

NADEŚLANE ARTYKUŁY

PRZEGLĄD PODSTAWOWYCH METOD REPREZENTACJI KSZTAŁTÓW 3D

Dariusz Frejlichowski, Patrycja Nuskiewicz

Zachodniopomorski Uniwersytet Technologiczny
Wydział Informatyki
ul. Żołnierska 49, 71-210 Szczecin
e-mail: dfrejlichowski@wi.zut.edu.pl, pnuskiewicz@wi.zut.edu.pl

Streszczenie: *W artykule przedstawiono przegląd istniejących rozwiązań w dziedzinie reprezentowania kształtów trójwymiarowych, stosunkowo nowym podejściu do opisywania obiektów w przetwarzaniu, rozpoznawaniu i indeksowaniu obrazów. Deskrytory kształtu 3D stają się coraz potrzebniejsze i coraz powszechniej stosowane. Wynika to z rozwoju sprzętu komputerowego, a co za tym idzie możliwości szybkiego przetwarzania skomplikowanych scen trójwymiarowych. Znajduje się przy tym kolejne obszary zastosowań trójwymiarowego opisu obiektów, m.in. w biometrii, systemach CAD i indeksowaniu.*

Słowa kluczowe: *Kształt 3D, indeksowanie obrazów, deskrytory kształtu 3D*

A discussion on selected problems of lossless binary digital images compression

Abstrakt: *In the paper a brief survey on existing approaches to the representation of three-dimensional shapes was presented. It is a new way of describing objects in image processing, recognition and indexing. The 3D shape descriptors become more and more useful and widely used. Rapid development of computer hardware and the possibilities of fast processing of three-dimensional scenes cause it. Many new applications of 3D shape description can be easily found, e.g. in biometrics, CAD systems and image retrieval.*

Keywords: *3D shape, image retrieval, 3D shape descriptors*

1. WPROWADZENIE

W ostatnich latach odnotować można wzrost zainteresowania aplikacjami wykorzystującymi obiekty trójwymiarowe. Sprzyjały temu znaczne ulepszenia narzędzi służących do tworzenia modeli 3D, skanerów trójwymiarowych oraz wzrost przepływu informacji, głównie w Internecie. Ponadto sprzęt komputerowy jest obecnie wystarczająco szybki i tani, aby dane trójwymiarowe mogły zostać przetworzone i wyświetlone w czasie rzeczywistym. Rozwój grafiki trójwymiarowej widoczny jest już nie tylko w ramach informatyki, ale także i w innych dziedzinach, takich jak: medycyna,

biologia molekularna, czy inżynieria mechaniczna. Ze względu na dużą złożoność modeli 3D, jak również ogromną ich ilość w różnego rodzaju bazach, pojawił się problem ich szybkiego i skutecznego wyszukiwania, rozpoznawania i klasyfikowania. Wybrany do tego celu algorytm powinien efektywnie odróżniać cechy charakterystyczne dla różnych klas kształtów trójwymiarowych oraz znajdować modele najbardziej podobne do zadanego na wejściu systemu. Wymagany jest również optymalny czas działania rozwiązania, tak aby wyniki wyszukiwania lub rozpoznawania przedstawiane były w rozsądnie krótkim czasie.

W niniejszym artykule opisane zostały najważniejsze używane dotychczas podejścia do problemu reprezentacji kształtów 3D. Stosowanie deskrytorów tego typu

pozwała na wydobycie najistotniejszych informacji o obiektach, redukując ilość potrzebnych danych (nie musimy używać dzięki temu całych modeli trójwymiarowych), ale także pozwala na uniezależnienie się od różnego rodzaju deformacji porównywanych obiektów.

Kształty trójwymiarowe posiadają cechy, które znacząco odróżniają je od obrazów dwuwymiarowych, w kontekście rozwiązania określonego problemu, np. efektywnego rozpoznawania ([12], [13]). Zazwyczaj, w przeciwieństwie do modeli 2D, obiekty 3D nie są zależne od położenia kamer, źródeł światła lub otaczających obiektów na scenie. W efekcie, obiekty nie zawierają odbić, cieni lub części innych obiektów. Ułatwia to proces wyznaczania miary podobieństwa pomiędzy różnymi modelami. Z drugiej strony, zauważalny jest dużo większy rozmiar opisu obiektów trójwymiarowych, co utrudnia wyszukiwanie odpowiadających sobie własności oraz parametrów modeli. Ponadto, większość obiektów, znajdujących się w dużych bazach danych, zawiera źle zorientowane, niepołączone lub brakujące wielokąty. Poprawienie jakości takich modeli jest trudne w realizacji, gdyż bardzo często wymaga ingerencji człowieka. Dlatego jednym z ważniejszych cech stawianych efektywnemu deskryptorowi kształtu trójwymiarowego jest odporność na błędy, wynikające z uszkodzenia lub braku wielokątów.

Ze względu na wspólne, charakterystyczne cechy prezentowane w artykule podejścia zostały podzielone na cztery grupy: deskryptory geometryczne, strukturalne, symetryczne oraz lokalne. W publikacji pominięto problem porównywania otrzymanych reprezentacji, koncentrując się na samym procesie ich tworzenia.

2. CHARAKTERYSTYKA WYBRANYCH METOD OPISU KSZTAŁTU OBIEKTÓW TRÓJWYMIAROWYCH

2.1. Deskryptory geometryczne

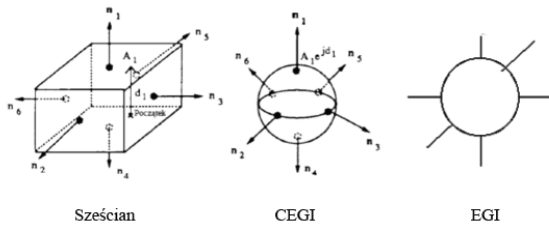
EGI, CEGI

Pierwszym deskryptorem kształtu, który w sposób znaczący wpłynął na późniejsze badania, był opisany w [5], pochodzący z roku 1983, deskryptor EGI (*Extended Gaussian Image*). Istotną wadą tej reprezentacji jest brak odporności na skalowanie i obrót – deskryptor zostaje tak samo przekształcony, jak zmodyfikowany obiekt. Natomiast zaletą tego podejścia jest niezależność od położenia obiektu na scenie, czyli jego przesunięcia (translacji) w przestrzeni. Dzięki tym właściwościom deskryptor EGI znalazł liczne zastosowania, m.in. w systemach rozpoznawania, wyznaczania orientacji obiektów 3D w przestrzeni, podziale map głębokości na mniejsze obiekty. Warto przy tym wspomnieć, że istnieją

algorytmy pozwalające na otrzymanie źródłowego obiektu na bazie jego deskryptora EGI. Cecha odwracalności deskryptora jest rzadko spotykaną właściwością w istniejących rozwiązaniach, dlatego jej spełnienie przez omawiane podejście na pewno należy uznać za istotną zaletę. Budowa deskryptora EGI oparta jest na obrazie gaussowskim, który otrzymujemy poprzez mapowanie powierzchni wektorów normalnych obiektu na sferę jednostkową (sferę gaussowską). Wektory normalne zostają umieszczone na sferze w taki sposób, aby punkt początkowy wektora znajdował się w punkcie środkowym sfery, a punkt końcowy – w punkcie na sferze, odpowiadającym orientacji danej powierzchni (ściany obiektu) w przestrzeni. Deskryptor EGI wzbogaca obraz gaussowski o informacje o rozmiarze powierzchni wszystkich ścian obiektu. Oznacza to, że dla każdego punktu na sferze gaussowskiej ustalona zostaje waga, równa powierzchni ściany obiektu. Wyznaczone wagi przedstawione są jako wektory równoległe do wektorów normalnych (danej powierzchni), o długości równej wartości obliczonej wagi.

Najważniejszą wadą deskryptora EGI jest brak informacji o położeniu ścian obiektu (np. nie można odczytać, które ściany obiektu są ze sobą połączone). Cecha ta powoduje niejednoznaczność – tylko w przypadku wielościanów wypukłych deskryptor jest unikalny. Powstało wiele propozycji rozwiązania powyższego problemu, np. algorytm opisany w [7] i [14], zachowujący informację o położeniu powierzchni poprzez sformułowanie równania ściany obiektu w podwójnej przestrzeni (jednoczesnej reprezentacji orientacji i położenia ściany obiektu).

Kolejny deskryptor, CEGI (*Complex Extended Gaussian Image*), jest rozszerzeniem poprzednio opisanego, przechowującym informację o orientacji, powierzchni ścian oraz pozycji obiektu w przestrzeni. W przypadku deskryptora EGI wielkość wagi, związanej z wektorem normalnym ściany modelu 3D, równa jest powierzchni określonej ściany. Aby powstał deskryptor CEGI należy dodać do każdej wagi część urojona – odległość pomiędzy powierzchnią ściany, a określonym początkiem (zwróconą w kierunku wektora normalnego danej ściany). Oznacza to, że waga (związana z określonym wektorem normalnym) jest liczbą zespoloną, której część rzeczywista równa jest powierzchni ściany, a urojona jest odległością mierzoną od określonego początku do powierzchni ściany obiektu. Na rys.1. przedstawiono graficzne interpretacje deskryptorów CEGI i EGI dla prostego obiektu 3D.



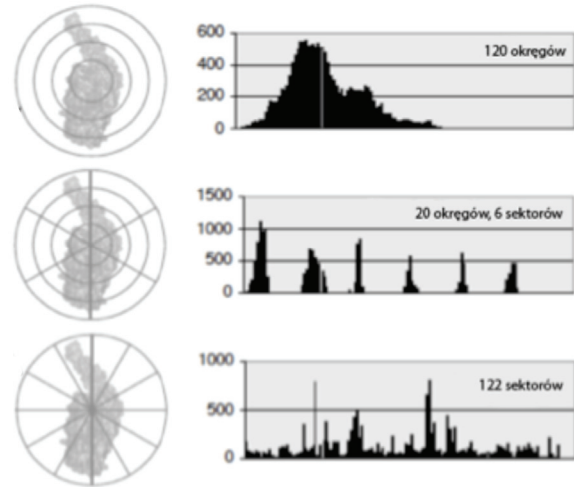
Rysunek. 1 Przykładowe interpretacje deskryptorów EGI i CEGI dla sześcianu, źródło: [18], [19].

W przypadku, gdy obiekt jest wypukły, deskryptor CEGI równy jest deskryptorowi EGI. Wartość wag deskryptora CEGI jest niezależna od położenia obiektu w przestrzeni. Dlatego porównując wagi, możliwe jest rozpoznanie oraz wyznaczenie orientacji obiektów – tak jak w przypadku deskryptora EGI. Dodatkowo, porównując różnice części urojonych wag zespolonych, można wyznaczyć zmianę odległości wzdłuż wektorów normalnych. Omówione właściwości deskryptora CEGI skłaniają ku stwierdzeniu, że jest on lepiej dostosowany do indeksowania większej bazy obiektów niż EGI.

Histogram kształtu

Histogram kształtu trójwymiarowego po raz pierwszy został zaproponowany w [1]. Wyznaczenie deskryptora opiera się tu na odpowiednim podziale przestrzeni, w której znajdują się obiekty. We wspomnianej publikacji opisano trzy techniki:

- Podział wykorzystujący okręgi o wspólnym punkcie środkowym. Promień zewnętrznego okręgu jest zależny od wielkości największego obiektu, znajdującego się w bazie danych. Główną cechą takiego deskryptora jest niezależność od obrotu obiektu 3D względem środka układu współrzędnych.
- Podział wykorzystujący sektory, rozpoczynające się w punkcie centralnym modelu. Budowa deskryptora, opartego na tej metodzie, jest bardziej złożona niż w poprzednim przypadku. W pierwszej fazie realizacji podziału, wykorzystany zostaje wielościan foremny, aby w sposób regularny wyznaczyć punkty na powierzchni sfery otaczającej obiekt. Zdefiniowane punkty, w połączeniu z diagramem Voronoi, automatycznie dzielą przestrzeń na sektory o wspólnym punkcie centralnym. Otrzymany opis jest inwariantem obrotu i skalowania.
- Podział będący połączeniem dwóch poprzednich technik, bardziej złożony, ale zawierający więcej informacji o obiekcie 3D.



Rysunek. 2 Histogramy kształtu, zbudowane techniką wykorzystującą kolejno: okręgi współśrodkowe, sektory oraz złożenie obu, źródło: [1].

Ilustracja graficzna opisanych różnych wersji histogramu kształtu 3D, z uwzględnieniem różnych technik podziału przestrzeni, została przedstawiona na rys.2. Z lewej strony pokazano sposób podziału przestrzeni, z prawej – histogram kształtu.

Rozkład kształtu

Kolejna metoda budowy deskryptorów kształtów trójwymiarowych koncentruje się na obliczeniu rozkładów kształtu na podstawie modelu. Pierwszym etapem jest tu wybór funkcji kształtu. Najważniejszą właściwością funkcji powinna być prostota oraz odporność na przekształcenia, którym mogą zostać poddane obiekty trójwymiarowe. W [12] i [13] przedstawione zostały następujące funkcje kształtu:

- **A3**: Miara kąta pomiędzy trzema losowymi punktami, znajdującymi się na powierzchni obiektu;
- **D1**: Miara odległości pomiędzy wybranym a losowym punktem, znajdującym się na powierzchni obiektu. Wybrany punkt to centroid powierzchni;
- **D2**: Miara odległości pomiędzy dwoma losowymi punktami, znajdującymi się na powierzchni obiektu;
- **D3**: Miara pierwiastka kwadratowego powierzchni trójkąta, którego wierzchołki są losowymi punktami, znajdującymi się na powierzchni obiektu;
- **D4**: Miara pierwiastka trzeciego stopnia objętości czworościanu, zbudowanego z czterech losowych punktów, znajdujących się na powierzchni obiektu.

Po wyborze funkcji kształtu wyznaczany jest rozkład kształtu. Dla danego obiektu trójwymiarowego obliczanych zostaje wiele przykładowych rozkładów kształtu. Na ich podstawie budowany zostaje histogram, przedstawiający ilość wcześniej obliczonych wartości mieszczących się w określonym przedziale. Rozkład

kształtu zostaje zrekonstruowany z wyznaczonego i zapisany w postaci wektora.

2.2. Deskryptory strukturalne

Wielo-rozdzielczościowy graf Reeba (MRG)

Pierwszy deskryptor strukturalny, oparty na teorii grafu Reeba, opisany został w [3]. Przedstawiono tam wielo-rozdzielczościowy graf Reeba (*Multiresolutional Reeb Graph – MRG*).

Graf Reeba jest jedną z podstawowych struktur danych, reprezentujących (kodujących) kształt i właściwości obiektu 3D. Pierwszym etapem jego budowy jest definicja funkcji μ (najczęściej równej funkcji wysokości), według której następuje podział modelu 3D na obszary. Każdy węzeł grafu Reeba reprezentuje składnik danego obszaru połączony krawędziami ze składnikami obszarów sąsiednich.

Graf MRG jest złożony z wielu grafów Reeba, wyznaczonych dla różnej liczby obszarów. Jako pierwszy wyznaczony zostaje graf dla jednego obszaru, składający się z jednego węzła, następnie liczba obszarów zostaje dwukrotnie zwiększona, itd. Wskutek podziału obiektu na różną liczbę obszarów, grafy MRG zawierają więcej informacji i lepiej aproksymują model 3D niż podstawowy graf Reeba. Węzły grafu Reeba o określonym poziomie rozdzielczości (określonej liczbie obszarów) mogą zostać wyznaczone poprzez połączenie węzłów grafu o większym poziomie rozdzielczości. Wykorzystując te właściwości, podobieństwo pomiędzy obiektami może zostać wyznaczone z użyciem strategii „od ogółu do szczegółu” (*coarse-to-fine*) dla różnych poziomów rozdzielczości.

Jednym z najważniejszych elementów budowy grafu Reeba i jednocześnie deskryptora MRG, jest wybór funkcji dzielącej obiekt 3D na obszary. W [3] wykorzystano funkcję, zdefiniowaną jako suma odległości geodezyjnej z punktu v do wszystkich punktów na płaszczyźnie S . Zapewnia ona niezmiennosc przy translacji obiektu oraz przy deformacji jego siatki, na przykład po pojawieniu się niewielkiego szumu i zakłóceń. Aby funkcja $\mu(v)$ zachowała niezmiennosc względem skalowania obiektu, dodano normalizację. Dodatkowo, z powodu braku punktu odniesienia (względem którego dokonywane są obliczenia), funkcja $\mu(v)$ jest stabilna, co oznacza, że zmiany w modelu nie wpływają na zmianę położenia punktu odniesienia, jak w przypadku standardowego obliczania odległości geodezyjnej.

Deskryptor szkieletowy

Deskryptor szkieletowy ([17]) jest reprezentacją „grafu szkieletowego” obiektu 3D. Graf zawiera we wszystkich węzłach informacje topologiczne, odnoszące się do lokalnych właściwości modelu oraz całości grafu. W

przeciwieństwie do opisanego poprzednio grafu MRG deskryptor szkieletowy jest bardzo intuicyjnym deskryptorem kształtu, ponieważ w przejrzysty i zrozumiały dla użytkownika sposób wizualizuje strukturę obiektu 3D. Dzięki tej właściwości możliwe jest wydajniejsze wyszukiwanie obiektów podobnych, także w przypadku, gdy obiekt z zapytania stanowi część większego obiektu w bazie. Dodatkowo mamy możliwość wskazania części modelu 3D, która ma zostać odnaleziona lub nadania wyższej wagi określonym częściom podczas wyszukiwania. Niemniej istotną zaletą deskryptora szkieletowego jest możliwość zrozumiałej wizualizacji wyników wyszukiwania, co pomaga zrozumieć i ocenić stopień podobieństwa porównywanych modeli. Kolejną zaletą deskryptora szkieletowego jest stała topologia grafu dla poruszającego się obiektu, dzięki czemu może on być stosowany dla obiektów w ruchu.

Proces wyszukiwania i porównywania obiektów za pomocą deskryptora szkieletowego rozpoczyna się od wyznaczenia szkieletu. W pierwszym etapie następuje ‘wokselizacja’ wszystkich modeli 3D oraz wyznaczenie płaszczyzny środkowej (*medial surface*). Następnie pomniejszany jest rozmiar uzyskanej płaszczyzny, tak aby mogła ona zostać przedstawiona w postaci grafu.

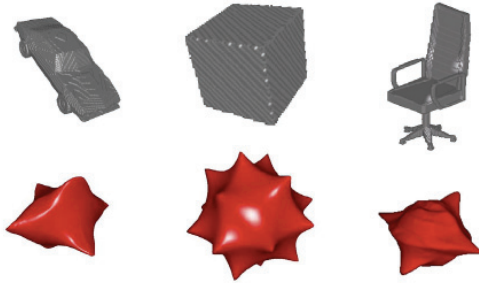
Deskryptor szkieletowy nie jest idealny dla wszystkich obiektów 3D. Trudność stanowią mogą proste modele, które odnajdywane zostają jako mniejsze części bardziej złożonych obiektów. Z drugiej strony, bardzo dużą zaletą grafów szkieletowych jest możliwość wpływania na wyniki rozpoznawania poprzez odpowiednie zdefiniowanie parametrów lokalnego dopasowania oraz złożoności szkieletu. Przy tym, czas obliczeń deskryptora jest proporcjonalny do liczby wokseli w danym obiekcie.

2.3. Deskryptory symetryczne

Deskryptor symetrii osiowej

Pierwsze badania nad symetrią kształtów skupione były na klasyfikacji obiektów ze względu na grupy symetrii. Jest to dobry sposób na zmniejszenie ilości informacji opisującej dany kształt, jeżeli występują w nim osie symetrii. Jednakże o wiele wydajniejsze jest stosowanie tzw. miar osiowych ([8]). Deskryptor symetrii osiowej (*Reflective Symmetry Descriptor – RSD*) jest dwuwymiarową funkcją opisującą symetrię względem każdej płaszczyzny poprowadzonej przez środek ciężkości obiektu (centroid). Ideą deskryptora jest wyznaczenie miary symetrii dla wszystkich płaszczyzn, nawet w przypadku, kiedy nie odpowiadają one idealnej symetrii osiowej modelu. Na rys. 3 znajdują się przykładowe obiekty oraz reprezentujące je RSD. Deskryptory zostały zwizualizowane poprzez skalowanie wektorów jednostkowych na kuli, proporcjonalnie do

miary symetrii płaszczyzny przechodzącej przez środek ciężkości obiektu i normalnej do wektora. Zauważyć można, że RSD zapewnia ciągłą miarę symetrii dla wszystkich płaszczyzn oraz posiada wierzchołki zgodne z płaszczyznami symetrii lub bliskie symetrii modelu. Deskryptor przechowuje bardzo dużo informacji o kształcie obiektu – poza opisem różniących się części obiektu, odzwierciedla ich przestrzenne zależności.



Rysunek. 3 Wizualizacja deskryptorów symetrii osiowej dla trzech obiektów, źródło: [8].

Deskryptor RSD ma szereg zalet. Po pierwsze daje w rezultacie charakterystyczny opis o cechach globalnych. Po drugie, z definicji wykorzystywana jest postać kanoniczna funkcji dwuwymiarowej, co ułatwia wspólną parametryzację w modelach trójwymiarowych. Dodatkowo, w określonych warunkach, reprezentacja jest odporna na szum i inne niewielkie zakłócenia, ale tylko dopóki wszystkie punkty symetrii tworzą wspólną całość.

Dzięki zastosowaniu sferycznej funkcji harmonicznej uzyskujemy opis obiektu niezależny od obrotu w trzech wymiarach ([9]). Kolejną zaletą tego rozwiązania jest możliwość jego umieszczenia w innych istniejących deskryptorach, jako dodatkowy etap w ich konstrukcji. W porównaniu z innymi podobnymi rozwiązaniami normalizującymi lub przekształcającymi zapis obiektu (np. normalizacja względem obrotu) możemy uzyskać lepsze wyniki przy rozpoznawaniu obiektów 3D, przy jednoczesnym zmniejszeniu rozmiaru uzyskiwanego opisu oraz czasu porównywania. Przykładowymi reprezentacjami kształtu, które mogą zostać opisane z użyciem sferycznego zapisu harmonicznego są: EGI ([5]), histogram kształtu ([1]), EXT ([15]), opisujący maksymalny „zasięg” kształtu wzdłuż promieni wychodzących z jednego punktu środkowego, RSD ([8]) oraz EDT.

Rozszerzony deskryptor kształtu

Symetria jest szczególnie przydatną cechą, ponieważ opisuje globalne informacje o kształcie obiektu. Wydobycie niewielkiej ilości informacji o symetrii pozwala na uzyskanie opisu obiektu wystarczającego do skutecznego wyznaczenia podobieństwa pomiędzy

modelami. Ponadto, jeżeli dwa obiekty trójwymiarowe różnią się przynajmniej w jednym punkcie, to w odpowiedzi uzyskamy informację, że są one inne. Wcześniejsze, prostsze techniki, klasyfikujące modele ze względu na rodzaj symetrii, zwracają informację binarną – symetria jest lub nie. W przypadku deskryptorów symetrycznych, uzyskujemy niejako strukturę ciągłą ([10]), gdzie przechowywane są miary symetrii, nawet w przypadku, kiedy obiekt nie jest symetryczny. Dzięki takiej ciągłej klasyfikacji symetrii możemy porównywać i rozpoznawać obiekty bez wykonywania procesu normalizacji, koniecznego przy większości innych deskryptorów. Dodatkowo deskryptor jest niezależny od obrotu obiektu.

Bazując na powyższych cechach symetrii kształtu w [10] zaproponowano rozbudowanie istniejących deskryptorów kształtu o informację o symetrii. Rozszerzono w tym celu sferyczną reprezentację harmoniczną, opisaną w poprzednim rozdziale. Wadą tej reprezentacji jest rozpatrywanie każdego składnika częstotliwości niezależnie i brak informacji charakteryzujących ich położenie względem siebie. Dlatego też dodano do sferycznej reprezentacji harmonicznej o charakterze lokalnym globalną informację o symetrii modeli 3D.

Pierwszy etap procesu budowy rozszerzonego deskryptora symetrycznego polega na przekształceniu sferycznego deskryptora kształtu w sferyczną reprezentację harmoniczną. W tym celu funkcja sferyczna wyrażona zostaje przez składniki częstotliwości oraz przechowana zostaje norma każdego z nich. Następnie obliczone zostają deskryptory symetryczne, dla k-krotnych osi symetrii funkcji sferycznej. Ostatnim etapem jest rozszerzenie sferycznej reprezentacji harmonicznej o informację o symetrii – skalowanie k-krotnych osi symetrii, aby otrzymać najlepszy rozmiar informacji o nieregularnej częstotliwości. Oznacza to, że dla każdego typu symetrii przechowywana jest kopia sferycznej reprezentacji harmonicznej. Dodatkowo, oddzielenie informacji o symetrii od informacji o częstotliwości, umożliwia efektywne porównanie dwóch modeli trójwymiarowych. Procesy porównywania symetrii i częstotliwości mogą być wykonywane niezależnie, a następnie wyniki mogą zostać połączone w jedną miarę podobieństwa.

2.4. Deskryptory lokalne

Opisywane dotychczas deskryptory kształtu trójwymiarowego miały charakter globalny. Powstały w celu ułatwienia procesu rozpoznawania obiektów 3D, jednak nie są wystarczająco skuteczne w sytuacji, gdy charakterystycznymi cechami określonej klasy obiektów są lokalne właściwości kształtu. Założenie podobieństwa obiektów w klasie na poziomie ogólnym (całego obiektu)

może się niekiedy okazać błędne. Dlatego też prowadzone są także prace z lokalnymi deskryptorami kształtu. W [16] uwaga skupiona została na analizie bazy danych obiektów i wyborze lokalnych właściwości, zapewniających najlepsze wyniki przy rozpoznawaniu modeli 3D. W tym celu obiekt reprezentowany był przez wiele lokalnych deskryptorów kształtu (każdy opisywał pewien podobszar kształtu). Podobieństwo pomiędzy dwoma modelami obliczone było na podstawie podobieństwa pomiędzy dwoma lokalnymi deskryptorami.

Inne przykładowe metody wyznaczania lokalnej reprezentacji kształtu bazują na wyborze podzbioru deskryptorów na podstawie istotności ([2]) lub prawdopodobieństwa ([6]). Tak zwane obrazy obrotowe ([6]) analizują obiekty poprzez utworzenie cylindrycznego odwzorowania lokalnych zbiorów punktów na powierzchni. Wadą w tym przypadku jest czas rozpoznawania, który znacznie wzrasta przy zwiększającej się liczbie porównywanych deskryptorów. Wyklucza to zastosowanie tego podejścia w dużych bazach danych. Aby przyspieszyć proces rozpoznawania modeli trójwymiarowych, można zastosować selekcję zbioru lokalnych deskryptorów. Najprostszą metodą jest wybór losowy, jednak w takim przypadku nie jest pewne, że wybrany deskryptor będzie wystarczająco charakterystyczny dla danego obiektu. Jest też bardzo prawdopodobne, że do uzyskania zadowalających wyników rozpoznawania będzie potrzebna większa liczba deskryptorów losowych.

Odrębny problem stanowi określenie, w jaki sposób wybierać podobszary obiektu, by można je było uznać za charakterystyczne dla niego. Pierwsze techniki bazowały na eksperymentach psychologicznych ([4]), według których system widzenia człowieka dzieli złożony kształt na części oraz wybiera określone właściwości, przed rozpoczęciem procesu rozpoznawania. W kolejnej metodzie ([2]) wskazywano części obiektu na podstawie właściwości jego krzywizny. Podobne rozwiązanie zastosowane zostało w [11].

System wyznaczania lokalnych deskryptorów kształtu opisany w [16] został podzielony na dwa etapy. Pierwszym jest trening deskryptorów dla obiektów znajdujących się w bazie, którego wynikiem jest funkcja dystynkcji (funkcja cech charakterystycznych dla obiektu). Na początku, kształty zostają znormalizowane względem rozmiaru, a na płaszczyźnie każdego obiektu wybierane są losowo punkty. W miejscu, w którym znajduje się wylosowany punkt, zbudowany zostaje deskryptor kształtu. Następnie, na podstawie wyników procesu rozpoznawania treningowej bazy modeli, wyznaczone zostaje prawdopodobieństwo dla każdego deskryptora, które określa najbardziej istotne (niepowtarzalne) cechy obiektu, wystarczająco dyskryminujące podczas procesu rozpoznawania.

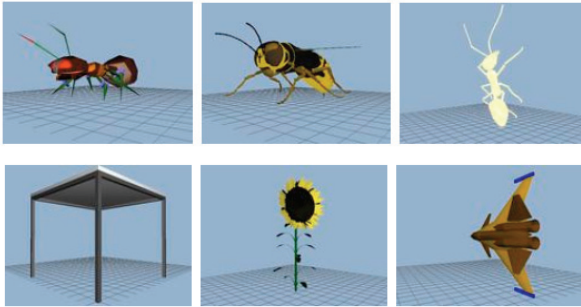
Następnie, zdefiniowana zostaje funkcja dystynkcji – mapująca prawdopodobieństwo deskryptora na przewidywany wynik procesu rozpoznawania. W tym celu dla każdego lokalnego deskryptora kształtu przeprowadzony zostaje proces rozpoznawania. Wyznaczona zostaje również jedna ze standardowych metryk rozpoznawania – *DCG (Discounted Cumulative Gain)*. Jej wartość zawiera się w przedziale $<0,1>$, gdzie lepsze wyniki procesu rozpoznawania znajdują się bliżej wartości 1.

Dla wszystkich deskryptorów obiektów znajdujących się w bazie obliczone zostaje prawdopodobieństwo oraz *DCG* procesu rozpoznawania. Następnie deskryptory zostają uporządkowane według rosnącego prawdopodobieństwa, a w przypadku jednakowego prawdopodobieństwa w drugiej kolejności pod uwagę są brane średnie wartości *DCG*. W rezultacie otrzymujemy histogram średniego wyniku procesu rozpoznawania (*DCG*) w funkcji prawdopodobieństwa (indeksami są osiągnięte wartości prawdopodobieństwa – funkcja dystynkcji).

3. PODSUMOWANIE

W artykule przedstawiony został wstępny przegląd istniejących rozwiązań w dziedzinie reprezentowania kształtu 3D. Opisane metody podzielone zostały na cztery grupy – geometryczne, strukturalne, symetryczne i lokalne. W przypadku każdej z nich przedstawiono podstawowe właściwości najpopularniejszych obecnie rozwiązań.

W wielu badaniach nad deskryptorami kształtu 3D wykorzystywana jest baza danych „Princeton Shape Benchmark” ([20]). Zawiera ona 1814 modeli obiektów trójwymiarowych, znalezionych w Internecie. Dla każdego obiektu dostępny jest plik, przechowujący informację o geometrii obiektu, obraz dwuwymiarowy obiektu w formacie JPEG oraz plik tekstowy, zawierający podstawowe informacje, np. adres strony internetowej, na której się znajduje, ilość wielokątów, maksymalne i minimalne wartości zmiennych na poszczególnych osiach. Baza danych ułatwia w znacznym stopniu rozwój badań nad deskryptorami kształtu 3D, ponieważ zawiera obiekty o bardzo zróżnicowanej liczbie poligonów (od kilkudziesięciu do kilkudziesięciu tysięcy), różnym rozmiarze, położeniu oraz obrotach w przestrzeni trójwymiarowej.



Rysunek. 4 Przykładowe obrazy obiektów trójwymiarowych, znajdujących się w bazie „Princeton Shape Benchmark”, źródło: [20].

Ponieważ wzrasta liczba zastosowań deskryptorów kształtu 3D (w rozpoznawaniu, wyszukiwaniu, systemach CAD i wielu innych) można się spodziewać w najbliższych latach pojawienia się wielu nowych podejść, tak jak to miało miejsce w przypadku kształtów 2D, gdzie w przybliżeniu można obecnie określać liczbę stosowanych na świecie metod na kilka setek.

Literatura

1. Ankerst M., Kastenmüller G., Kriegel H., Seidl T., “3D Shape Histograms for Similarity Search and Classification in Spatial Databases”, 6th International Symposium on Spatial Databases, pp. 207-226, 1999
2. Gal R., Cohen-Or D., "Salient Geometric Features for Partial Shape Matching and Similarity", ACM Transactions on Graphics, vol. 25, pp. 130-150, 2006
3. Hilaga M., Shinagawa Y., Kohmura T., Kunii T. L., „Topology Matching for Fully Automatic Similarity Estimation of 3D Shapes”, 28th Annual Conference on Computer Graphics and Interactive Techniques, pp. 203-212, 2001
4. Hoffman D. D., Singh M., “Salience of Visual Parts”, Cognition, vol. 63, pp. 29-78, 1997
5. Horn B., “Extended Gaussian Images”, Proc. of the IEEE A.I. Memo No. 740, Vol. 72 (12), pp. 1671–1686, 1984
6. Johnson A., “Surface Landmark Selection and Matching in Natural Terrain”, IEEE Conf. on Computer Vision and Pattern Recognition, pp. 413-420, 2000
7. Kang S., Ikeuchi K., “Determining 3-D Object Pose Using the Complex Extended Gaussian Image”, CVPR, pp. 580–585, 1991
8. Kazhdan M., Chazelle B., Dobkin D., Funkhouser T., Rusinkiewicz S., “A Reflective Symmetry Descriptor for 3D Models”, Algorithmica, vol 38, pp. 201-225, 2003
9. Kazhdan M., Funkhouser T., Rusinkiewicz S., “Rotation Invariant Spherical Harmonic Representation of 3D Shape Descriptors”, 2003 Eurographics/ACM SIGGRAPH Symposium on Geometry Processing, pp. 156-164, 2003
10. Kazhdan M., Funkhouser T., Rusinkiewicz S., “Symmetry Descriptors and 3D Shape Matching”, 2004 Eurographics/ACM SIGGRAPH Symposium on Geometry Processing, pp. 115-123, 2004
11. Novotni M., Degener P., Klein R., “Correspondence Generation and Matching of 3D Shape Subparts”, Technical Report CG-2005-2, Universität Bonn, 2005
12. Osada R., Funkhouser T., Chazelle B., Dobkin D., “Matching 3D Models with Shape Distributions”, International Conference on Shape Modeling and Applications, pp. 154-166, 2001
13. Osada R., Funkhouser T., Chazelle B., Dobkin D., “Shape Distributions”, ACM Transactions on Graphics, vol. 21, pp. 807-832, 2002
14. Roach J., Wright J., Ramesh V., “Spherical Dual Images: A 3D Representation Method for Solid Objects that Combines Dual Space and Gaussian Spheres”, CVPR, pp. 236-241, 1986
15. Saupé D., Vranić D. V., “3D Model Retrieval with Spherical Harmonics and Moments”, 23rd DAGM-Symposium on Pattern Recognition, pp. 392-397, 2001
16. Shilane P., Funkhouser T., “Selecting Distinctive 3D Shape Descriptors for Similarity Retrieval”, IEEE International Conference on Shape Modeling and Applications, 2006
17. Sundar H., Silver D., Gagvani N., Dickinson S., “Skeleton Based Shape Matching and Retrieval”, Shape Modeling International, pp. 130-139, 2003
18. “Extended Gaussian Images”, <http://www.cs.cf.ac.uk/Dave/AI2/node197.html>
19. Kazhdan M., „Presentation of Extended Gaussian Images”, <http://www.cs.jhu.edu/~misha/Fall04/EG11.ppt>
20. „Princeton Shape Benchmark”, <http://shape.cs.princeton.edu/benchmark/>