

NEURAL IMAGE DATA COMPRESSION IN THE PROCESS OF IDENTIFICATION OF SELECTED AGRICULTURAL OBJECTS

Summary

The aim of this work was to discuss issues relating to synthetic neural compression of graphical data using a topology of artificial neural networks. Utilitarian aspect of the analysis was the implementation of the proposed image compression technology in the original system "Sunflower.b", supporting the processing of photos of selected agricultural objects.

NEURONOWA KOMPRESJA DANYCH GRAFICZNYCH W PROCESIE IDENTYFIKACJI WYBRANYCH OBIEKTÓW ROLNICZYCH

Streszczenie

Celem pracy było omówienie problematyki dotyczącej neuronowej kompresji danych graficznych z wykorzystaniem wybranej topologii sztucznej sieci neuronowej. Aspektem użytkowym przeprowadzonej analizy była implementacja proponowanej metodyki kompresji obrazów graficznych w wytworzonym, oryginalnym systemie informatycznym „Sunflower.b”, wspomagającym proces przetwarzania zdjęć wybranych obiektów rolniczych.

1. Wprowadzenie

Kompresja obrazu jest mechanizmem umożliwiającym redukcję wielkości pliku z danymi graficznymi, co pozwala na ich wydajniejsze wykorzystanie, np. przy składowaniu danych, przetwarzaniu lub transmisji. Ważnym aspektem realizowanego procesu kompresji jest zachowanie (w przetransformowanych danych graficznych) zadowalającego poziomu jakości (ilości) informacji w stosunku do danych zakodowanych w pliku oryginalnym. Sprowadza się to do stratnego lub bezstratnego procesu kompresji posiadanego obrazu. Kompresję ogólnie można zdefiniować jako proces przekształcenia źródłowego zbioru danych w reprezentację o mniejszej liczbie bitów. Proces odwrotny (czyli rekonstrukcja oryginalnego zbioru danych) jest nazywany dekompresją. Obecnie przy tworzeniu nowoczesnych narzędzi do redukcji rozmiaru obrazu cyfrowego, pojęcie kompresji nabiera coraz większego znaczenia i podlega istotnym zmianom formalnym. Według rozszerzonej definicji, kompresja polega również na wyznaczeniu możliwie użyteczniej, w danym zastosowaniu, przesyłanej lub gromadzonej informacji wejściowej.

Cele kompresji mogą być różnorodne. Na ogół zależy to od rodzaju zastosowań i charakteru posiadanych danych. Zwykle podczas wyboru oraz tworzenia metod kompresji dąży się do uzyskania możliwie największej efektywności. Najbardziej powszechnym rozumieniem pojęcia efektywności kompresji jest minimalizacja rozmiaru danych skompresowanych w stosunku do zbioru oryginalnego. Miarą tej efektywności jest stopień kompresji CR (ang. *Compression Ratio*), procent kompresji CP (ang. *Compression Percentage*) oraz średnia bitowa BR (ang. *bit rate*). Stopień kompresji to stosunek liczby bitów reprezentacji wejściowej do liczby bitów danych po kompresji. Procent kompresji stosuje się do oceny skuteczności archiwizatorów. Średnia bitowa jest użyteczna przy zastosowaniach transmisyjnych i charakteryzuje

rozmiar danych wyjściowych. Do efektywności kompresji można zaliczyć także minimalizację czasu procesu kompresji oraz dekompresji.

Metody kompresji można podzielić na dwie zasadnicze kategorie: bezstratne i stratne. W kompresji bezstratnej, ciąg danych uzyskany po transformacji jest identyczny ze zbiorem źródłowym (z dokładnością do pojedynczego bitu). Dlatego kompresję bezstratną można inaczej nazwać odwracalną lub bezstratną numerycznie. Wykorzystuje się ją do takich zastosowań, które wymagają wiernej rekonstrukcji danych źródłowych, takich jak np. archiwizacja dokumentów tekstowych, obrazów medycznych i wielu innych.

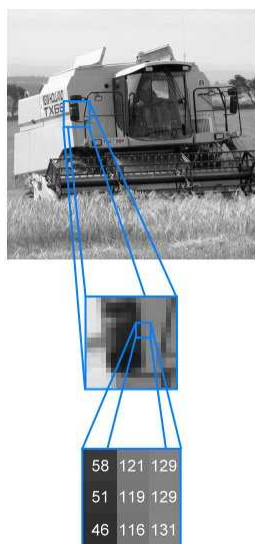
Kompresja stratna nie pozwala odtworzyć danych źródłowych z dokładnością do jednego bitu. Dlatego nazwa się ją kompresją stratną numerycznie lub nieodwracalną. Ma ona ważne zastosowanie w analizie danych graficznych, ponieważ podczas stratnej kompresji obrazu mamy do czynienia z pojęciem tzw. bezstratności wizualnej. Oznacza to, że usunięcie części informacji zawartej w postaci prezentowanego obrazu, nie powoduje dostrzegalnych zmian w obserwowanym obrazie (i jest na ogół niezauważalne przez obserwatora). Fakt ten wynika bezpośrednio ze specyfiki budowy oka oraz sposobu akwizycji oraz przetwarzania informacji graficznej, realizowanego przez człowieka w procesie wizualnej percepcji otaczającej rzeczywistości. Warto zauważyć, że kompresja obrazu może stać się zauważalna, jeżeli zostaną zmienione warunki obserwacji lub obraz zostanie poddany dalszemu przetworzeniu.

Podsumowując, istotą kompresji obrazu jest zmniejszenie rozmiaru danych nie zmieniając (lub nieznacznie zmieniając) przy tym ilości zawartych w nim informacji. Dane graficzne często charakteryzują się nadmiarowością (np. zawierają powtarzające się schematy, barwy itd.). Celem kompresji jest zakodowanie tych danych w inny sposób, zmniejszając przy tym rozmiar pliku

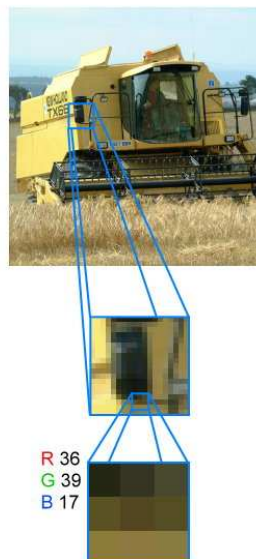
wynikowego. Ważne jest zastosowanie odpowiedniego systemu kompresji, który pozwoli na utratę nieistotnych danych przy równoczesnym zachowaniu danych o kluczowym znaczeniu.

2. Reprezentacja rastrowego obrazu cyfrowego

Obraz rastrowy zapisywany jest w pamięci komputera za pomocą bitów. Strukturalnie jest to tabela zawierająca informacje o pozycji piksela oraz wartość koloru do niego przypisanego. Jedną z istotnych cech charakterystycznych danych graficznych zapisanych w pamięci maszyny cyfrowej jest rozdzielczość, czyli ilość pikseli zapisanych w pionie i w poziomie (w postaci ww. tabeli). Najprostszym przykładem takiej tabeli jest obraz 2-bitowy, gdzie „1” oznacza kolor czarny, a „0” kolor biały. Gdy mamy do czynienia z obrazem w odcieniach szarości, można posłużyć się mapą, w której każdy piksel ma przypisaną liczbę odpowiadającą poziomowi jasności danego punktu, zawierającą się w przedziale od 0 do 255 (8-bitowy zapis poziomu szarości). Tzw. „grayscale”, czyli skala szarości wykorzystuje tylko jeden kanał barwny: czarny (oraz jego odcienie). W pamięci komputera każdy piksel reprezentowany jest zatem przez 8 bitów ($2^8 = 256$).



Rys. 1. Zapis obrazu w odcieniach szarości
Fig. 1. Record of image in grayscale



Rys. 2. Zapis obrazu RGB
Fig. 2. Record of image in RGB format

W przypadku obrazu kolorowego w każdej komórce zawarte są informacje o stopniu jasności jednego z trzech podstawowych kolorów. Jednym z najpopularniejszych modeli koloru jest model przestrzeni barw RGB, gdzie R (ang. Red) to kolor czerwony, G (ang. Green) to kolor zielony, a B (ang. Blue) niebieski. Podobnie jak w zdjęciach z 8-bitową paletą odcieni szarości tak i tutaj można dobrać ilość używanych kolorów.

Standarem jest paleta 16-bitowa ($2^{16} = 65.536$), która gwarantuje bardzo dobre-odzworowanie barw.

3. Format zapisu cyfrowego obrazu rastrowego

Obraz cyfrowy (np. w postaci zdjęcia) można zapisać w pamięci komputera w postaci pliku graficznego. Wybór formatu zależy m.in. od rodzaju grafiki oraz od późniejszego przeznaczenia wytworzonego obrazu. W przypadku grafiki rastrowej, przez lata rozwoju informatyki, wytworzono wiele różnych formatów, jednak najpopularniejszymi obecnie są: BMP, JPG, GIF, TIF oraz PNG.

BMP (bitmapa) to standard opracowany pierwotnie jako natywny dla systemu OS/2 (i wykorzystany później przez Microsoft), który szczególnie nadaje się do obrazów przeznaczonych do edycji. Format ten zawiera w sobie bezstratną kompresję RLE informacji o kolorach. Jednak kompresja ta w większości przypadków nie jest wykorzystywana. Jego największą wadą jest duży rozmiar pliku wyjściowego oraz fakt, że obsługuje wyłącznie model koloru RGB. Plik BMP składa się z trzech części:

- nagłówka,
- palety typu RGB (opcjonalnie),
- oraz danych obrazowych.

Nagłówek pliku BMP składa się z 54 bajtów i zawiera w swojej strukturze min. informacje o: sygnaturze pliku, długości całego pliku w bajtach, szerokości obrazu w pikselach, wysokości obrazu w pikselach, liczbie warstw kolorów, liczbie bitów na piksel, liczbie kolorów w paletce.

Po części nagłówkowej pliku BMP zazwyczaj znajdują się paleta. Można tu wyodrębnić 2 przypadki:

- w trybach 8-bitowych paleta zbudowana jest z reprezentacji kolorów, gdzie każdy kolor opisany jest za

pomocą 4 bajtów (B, G, R, nieużywany lub alpha: każdą ze składowych koloru należy podzielić przez 4),

- w trybach o większej liczbie kolorów (niż 256) paleta nie występuje, a kolor piksela w obrazie zapisywany jest przy pomocy numeru koloru, który jest wyliczany ze wzoru:

$$\text{numer koloru} = \mathbf{R} + 256 * \mathbf{G} + 65536 * \mathbf{B}, \quad (1)$$

gdzie:

R - wartość składowej czerwonej,

G - wartość składowej zielonej,

B - wartość składowej niebieskiej.

W następnej kolejności w pliku usytuowane są informacje opisujące dane obrazowe.

JPEG (ang. *Joint Photographic experts Group*) to standard stratnej kompresji obrazów rastrowych. Ten format pliku graficznego jest dedykowany do zapisu zdjęć i publikacji w Internecie. Małe rozmiary pliku zależą od stopnia kompresji, który można „ręcznie” ustawić przed zapisaniem pliku. JPG oferuje 24 - bitową głębię barw, co daje możliwość wyświetlenia 16,7 miliona kolorów ($2^{24} = 16.777.216$).

Format GIF (ang. *Graphics Interchange Format*) pozwala na użycie tylko 256 kolorów. Jego zaletą są małe rozmiary pliku oraz możliwość używania „przezroczystych” kolorów. Używany jest do publikacji w Internecie. W wyniku jedynie 8-bitowej palety barw GIF nie nadaje się do zapisu obrazów o wysokiej rozdzielczości oraz bogatej kolorystyce. Jest to format rastrowy kompresowany bezstratnie. Z racji tego, że większość obrazów posiada paletę 24-bitową (ok. 16,7 milionów kolorów), przed zapisaniem obrazu GIF następuje szereg procesów stratnych, do których należy zaliczyć:

- wyznaczenie maksimum 256 kolorów, które jak najwierniej oddają oryginalny zestaw kolorów obecnych w obrazie; jest tzw. kwantyzacja kolorów,
- zapis tych kolorów do palety (może być również stratny),
- opcjonalnie: rozsiewanie lub rozpraszenie (ang. *dithering*).

TIF (ang. *Tagged Image file Format*) format o 64-bitowej głębi barw. Używany jest do zapisywania obrazów o dużej rozdzielczości. Często wykorzystywany w drukarniach. TIF to format rastrowy wykorzystujący kompresję bezstratną z wykorzystaniem algorytmu LZW - Lempel-Ziv-Welch [5] będącego modyfikacją kodowania LZ78 opracowanego w 1978 roku przez Jacoba Ziva i Abrahama Lempela [6].

PNG (ang. *Portable Network Graphics*), następcą formatu GIF, wytworzony głównie z myślą o publikacji w Internecie, obsługujący przezroczystość (24-bitowa paleta barw). Jest to format rastrowy wykorzystujący kompresję bezstratną.

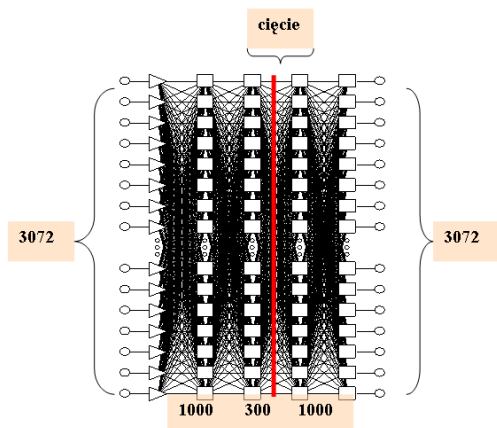
4. Problem redukcji wymiaru z wykorzystaniem modelowania neuronowego

Popularną metodą redukcji wymiaru składowych wektora danych, często stosowaną w statystycznej obróbce danych empirycznych, jest macierzowa analiza składowych głównych, zwana w skrócie metodą PCA (ang. *Principal Components Analysis*). W istocie jest to transformacja macierzowa, która redukuje liczbę zmiennych (cech) do

dowolnej, zadanej wartości w taki jednak sposób, aby zachować maksymalną wariancję danych wejściowych. W tak przetworzonych danych oczekuje się zachowania maksymalnej ilości informacji, zbliżonej do tej, jaką zawiera oryginał. Metoda PCA wyznacza, w swojej istocie, transformację liniową sprowadzającą się do rotacji danych do nowego układu współrzędnych. Układ ten utworzony zostaje przez wektory własne macierzy autokorelacji wyznaczonej dla analizowanych danych [1].

Istotnym problemem aplikacyjnym, związanym z analizą głównych składowych, jest jej liniowy charakter. Z tego też powodu nie może ona być użyta do redukcji wymiaru danych wejściowych w przypadku rozwiązywania problemów nieliniowych. Alternatywne podejście, wolne od wskazanego wyżej ograniczenia, polega na wykorzystaniu szczególnej topologii jaką reprezentują autoasocjacyjnej sieci neuronowej. Sieci autoasocjacyjne, występujące na ogół w postaci perceptronów wielowarstwowych typu MLP (ang. *MultiLayer Perceptron*), odtwarzają na swoich wyjściach wartości, które zostały podane na wejściu. Sens użycia sieci autoasocjacyjnej polega na tym, że warstwy ukryte liczą mniej neuronów niż warstwy wejściowa oraz wyjściowa. Fakt ten powoduje, że w trakcie pracy nauczonej sieci neuronowej następuje, w jej strukturze wewnętrznej, redukcja liczby danych zawartych w wektorze wejściowym. W szczególności, technika ta może stanowić efektywne narzędzie do stratnej kompresji różnego rodzaju danych.

Pomysł zrealizowania nieliniowej redukcji wymiaru polega w istocie na zastosowaniu topologii sieci neuronowej zbudowanej z neuronów nieliniowych, tzn. posiadających w swojej strukturze nieliniowy potencjał membranowy PSP (ang. *Post-Synaptic Potential*), np. neuronów sigmoidalnych. Jednak, aby w pełni wykorzystać możliwości, jakie stwarza neuronowy model nieliniowy, potrzebna jest sieć o strukturze posiadającej dodatkową warstwę ukrytą neuronów, dla każdej z realizowanych transformacji, zarówno dla kompresji jak i dla dekompresji. Dlatego tworząc sieć do nieliniowej kompresji danych należy wygenerować topologię sieci neuronowej o pięciu warstwach [2]. Ukryta warstwa środkowa jest warstwą redukującą wymiar sygnału wejściowego, zaś warstwa znajdująca się pomiędzy nią i warstwą wyjściową dokonuje właśnie wymaganej nieliniowej kompresji wejściowych danych. Odpowiednio dwuwarstwowa struktura sieci, znajdująca się pomiędzy warstwą ukrytą a warstwą wyjściową, realizuje transformację odwrotną dekompresując, „zagęszczony uprzednio”, sygnał. Przykładową autoasocjacyjną sztuczną sieć neuronową typu MLP, o 3072 wejściach i 3072 wyjściach, przedstawia rys. 3. [1].



Rys. 3. Autoasocjacyjna sieć neuronowa typu MLP
Fig. 3. Autoassociative neural network type MLP

5. Implementacja wygenerowanego kodu źródłowego autoasocjacyjnej sieci neuronowej

W celu wygenerowania zbiorów uczących (niezbędnych w procesie uczenia autoasocjacyjnej sieci neuronowej) wykorzystano specjalnie do tego celu wytworzony, oryginalny system informatyczny: „Przetwarzanie Obrazu v.1.0” [4]. Program ten pozwala m.in. na konwersję zdjęć do postaci zbioru uczącego, akceptowalnego przez symulator sztucznych sieci neuronowych, zaimplementowany w komercyjnym pakiecie statystycznym „STATISTICA v. 7.0”.

Po zaimportowaniu wygenerowanych (przez system informatyczny „Przetwarzanie Obrazu” v.1.0”) plików do edytora danych symulatora sieci neuronowych pakietu „STATISTICA v. 7.0”, posługując się standardową procedurą, wytworzono autoasocjacyjną sieć neuronową typu MLP składającą się z 5-ciu warstw i posiadającą następującą strukturę: 256 – 120 – 55 – 120 – 256.

W celu implementacji sieci w tworzonej aplikacji „Sunflower.b” skorzystano z modułu „Generator kodu” będącego elementem pakietu „STATISTICA v.7.0”. Narzędzie to pozwala na efektywne wygenerowanie kodu

źródłowego wytworzonej topologii neuronowej w języku C, Visual Basic lub w C#.



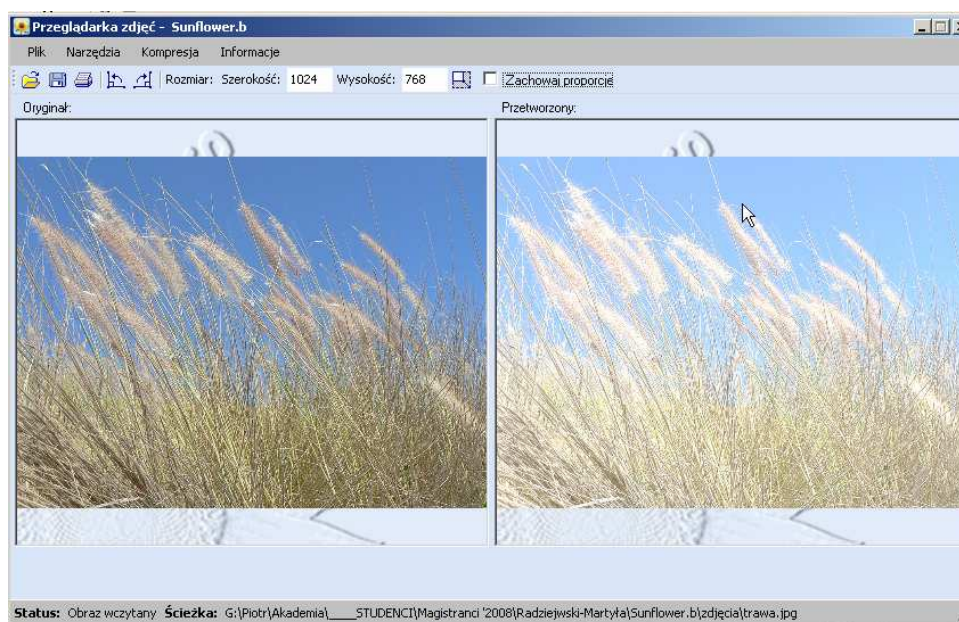
Rys. 4. Okno powitalne programu „Sunflower.b”
Fig. 4. Intro of computer system “Sunflower.b”

6. Obsługa aplikacji Sunflower.b

Neuronowa kompresja obrazów wprowadzonych do systemu „Sunflower.b” odbywa się automatycznie po wybraniu opcji „Kompresja”, której zakładka znajduje się w menu głównym programu (rys. 5). W oknie dialogowym wybiera się przycisk „Wczytaj”, który umożliwi wczytanie pliku przeznaczonego do przetworzenia. Następnym krokiem jest skompresowanie wcześniej wprowadzonych danych przez użytkownika. W tym celu należy wybrać opcję „Kompresja”. Ostatnią czynnością jest zapis przetworzonych danych do pliku wynikowego, którym jest skompresowany obraz.

Wymagania niefunkcjonalne systemu „Sunflower.b”:

- komputer klasy PC,
- system operacyjny *Microsoft Windows 2000/XP*,
- dostęp do Internetu,
- zainstalowane środowisko *.NET Framework*,
- pamięć operacyjna RAM 512MB (lub więcej),
- około 1GB wolnej przestrzeni dyskowej.

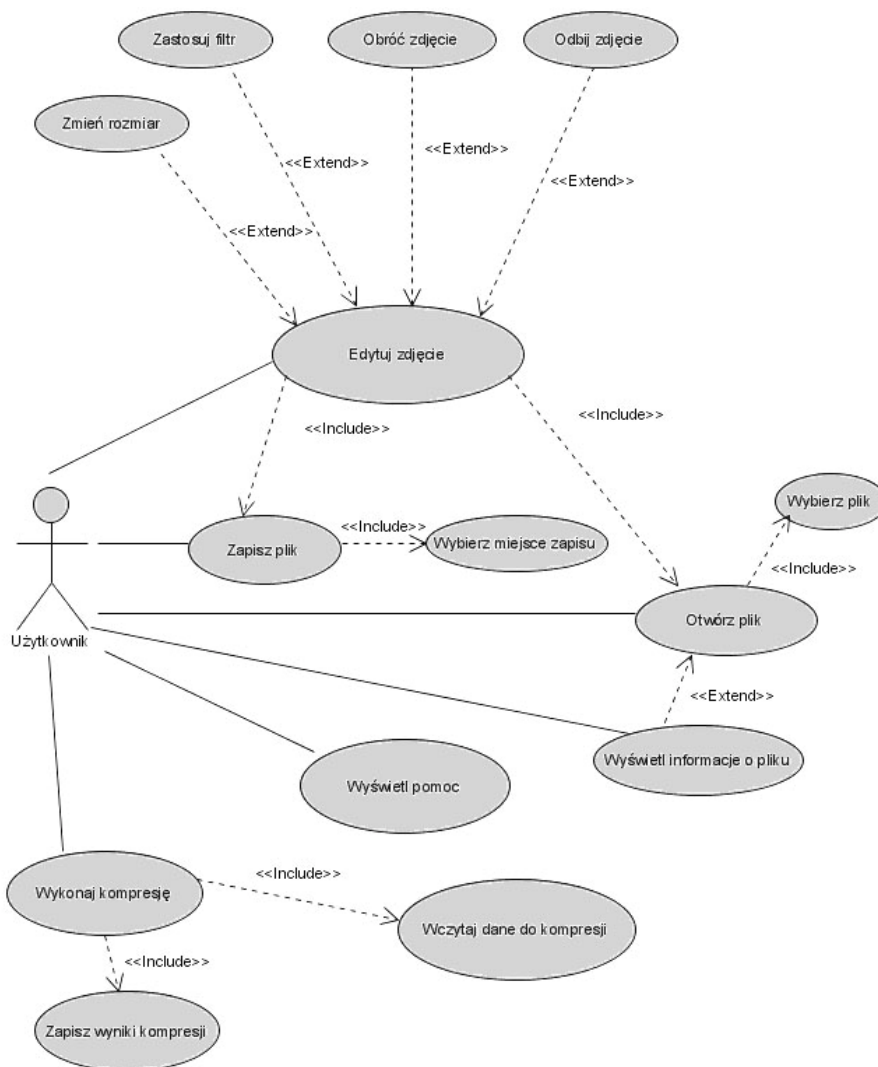


Rys. 5. Główne okno programu „Sunflower.b”
 Fig. 5. Interface of computer system „Sunflower.b”



Rys. 6. Główne okno programu „Sunflower.b”
 Fig. 6. Interface of computer system „Sunflower.b”

Diagram przypadków użycia przedstawiony poniżej obrazuje możliwości pracy w systemie „Sunflower.b”.



Rys. 7. Diagram przypadków użycia
 Fig. 7. Use case diagram

7. Uwagi końcowe

Na podstawie przeprowadzonych rozważań a następnie doświadczeń z użyciem aplikacji „Sunflower.b” w

kontekście wykorzystania autoasocjacyjnych sieci neuronowych typu MLP, jako instrumentu do kompresji danych graficznych można sformułować następujące uwagi końcowe:

- sieć neuronowa typu MLP wytworzona z wykorzystaniem symulatora, zaimplementowanego w pakiecie „STATISTICA v.7.0 – Sieci Neuronowe”, może być wykorzystana do stratnej kompresji danych wstępujących w formie graficznej,
- wydajność modułu do kompresji danych w proponowanym systemie informatycznym „Sunflower.b” (z zaimplementowaną siecią neuronową typu MLP) zależy od struktury przygotowanego zbioru uczącego,
- wytworzony system informatyczny został zaprojektowany z myślą o jego dalszej rozbudowie.

8. Literatura

[1] Boniecki P., Przybył J.: Autoasocjacyjna sieć neuronowa jako narzędzie do nieliniowej kompresji danych. *Journal of*

Research and Applications in Agricultural Engineering, 2006, str. 37-41.

- [2] Kramer M.A.: Nonlinear principal components analysis using autoassociative neural networks. *AIChe Journal* 1991, 37 (2), 233-243.
- [3] Jaskiewicz A.: Inżynieria oprogramowania.. Wydawnictwo Helion S.A., Gliwice 1997.
- [4] Nowakowski K., Boniecki P., Weres J., Mueller W.: Przetwarzanie graficznych danych empirycznych dla potrzeb edukacji sztucznych sieci neuronowych, modelujących wybrane zagadnienia inżynierii rolniczej. *Inżynieria Rolnicza*, Kraków 2007, 2(90), str. 243-248.
- [5] Welch T. A.: A technique for high-performance data compression. *Computer* 1984, nr 6, p. 8-19.
- [6] Ziv J., Lempel A.: Compression of individual sequences via variable-rate encoding. *IEEE Transactions on Information Theory*, 1978, IT-24(5): p.530-536.
- [7] Tadeusiewicz R., Lula P.: *Statistica Neural Networks PL: wprowadzenie do sieci neuronowych*. StatSoft Polska, Kraków 2001.