

Krzysztof Gdawiec¹, Wiesław Kotarski²

¹ Instytut Matematyki, Uniwersytet Śląski
kgdawiec@math.us.edu.pl

² Instytut Informatyki, Uniwersytet Śląski
kotarski@math.us.edu.pl

FRAKTALNE ROZPOZNAWANIE OBIEKTÓW DWUWYMIAROWYCH

Streszczenie

Od momentu odkrycia fraktale znalazły szerokie zastosowania. Jednym z takich zastosowań jest wykorzystanie ich do rozpoznawania kształtów. W niniejszym artykule przedstawimy dwie metody fraktalnego rozpoznawania oparte o opis fraktalny powstający w wyniku fraktalnej kompresji obrazów. Następnie omówimy prostą modyfikację tych metod oraz wyniki przeprowadzonych badań.

1 Wstęp

W latach 70-tych dwudziestego wieku Benoit Mandelbrot przedstawił światu nowe pole matematyki. Nazwał je geometrią fraktalną (*fractus* – z łaciny ułamkowy). Geometria fraktalna zmienia to jak patrzymy na wszystko, daje narzędzia do opisu wielu naturalnych obiektów, których nie mogliśmy opisać przy pomocy klasycznej geometrii euklidesowej. Fraktale oczywiście istniały już znacznie wcześniej, ale były postrzegane jako obiekty wyjątkowe, matematyczne monstra. Mandelbrot obalił ten pogląd. Zauważył, że to co do tej pory wydawało się być wyjątkiem tak naprawdę jest regułą.

Fraktale znalazły szerokie zastosowanie w wielu dziedzinach m.in. kompresji obrazów, generowaniu linii brzegowych, gór, chmur, roślin czy też w medycynie i ekonomii. W niniejszym artykule zajmiemy się jednym z zastosowań fraktali, a mianowicie fraktalnym rozpoznawaniem. Na początku wprowadzimy pojęcie fraktala i kilka podstawowych definicji. Następnie w skrócie zaprezentujemy ideę fraktalnej kompresji obrazu, którą później wykorzystamy do fraktalnego rozpoznawania. Przedstawimy dwie metody fraktalnego rozpoznawania: metodę opartą o współczynniki układu PIFS [2, 3] oraz metodę Curtisa [6, 7, 8]. Następnie zaprezentujemy modyfikację obu tych metod i wyniki badań.

2 Fraktale

Fraktale można zdefiniować na różne sposoby, więc w tym rozdziale podamy definicję, którą będziemy się posługiwać w pracy. Najpierw jednak przedstawimy potrzebne pojęcia i oznaczenia.

Ustalmy dowolną przestrzeń metryczną zupełną (X, ρ) i oznaczmy przez $H(X)$ przestrzeń podzbiorów zwartych w X . W tej przestrzeni wprowadzamy funkcję $h : H(X) \times$

$H(X) \rightarrow \mathbb{R}_+$ daną wzorem:

$$h(R, S) = \max\{D(R, S), D(S, R)\},$$

gdzie $R, S \in H(X)$ oraz odwzorowanie $D : H(X) \times H(X) \rightarrow \mathbb{R}_+$ jest określone następująco:

$$D(R, S) = \max_{x \in R} \min_{y \in S} \rho(x, y).$$

Okazuje się, że funkcja h jest metryką tzw. metryką Hausdorffa, a przestrzeń $(H(X), h)$ jest przestrzenią metryczną zupełną [1]. Kolejnym ważnym pojęciem w naszych rozważaniach jest pojęcie układu IFS (Iterated Function System).

Układem IFS nazywamy zbiór $W = \{f_1, \dots, f_n\}$, gdzie $f_i : X \rightarrow X$ jest przekształceniem zwięzającym dla $i = 1, \dots, n$. Tak zdefiniowany układ IFS wyznacza tzw. operator Hutchinsona dany wzorem:

$$\forall_{A \in H(X)} \quad W(A) = \bigcup_{i=1}^n f_i(A) = \bigcup_{i=1}^n \{f_i(a) : a \in A\}.$$

Operator Hutchinsona jest zwięzający ze względu na to, że wyznaczający go układ IFS składa się z odwzorowań zwięzających [1].

Rozważmy teraz następujący ciąg rekurencyjny:

$$\begin{cases} W^0(A) = A \\ W^k(A) = W(W^{k-1}(A)), \quad k \geq 1 \end{cases}$$

gdzie $A \in H(X)$.

Z twierdzenia Banacha o punkcie stałym wynika następujące twierdzenie.

Twierdzenie 2.1. *Niech (X, ρ) będzie przestrzenią metryczną zupełną, a $W = \{f_1, \dots, f_n\}$ będzie układem IFS. Wówczas istnieje dokładnie jeden zbiór $B \in H(X)$ taki, że $W(B) = B$. Ponadto ciąg zdefiniowany powyżej jest zbieżny oraz*

$$\forall_{A \in H(X)} \quad \lim_{k \rightarrow \infty} W^k(A) = B.$$

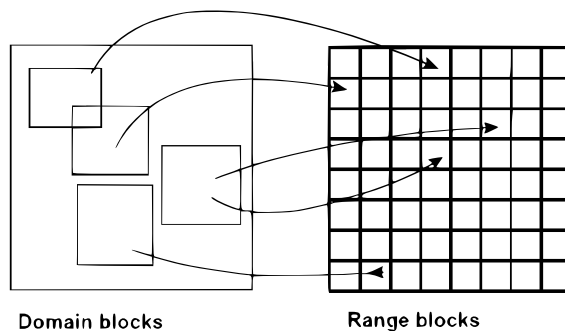
Granicę z powyższego twierdzenia będziemy nazywać *atraktorem* układu IFS W albo fraktalem.

3 Fraktalna kompresja obrazów

Do fraktalnego rozpoznawania obiektów potrzebujemy ich opisu fraktalnego. Aby znaleźć ten opis wykorzystamy fraktalną kompresję obrazów [4]. Poniżej przedstawimy najprostszy algorytm fraktalnej kompresji, który wykorzystamy później w procesie rozpoznawania.

Przy fraktalnej kompresji obrazów wprowadza się dodatkowo pojęcie układu PIFS (Partitioned Iterated Function System). *Układ PIFS* jest to zbiór $P = \{(F_1, O_1), \dots, (F_n, O_n)\}$, gdzie F_i jest przekształceniem zwięzającym, O_i jest obszarem obrazu, który przekształca odwzorowanie F_i ($i = 1, \dots, n$). W przypadku przestrzeni obrazów jako odwzorowania zwięzające wykorzystuje się przekształcenia afiniczne $F : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ postaci:

$$F \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{bmatrix} a & b & 0 \\ c & d & 0 \\ 0 & 0 & s \end{bmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} + \begin{pmatrix} e \\ f \\ o \end{pmatrix},$$



Rysunek 1. Fraktalna kompresja obrazu.

gdzie współczynniki $a, b, c, d, e, f \in \mathbb{R}$ określają przekształcenia geometryczne, a współczynniki $s, o \in \mathbb{R}$ odpowiadają za kontrast i jasność. W dalszym ciągu przy rozpoznawaniu ograniczamy się wyłącznie do obrazów binarnych, więc w tym wypadku przekształcenia afiniczne przyjmują uproszczoną postać:

$$F\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} a_5 \\ a_6 \end{bmatrix}.$$

Algorytm kompresji wygląda następująco. Dzielimy obraz na ustaloną z góry liczbę nienachodzących na siebie bloków przeciwdziedzinowych (Range Blocks). Tworzymy listę bloków dziedzinowych (Domain Blocks). Lista składa się z nachodzących na siebie obszarów obrazu, większych od bloków przeciwdziedzinowych (najczęściej dwa razy większych) oraz przekształconych przez następujące przekształcenia:

$$\begin{aligned} \begin{bmatrix} x' \\ y' \end{bmatrix} &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}, & \begin{bmatrix} x' \\ y' \end{bmatrix} &= \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}, \\ \begin{bmatrix} x' \\ y' \end{bmatrix} &= \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}, & \begin{bmatrix} x' \\ y' \end{bmatrix} &= \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}. \end{aligned}$$

Następnie, dla każdego bloku przeciwdziedzinowego R szukamy bloku dziedzinowego D takiego, że $\rho(R, F(D))$ jest najmniejsze, gdzie ρ jest metryką, a F jest przekształceniem wyznaczonym przez położenie bloków R i D , wielkość tych bloków względem siebie oraz jedno z czterech przekształceń zdefiniowanych powyżej. Jest to najbardziej czasochłonny krok w algorytmie.

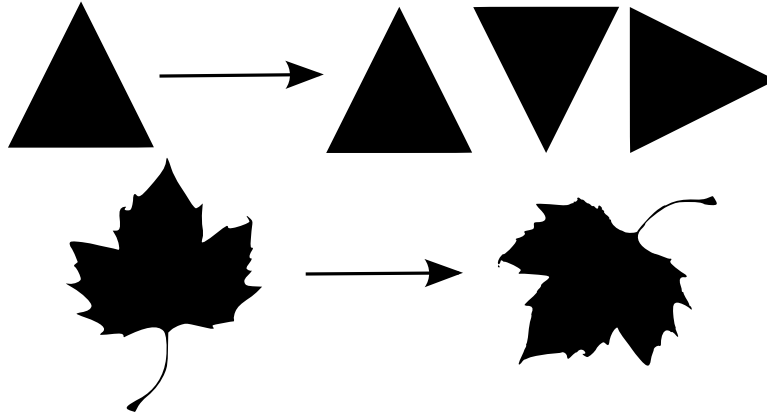
Rysunek 1 przedstawia ideę fraktalnej kompresji obrazów.

4 Fraktalne rozpoznawanie

W niniejszym rozdziale zostaną przedstawione dwie metody fraktalnego rozpoznawania kształtów dwuwymiarowych – metoda korzystająca ze współczynników układu PIFS oraz metoda opierająca się na niezmienniczości atraktora ze względu na układ IFS, który go generuje.

4.1 Metoda oparta o współczynniki układu PIFS

Metoda oparta o współczynniki układu PIFS została przedstawiona w pracach [2, 3]. Została ona wykorzystana do rozpoznawania twarzy na obrazach w odcieniach szarości.



Rysunek 2. Przykłady nadania obiektowi odpowiedniej orientacji.

W przypadku rozpoznawania tą metodą obiektów dwuwymiarowych obrazy w odcieniach szarości zamieniamy na obrazy binarne. Metoda składa się z następujących etapów:

- binaryzacja obrazu,
- nadanie obiektowi odpowiedniej orientacji,
- znalezienie znormalizowanego układu PIFS tzn. takiego dla którego przestrzenią jest $[0, 1]^2$,
- rozpoznawanie, w którym
 - obliczamy miarę podobieństwa między otrzymanym układem PIFS, a układem PIFS z bazy,
 - powtarzamy poprzedni punkt dla wszystkich układów PIFS z bazy,
 - wybieramy obraz z bazy, dla którego miara podobieństwa jest najmniejsza.

Nadanie obiektowi odpowiedniej orientacji polega na obrocie obiektu o taki kąt, że po obrocie spełnia on następujące warunki: pole najmniejszego prostokąta zawierającego obiekt jest najmniejsze, wysokość tego prostokąta jest mniejsza od szerokości oraz lewa połowa obiektu ma co najmniej tyle pikseli co prawa.

Na rysunku 2 przedstawiono przykłady dobrych orientacji dla dwóch obiektów. W przypadku trójkąta widzimy trzy różne orientacje. Jeśli dodajemy do bazy obiekt o kilku różnych orientacjach, to wtedy dla każdej z orientacji znajdujemy znormalizowany układ PIFS i dodajemy go do bazy. W przypadku obiektu o kilku różnych orientacjach do procesu rozpoznawania wybieramy tylko jedną z nich.

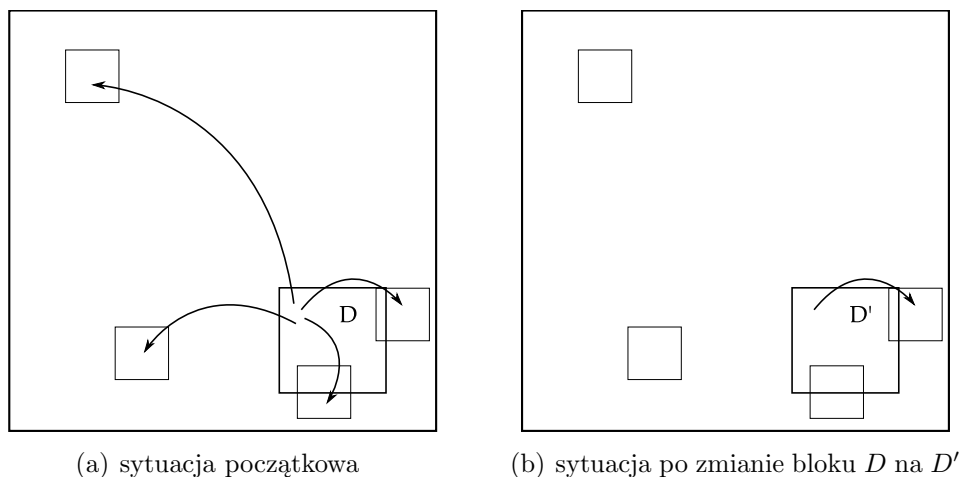
W procesie rozpoznawania wykorzystujemy następującą miarę podobieństwa:

$$d(P_1, P_2) = \sum_{i=1}^n \rho(F_i, G_i),$$

gdzie $P_1 = \{(F_1, O_1), \dots, (F_n, O_n)\}$, $P_2 = \{(G_1, Q_1), \dots, (G_n, Q_n)\}$ są układami PIFS dla obrazów oraz

$$\rho(F, G) = \sum_{i=1}^6 |a_i - b_i|,$$

gdzie a_i, b_i są to odpowiadające sobie współczynniki przekształceń F i G .



Rysunek 3. Przekształcenia przy obrazie traktowanym jako całość.

4.2 Metoda Curtisa

Nazwa tej metody pochodzi od nazwiska jej twórcy. Metoda Curtisa została przedstawiona w pracach [6, 7, 8]. Wykorzystuje ona fakt, że atraktor jest niezmiennikiem swojego układu IFS, o czym mówi wcześniej zaprezentowane twierdzenie 2.1. Metoda składa się z następujących etapów:

- binaryzacja obrazu,
- nadanie obiektowi odpowiedniej orientacji,
- znalezienie znormalizowanego układu PIFS (w przypadku dodawania do bazy),
- rozpoznawanie, w którym
 - przekształcamy obraz za pomocą układu PIFS z bazy,
 - liczymy błąd średniokwadratowy dla tak powstałego obrazu i obrazu oryginalnego,
 - powtarzamy poprzednie dwa punkty dla każdego układu PIFS z bazy,
 - wybieramy obraz z bazy, dla którego błąd jest najmniejszy.

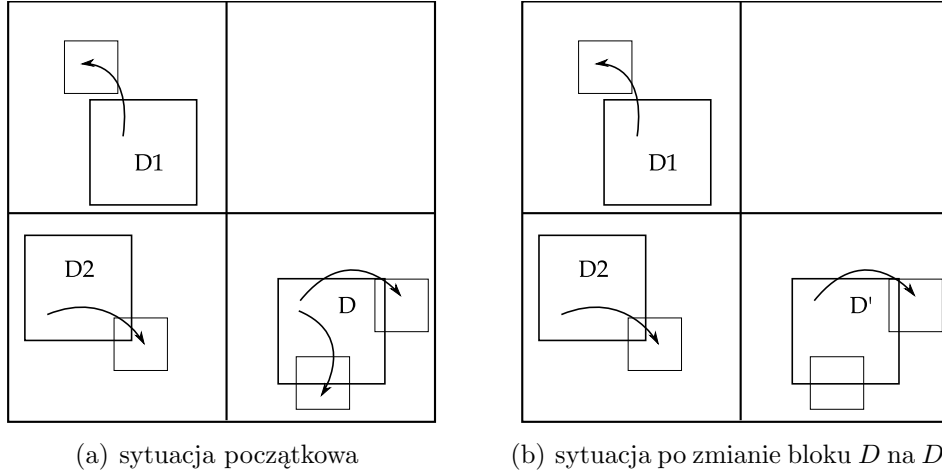
5 Modyfikacja metod

Przedstawimy teraz modyfikację prezentowanych wcześniej metod fraktalnego rozpoznawania. Modyfikacja opiera się na pewnej obserwacji, którą wyjaśnimy korzystając z rysunku 3.

Na rysunku 3(a) przedstawiona została sytuacja odpowiadająca metodom prezentowanym wcześniej kiedy to jednemu blokowi D odpowiadało kilka bloków przeciwdziedziny. Załóżmy teraz, że blok D zostaje zamieniony na blok D' , jak na rysunku 3(b). Blok D' nie odpowiada już niektórym blokom przeciwdziedziny, jak wcześniej. Powoduje to, że zmieniają się przekształcenia w układzie PIFS.

Podzielmy teraz obraz na kilka nienachodzących na siebie podobrazów np. na 4 podobrazy (rysunek 4(a)). Ponownie rozważmy ten sam blok dziedzinowy D i te same bloki

przeciwdziedziny. Tym razem blok D odpowiada tylko blokom należącym do tego samego podobrazu, blokom przeciwdziedziny z innych podobrazów odpowiadają inne bloki dziedziny. Ponownie założmy, że blok D ulega zamianie na blok D' (rysunek 4(b)). Zmiana bloku ma wyłącznie wpływ na podobraz, na którym znajduje się blok D' . Dopasowanie w innych podobrazach nie ulega zmianie. A zatem lokalna zmiana bloku D ma jedynie lokalny wpływ na przekształcenia układu PIFS, a nie globalny jak to miało miejsce w przypadku traktowania obrazu jako całości.



Rysunek 4. Przekształcenia po wprowadzeniu podziału obrazu na podobrazy.

Modyfikacja metod będzie polegać na tym, że w kroku, na którym odbywa się fraktalna kompresja obrazu podzielimy na kilka nienachodzących na siebie obrazów. Następnie każdy podobraz poddawany jest fraktalnej kompresji. Układem PIFS dla obrazu będzie teraz suma układów PIFS dla podobrazów. Rozpoznawanie w przypadku metody Curtisa pozostaje bez zmian, zaś w przypadku metody opartej o współczynniki układu PIFS wprowadzamy nową miarę podobieństwa:

$$S(P_1, P_2) = \sum_{i=1}^n d(T_i, Q_i),$$

gdzie $P_1 = \{T_1, \dots, T_n\}$, $P_2 = \{Q_1, \dots, Q_n\}$, n jest liczbą podobrazów, na które dzieliśmy obrazy, $T_1, \dots, T_n, Q_1, \dots, Q_n$ są układami PIFS podobrazów.

6 Wyniki

Eksperymenty przeprowadzono na trzech zbiorach testowych. W pierwszym zbiorze znalazły się obiekty bazowe poddane przekształceniom elementarnym (obrót, skalowanie, translacja). W drugim zbiorze znajdowały się obiekty poddane przekształceniom elementarnym i dodatkowo wprowadzono niewielkie lokalne zniekształcenia kształtu. W ostatnim, trzecim zbiorze podobnie jak poprzednio znalazły się obiekty poddane przekształceniom elementarnym oraz dodatkowo obiekty poddano lokalnie dużym zniekształceniom.

W przypadku testu każdej z metod do fraktalnej kompresji obrazu użyto 256 przekształceń. W zmodyfikowanych metodach zastosowano podział na 4 (2×2) oraz 16 (4×4) podobrazów. Każdy podobraz był kompresowany przy użyciu 64 (8×8) przekształceń dla podziału na 4 podobrazy oraz 16 (4×4) przekształceń dla podziału na 16 podobrazów.

Tabela przedstawia uzyskane wyniki.

Metoda	Skuteczność [%]			
	Elemen.	Lok. niewielkie	Lok. duże	Średnia
PIFS (oryginalna)	55.55	60.00	37.78	51.11
PIFS (2 × 2, 8 × 8)	66.68	68.89	37.78	57.78
PIFS (4 × 4, 4 × 4)	91.11	86.67	40.00	72.59
Curtis (oryginalna)	86.67	86.67	66.67	80.00
Curtis (2 × 2, 8 × 8)	86.67	86.67	73.33	82.22
Curtis (4 × 4, 4 × 4)	86.67	86.67	75.55	82.96

7 Wnioski i dalsze badania

Z przeprowadzonych testów wynika, że za pomocą zaproponowanej modyfikacji uzyskuje się znaczący wzrost skuteczności rozpoznawania w przypadku metody opartej o współczynniki układu PIFS oraz niewielki wzrost w przypadku metody Curtisa dla obiektów lokalnie znacznie zniekształconych. Ponadto, w przypadku metody opartej o współczynniki PIFS zanotowano wzrost szybkości rozpoznawania w porównaniu do oryginalnej metody. Jest to spowodowane tym, że w przypadku podziału na podobrazy i ich kompresję przeszukujemy mniejsze listy bloków dziedzinowych.

Dalsze badania będą polegały na uwzględnieniu w mierze podobieństwa liczby pasujących do siebie podobrazów, co może zaowocować dalszym wzrostem skuteczności rozpoznawania. Przeprowadzimy również testy dla innych podziałów na niezależne obrazy żeby zbadać wpływ różnych podziałów na skuteczność metod. Ponadto wprowadzimy prezentowaną modyfikację do innych znanych metod fraktalnego rozpoznawania.

Literatura

- [1] M. Barnsley, *Fractals everywhere*, Academic Press, Boston, (1988).
- [2] S. Chandran, S. Kar, *Fractal based unstructured object recognition*, <http://www.cse.iitb.ac.in/~sharat/papers/fractal.ps.gz>
- [3] S. Chandran, S. Kar, *Retrieving faces by the PIFS fractal code*, Workshop on Applications of Computer Vision, pp. 8-12, (2002).
- [4] Y. Fisher, *Fractal image compression*, Springer-Verlag, New York, (1995).
- [5] S. Mozaffari, K. Faez, M. Ziaratban, *Character representation and recognition using quadtree-based fractal encoding scheme*, IEEE ICDAR, pp. 819-823, (2005).
- [6] G. Neil, K. M. Curtis, *Scale and rotationally invariant object recognition using fractal transformations*, IEEE ICASP, 6, pp. 3458-3461, (1996).
- [7] G. Neil, K. M. Curtis, *Shape recognition using a novel fractal technique*, IEEE ICECS, pp. 724-727, (1996).
- [8] G. Neil, K. M. Curtis, *Shape recognition using fractal geometry*, Pattern Recognition, vol. 30, No. 12, pp. 1957-1969, (1997).
- [9] P-Y Wu, *Fast fractal image compression*, ITCC, pp. 54-61, (2000).