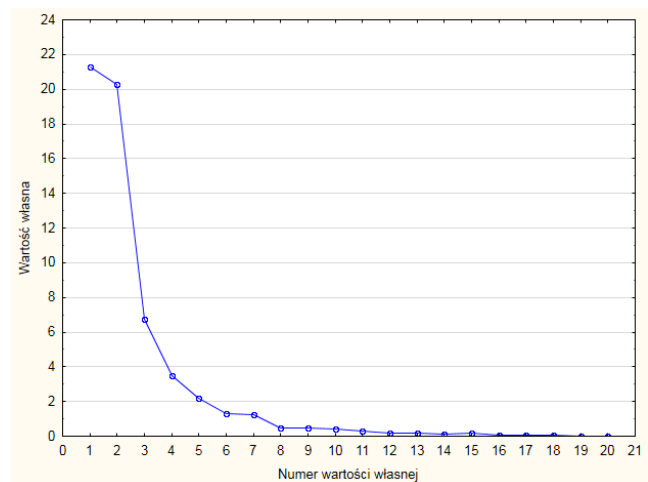
Jest to metoda badania struktury wewnętrznych zależności obserwacji wielowymiarowych. Każda zmienna obserwowalna przedstawiana jest jako kombinacja liniowa pewnej liczby nieobserwowalnych zmiennych, zwanych czynnikami, wspólnych dla całego zbioru zmiennych wejściowych oraz jednego nieobserwowalnego czynnika swoistego dla tej zmiennej – czynniki wspólne i czynniki swoiste nie są ze sobą skorelowane. Analiza czynnikowa prowadzi do znalezienia takiego zbioru czynników wspólnych, który pozwala na wyjaśnienie struktury powiązań między zmiennymi obserwowalnymi. Jako metodę szacowania ładunków czynnikowych wybrano metodę centroidalną. Maksymalną liczbę iteracji, w których szacowane będą zasoby zmienności wspólnej zmiennych wejściowych ustalono na pięćdziesiąt. Szacowanie może zostać zakończone bez osiągnięcia podana liczba iteracji w przypadku uzyskania w danej iteracji zmiany mniejszej niż przyjęta wartość równa 0,01. Założono także maksymalną liczbę czynników równą dwadzieścia oraz minimalną wartość własną równą zero.

Tabela 3 przedstawia skumulowane wartości i wartości własne oraz skumulowane udziały procentowe i udziały procentowe sumy wariancji dla sześciu dziesięciu

najistotniejszych składowych głównych. Wartości Według kryterium wartości własnych, do dalszej analizy należało wykorzystać siedem pierwszych czynników. Niestety, przy zastosowaniu tego kryterium nastąpiła utrata ponad 30% informacji zawartych w zmiennych wejściowych (sumy wariancji zmiennych wejściowych). Zasadność przyjęcia wspomnianej wyżej liczby czynników wynikała także z kryterium osypiska.

Tabela 3. Statystyki analizy czynnikowej.

Numer wartości	Wartość własna	Proc. udział sumy wariancji	Skumulowana wartość własna	Skumulowany proc. Udział sumy wariancji
1	21,2611	25,3108	21,2611	25,3108
2	20,2950	24,1607	41,5561	49,4715
3	6,7192	7,9991	48,2753	57,4707
4	3,4661	4,1264	51,7415	61,5971
5	2,2154	2,6374	53,9569	64,2345
6	1,2978	1,5451	55,2548	65,7796
7	1,2465	1,4839	56,5013	67,2635
8	0,4828	0,5748	56,9842	67,8383
9	0,4488	0,5343	57,4330	68,3727
10	0,3957	0,4711	57,8288	68,8438



Rys.5. Wykres wartości własnych macierzy korelacji

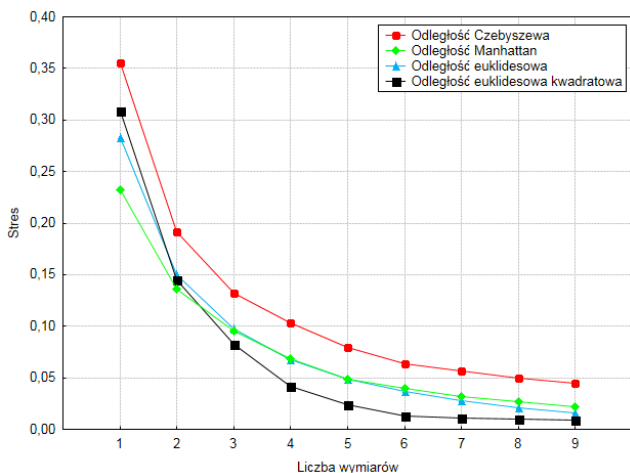
Na rysunku 5 widać, że osypisko zaczyna się od czynnika ósmego. Na tej podstawie, w dalszej analizie należało uwzględnić siedem czynników. Z takim wyborem wiązałaby się jednak utrata ok. 63,5% informacji zawartych w zmiennych wejściowych, czego nie można było zaakceptować. Wynika z tego, że liniowe metody redukcji wymiarowości nie są odpowiednie dla analizowanego zbioru danych.

Skalowanie wielowymiarowe

Skalowanie wielowymiarowe jest metodą statystyczną, która dokonuje redukcji liczby wymiarów na zasadzie dopasowania konfiguracji odległości punktów pomiarowych w przestrzeni wyników do przestrzeni zredukowanej. Analizy oparte na macierzach kowariancji lub korelacji pozwalają na przeprowadzenie klasycznych testów istotności, ale za to wymagają założenia o normalności rozkładu wielowymiarowego. Skalowanie wielowymiarowe nie wymaga żadnych założeń o rozkładach, pozwala na stosowanie dowolnej miary odległości i daje mniejszą liczbę wymiarów [8].

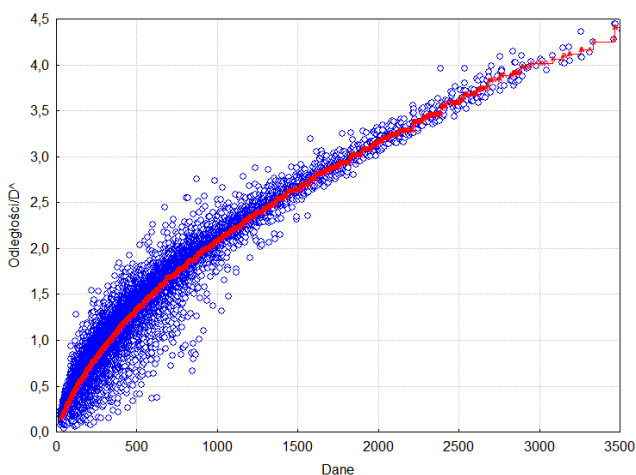
Wartości cech zostały wstępnie poddane standaryzacji. Następnie, utworzono cztery macierze odległości, przy czym dla każdej macierzy zastosowano inną miarę odległości. Wykorzystano odległości: euklidesową, euklidesową kwadratową, Manhattan, Czebyszewa.

Wykonano analizę dla dziewięciu wymiarów oraz czterech przyjętych miar odległości. Do oceny dopasowania macierzy odległości odtworzonych do macierzy odległości obserwowanych wykorzystano funkcję stresu. Im mniejsza jest wartość funkcji stresu, tym lepsze jest dopasowanie pomiędzy wspomnianymi wcześniej macierzami. W celu podjęcia decyzji dotyczącej docelowej liczby wymiarów, do której zostanie zredukowana przestrzeń wielowymiarowa, skorzystano z kryterium osypiska. Na rysunku 6 przedstawiono wykresy funkcji stresu w zależności od liczby wymiarów dla czterech macierzy odległości. Dla każdego rodzaju odległości osypisko zaczyna się od czwartego wymiaru. Na tej podstawie, przestrzeń wielowymiarowa została zredukowana do trzech wymiarów. Biorąc pod uwagę wartości funkcji stresu dla trzech wymiarów oraz zaproponowane oceny dopasowania można stwierdzić, że dla odległości Czebyszewa występuje dopasowanie słabe natomiast dla odległości euklidesowej kwadratowej dobre, a dla pozostałych miar odległości – średnie.



Rys.6. Wykres osypiska prezentujący zależność funkcji stresu od liczby wymiarów

Na rys. 7 przedstawiono diagram Shepparda dla przykładowej odległości euklidesowej kwadratowej. Przedstawiają one odtworzone odległości wykreślone na osi pionowej względem pierwotnych podobieństw wykreślonych na osi poziomej.



Rys.7. Diagram Shepparda dla odległości euklidesowej kwadratowej

Przedstawiony diagram pokazuje także przebiegi funkcji krokowych. Linie te przedstawiają wartości D^2 , czyli wyniki transformacji monotonicznej danych wejściowych. Jeśli

wszystkie odtworzone odległości znajdowałyby się na linii krokowej, to porządek rangowy odległości byłby dokładnie odtworzony. Wielkość odchylenia od linii krokowej jest interpretowana jako miara dopasowania. Postać diagramów potwierdza dobre dopasowanie dla odległości euklidesowej kwadratowej.

Podsumowanie

W artykule zaprezentowano wyniki redukcji wymiaru przestrzeni cech analizowanego zbioru danych. Zastosowano trzy metody ekstrakcji cech: analizę składowych głównych, analizę czynnikową oraz skalowanie wielowymiarowe. Ocena wyników dokonano na podstawie skategoryzowanych wykresów rozrzutu, prezentujących położenie obserwacji w dwuwymiarowym układzie nowych współrzędnych. Dla żadnej z zastosowanych metod nie uzyskano wyraźnych segmentów, które grupowałyby przypadki należące do klas stabilny i niestabilny. Jednak wykorzystując analizę składowych głównych uzyskano obszary dla których jednoznacznie można stwierdzić niestabilny stan procesu spalania. Ponadto, obliczone wartości wskaźnika P (istotności zmiennych) zbliżone do 1 dla wszystkich cech geometrycznych oznaczają istotny wpływ tych cechy na wyniki analizy. W związku z tym, wybór przedstawionych cech geometrycznych wydaje się zasadny do analizowania stanu procesu spalania. Dalsze prace skoncentrowane będą na ocenie stanu procesu współspalania węgla i biomasy za pomocą klasyfikatorów oraz możliwości ich wykorzystania do sterowania procesem spalania w czasie rzeczywistym.

Autorzy: mgr inż. Daniel Sawicki, dr hab. inż. Andrzej Kotyra prof. PL, Politechnika Lubelska, Instytut Elektroniki i Techniki Informatycznych, ul. Nadbystrzycka 38a 20-618 Lublin, E-mail: d.sawicki@pollub.pl, Khairullina Perdes, Kazakh National Technical University after K. I. Satpaev, Satpaeva 22, Almaty, Kazakhstan

LITERATURA

- [1] Ballester J, García-Armingol T.: Diagnostic techniques for the monitoring and control of practical flames. Prog. Energy Combust 2010, Volume: 36, 375-411
- [2] Cieszczyk S.: Non-Luminous Flame Temperature Determination Method Based on CO₂ Radiation, *Acta Physica Polonica*, (2014), Volume: 126, 1235-1240
- [3] Lu G, Yan Y, Colechin M, Hill R. Monitoring of oscillatory characteristics of pulverized coal flames through image processing and spectral analysis. IEEE Trans Instrum Meas 2006; Volume: 55, 226-31.
- [4] Cieszczyk, S., Lawicki T., Miaskowski, A.: The Curvelet Transform Application to the Analysis of Data Received from GPR Technique, *Elektronika I Elektrotechnika*, 2013, Volume: 19, Issue: 6, 99-102
- [5] Lu G, Yan Y. Temperature profiling of pulverized coal flames using multicolour pyrometric and digital imaging techniques. IEEE, 2006, 55, 1303-1308
- [6] González-Cencerrado A., Peña B., Gil A.: Coal flame characterization by means of digital image processing in a semi-industrial scale PF swirl burner, *Applied Energy*, 2012, 94, 375-384
- [7] Sawicki D.: Using PCA in flame vision monitoring system, *Proc. SPIE 9290*, Photonics Applications in Astronomy, Communications, Industry, and High-Energy Physics Experiments, 2014, 92900D
- [8] Woźniak A., Sikora J.: Wykorzystanie skalowania wielowymiarowego w analizie potencjału infrastrukturalnego gospodarstw rolnych w wybranych gminach województwa małopolskiego, *Infrastruktura I Ekologia Terenów Wiejskich*, 2006, nr.3, 161-178