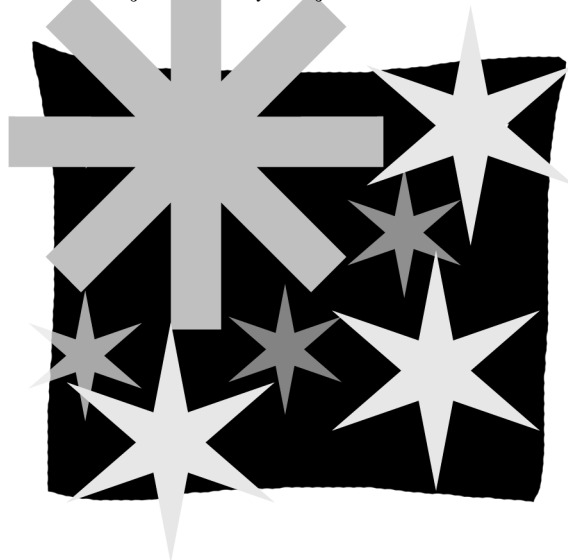


Wyszukiwanie obrazów w bazach danych

Lukasz Matuszczak
Marek Wojciechowski

Zadaniem współczesnych systemów zarządzania bazami danych jest już nie tylko efektywna obsługa danych alfanumerycznych. Upowszechnienie się tanich, pojemnych pamięci masowych, w połączeniu ze wzrostem mocy obliczeniowej serwerów i stacji roboczych, przy jednoczesnym spadku ich cen, spowodowało w ostatnich latach rozwój technologii multimedialnych. Podstawowe wymagania stawiane systemom zarządzania bazami danych w zakresie obsługi danych multimedialnych (obrazów, sekwencji audio, filmów wideo) można ująć jednym stwierdzeniem: systemy te powinny oferować dla danych multimedialnych te same mechanizmy, które są dostępne w kontekście danych alfanumerycznych. Wprowadzenie typu danych BLOB (ang. binary large object) umożliwiło składowanie danych multimedialnych w bazie danych, automatycznie udostępniając takie mechanizmy jak: przetwarzanie transakcyjne, kontrola dostępu do danych, odtwarzanie po awarii. Sporym wyzwaniem dla twórców systemów zarządzania bazami danych jest jednak w dalszym ciągu obsługa zapytań do multimedialnych baz danych.

Kryteria selekcji pojawiające się w zapytaniach do multimedialnych baz danych można podzielić na dwie klasy: kryteria dotyczące metadanych o obiektach i dotyczące zawartości. Metadane o obiekcie multimedialnym to alfanumeryczne dane o jego pochodzeniu (np. nazwa źródłowego pliku) i formacie (format pliku, metoda kompresji, wymiary obrazu, częstotliwość próbkowania dla sekwencji audio, liczba klatek na sekundę dla sekwencji wideo, itd.). Obsługa zapytań o metadane o obiektach realizowana jest w oparciu o tradycyjne techniki przetwarzania zapytań dla danych alfanumerycznych. Zapytania odwołujące się do zawartości obiektów multimedialnych wymagają technik automatycznej ekstrakcji ich właściwości i/lub zaawansowanych modeli pozwalających na składowanie informacji o semantycznej zawartości obiektów.



Celem niniejszego artykułu jest przybliżenie czytelnikom zagadnień związanych z przetwarzaniem zapytań w multimedialnych bazach danych na przykładzie jednego z rodzajów multimediów: obrazów. Obrazy są obecnie jedynym typem danych multimedialnych, dla których Oracle wspiera wyszukiwanie ze względu na zawartość. Artykuł stanowi próbę oceny użyteczności, kompletności i dojrzałości mechanizmów wyszukiwania obrazów w Oracle9i na tle opracowanych przez środowiska naukowe systemów prototypowych i zatwierdzonych w ostatnich latach standardów.

Rozwiązania prototypowe

Problem przetwarzania danych multimedialnych w bazach danych stał się prawdziwym wyzwaniem dla środowisk naukowych, co zaowocowało niezliczoną wręcz liczbą systemów i języków eksperymentalnych. Szczególnie wiele rozwiązań zostało zaproponowanych dla problemu przeszukiwania dużych kolekcji obrazów. Dużą uwagę poświęcono sposobom reprezentacji semantycznej zawartości obrazów i konstrukcji języków zapytań, które często odchodziły od postaci ściśle tekstowej. Miało to nie tylko ułatwić korzystanie z nich użytkownikom, ale umożliwić również tworzenie zapytań dotychczas bardzo trudnych do skonstruowania. Przykładem mogą tu być języki, w których użytkownik rysuje na ekranie pewien obraz, a następnie próbuje odnaleźć w bazie danych obraz najbardziej podobny do narysowanego. Poniżej przedstawiono kilka przykładów systemów prototypowych, które pokazują różnorodne podejścia do kwestii wyszukiwania obrazów.

QBIC - IBM's Query By Image Content

System QBIC [3] umożliwia przeszukiwanie dużych kolekcji obrazów w oparciu o ich własności wizualne. Uwzględniane właściwości to procentowy udział kolorów w obrazie, kształty i tekstury oraz lokalizacja kolorów, kształtów i tekstur na obrazie. QBIC umożliwia na przykład wyszukanie obrazów, które prezentują czerwony okrągły obiekt, znajdujący się w górnym lewym rogu, umieszczony na zielonym tle. Zapytania w systemie QBIC mają charakter wizualny. Użytkownik specyfikuje wymagane kolory, tekstury, kształty i ich lokalizację za pomocą graficznego interfejsu. Interfejs graficzny udostępnia paletę kolorów i suwaki do definiowania udziału poszczególnych kolorów w obrazie oraz narzędzie do szkicowania obrazu wzorcowego w oparciu o predefiniowane kształty geometryczne i paletę tekstur. Funkcjonalność wyszukiwarki QBIC jest obecnie dostępna w komercyjnym systemie zarządzania bazą danych IBM DB2 w ramach rozszerzenia oferowanego pod nazwą DB2 Image Extender [2].

SCORE (a System for Content based REtrieval of pictures)

SCORE [1] proponuje specjalny model reprezentacji zawartości obrazów. Każdy z obrazów w bazie danych jest opisany przez zmodyfikowany diagram encji-związków. Encje nie oznaczają tu jednak typów, ale konkretne obiekty. Podobnie symbol związku dotyczy jednego konkretnego powiązania, a nie zbioru powiązań. SCORE wyróżnia dwa typy związków. *Akcje* opisują pewne sytuacje obecne na obrazku (np. pies *goni* kota). *Relacje przestrzenne* określają względne pozycje obiektów na obrazie (*na lewo, pod, z przodu*).

Tworzenie zapytania polega na graficznym wyborze kilku obiektów z palety ikon. Następnie użytkownik określa dodatkowe parametry i atrybuty obiektów (kolor, rozmiar, liczba) oraz definiuje żądane związki między obiektami. Wykonanie zapytania uwzględnia przybliżone dopasowanie wartości atrybutów oraz akcji (np. sosna jest drzewem) i reguły dedukcji dla związków przestrzennych (np. przechodność relacji *na lewo*).

SMDS (Structured Multimedia Database System)

SMDS [8] jest jedną z pierwszych prób stworzenia podstaw teoretycznych dla technologii multimedialnych. Formalnie zdefiniowano instancję medium, która reprezentuje jeden konkretny typ medium, np. dane audio, wideo, obrazy, dokumenty. Instancja medium zawiera w sobie poszczególne egzemplarze danego typu (np. ścieżka audio) oraz cechy opisujące zawartość tych egzemplarzy. Określono również formalnie strukturę bazy danych. Struktura taka oprócz instancji mediów obejmuje również elementy pozwalające na osłabianie treści zapytań, np. hierarchie generalizacji cech (mustang jest przykładem ford), lub dopuszczalne substytuty dla wartości atrybutów (kolor czerwony można zastąpić pomarańczowym).

Zaproponowane definicje są na tyle ścisłe, że pozwalają stworzyć język zapytań wykorzystujący programowanie w logice (podobny do języka PROLOG). Ponieważ formułowanie zapytań za pomocą programowania w logice jest dość trudne, zaproponowano również język oparty na składni SQL, który na niższym poziomie wykorzystuje odpowiednie formuły logiczne. Poniższe zapytanie w języku SMDS-SQL znajduje obrazy, na których widoczny jest biały ford.

```
SELECT M
FROM smds source1 M
WHERE FindType(M)=image
AND FindObjWithFeature(ford)
AND Color(ford, white, S)
```

CSQL (Cognition and Semantics-based Query Language)

Język CSQL [7] jest częścią systemu zarządzania bazą danych SEMCOG, który służy do przechowywania obrazów statycznych. System wprowadza hierarchiczną strukturę do modelowania obrazów, która wspiera zarówno zapytania na poziomie całych obrazów, jak i obiektów składowych. Dodatkowo uwzględniono dwojaki charakter przechowywanych obrazów, ich cechy *wizualne* i *semantyczne*. Uwzględnienie cech wizualnych pozwala np. na tworzenie zapytań o podobieństwo dwóch obrazów (w kształtach, kolorach, rozmiarach), natomiast cechy semantyczne

umożliwiają wyszukiwanie obrazów na podstawie opisu znaczenia obrazu, który jest definiowany ręcznie przez użytkownika lub półautomatycznie przez algorytmy przetwarzania obrazów.

Model zakłada, że obraz jest obiektem złożonym z wielu obiektów składowych, które mają pewne znaczenie i cechy wizualne (np. człowiek, samochód). Struktura każdego obiektu obejmuje więc jego obraz (zbiór pikseli), cechy semantyczne i relacje przestrzenne. Obiekty mogą również zawierać kolejne podobiekty i relacje przestrzenne między nimi. W zapytaniach można specyfikować kryteria selekcji odwołujące się do semantyki obrazów, wizualnego podobieństwa obrazów oraz relacji przestrzennych, określonych w modelu obrazu. Poniższe, przykładowe zapytanie wyszukuje wszystkie obrazy, na których widoczna jest osoba na prawo od obiektu podobnego do zadane-go obrazu.

```
SELECT image P
WHERE P contains X
AND P contains Y
AND X is_a człowiek
```



```
AND Y i_like
AND X to_the_right_of Y
```

MOQL (Multimedia Object Query Language)

Język MOQL [6] jest rozszerzeniem języka OQL opracowanego dla obiektowych baz danych przez ODMG. W tym wypadku celem twórców była reprezentacja dowolnych danych multimedialnych, a nie tylko obrazów. Sposób przechowywania nie został konkretnie określony, natomiast założono, że oprócz samych mediów, dostępne będą również informacje semantyczne o interesujących obiektach wchodzących w ich skład.

Rozszerzenia języka OQL dotyczą przede wszystkim nowych wyrażeń, jakie można stosować w ramach klauzuli *WHERE*. Są to m.in.: predykaty i funkcje przestrzenne – np. *intersect*, predykaty i funkcje temporalne – np. *overlap*, predykat *CONTAINS*. Składnia zapytań została również poszerzona o specjalną klauzulę *PRESENT*, która daje szerokie możliwości definiowania sposobu prezentacji wyników.

Poniższe zapytanie wyszukuje pary (obraz, wideo), gdzie klip wideo obejmuje wszystkie samochody widoczne na obrazku. Obraz i wideo są prezentowane w oknach o określonej pozycji i rozmiarze. Pokaz obrazu trwa 20 sekund i rozpoczyna się 10 sekund przed początkiem klipu wideo, który jest odtwarzany przez 30 minut.

```
SELECT m,v
FROM Images m, Videos v
WHERE FOR ALL c IN (SELECT r FROM Cars r WHERE m CONTAINS r) v CONTAINS c
PRESENT atWindow(m, (0,0), (300, 400))
AND atWindow(v, (301, 401), (500, 700))
AND play(v, 10, normal, 30*60) parStart display(m, 0, 20)
```

Obowiązujące standardy

Proponowane w rozwiązaniach prototypowych rozszerzenia języka SQL o konstrukcje wspierające obsługę obrazów i innych rodzajów multimedii nie zyskały uznania organizacji odpowiedzialnych za standard SQL. Zamiast

rozbudowy SQL o kolejne, trudne w implementacji i potencjalnie niezgodne ze sobą rozszerzenia dla danych multimedialnych, tekstowych i przestrzennych, zwyciężyła koncepcja opracowania nowego standardu, obejmującego specyfikację bibliotek opartych o typy obiektowe SQL99 [10], przeznaczonych do obsługi poszczególnych specjalistycznych rodzajów danych i aplikacji. Nowy standard - SQL/MM: SQL Multimedia and Application Packages - składa się z wielu części [9]. Obecnie tylko jedna z nich dotyczy danych multimedialnych - *Part 5: Still Image* [4], poświęcona obrazom. Pozostałe części dotyczą tekstowych i przestrzennych baz danych oraz eksploracji danych (ang. data mining). SQL/MM w obecnej postaci nie zawiera specyfikacji dla danych audio i wideo.

Specyfikacja SQL/MM: *Still Image* definiuje strukturalne typy SQL umożliwiające składowanie obrazów w bazie danych, ich odczyt i różnego rodzaju modyfikacje oraz wyszukiwanie obrazów spełniających kryteria selekcji odwołujące się do atrybutów wizualnych obrazu. W standardzie SQL/MM obrazy są reprezentowane za pomocą typu *SI_StillImage*. *SI_StillImage* składa się z dwuwymiarowego obrazu w postaci kolekcji pikseli. Standard zakłada, że obrazy mogą być składowane w różnych formatach (JPEG, GIF, TIFF, itd.), zależnie od możliwości konkretnej implementacji. Oprócz samego obrazu (składowanego w postaci dużego obiektu binarnego - BLOB), *SI_StillImage* zawiera również takie informacje o obrazie jak format i wymiary w pikselach. Metody typu *SI_StillImage* pozwalają m.in. na skalowanie i obroty obrazu, obcinanie obrazu, konwersje formatów oraz generację miniaturki obrazu w mniejszej rozdzielczości (ang. thumbnail).

Oprócz podstawowego typu *SI_StillImage*, SQL/MM w zakresie obsługi obrazów definiuje również kilka typów służących do reprezentacji różnych właściwości obrazu. Typ *SI_AverageColor* reprezentuje średni kolor obrazu, *SI_ColorHistogram* dostarcza informacji o udziale kolorów w obrazie, *SI_PositionalColor* reprezentuje lokalizację poszczególnych kolorów na obrazie, a *SI_Texture* służy do zapamiętania informacji o tzw. teksturze obrazu. Dla każdego z wymienionych typów istnieje metoda typu *SI_StillImage* zwracająca daną właściwość wskazanego obrazu. Ponadto, typ *SI_StillImage* i każdy z typów reprezentujących właściwości obrazu posiada metodę *SI_Score*, służącą do wyznaczenia odległości między dwoma obrazami w kontekście danej właściwości. Odległość przyjmuje wartości większe lub równe 0, większa odległość oznacza większą różnicę.

Poniższe przykładowe zapytanie wykorzystujące typy SQL/MM zwraca wzory tkanin (składowane w kolumnie *Wzor* tabeli *WZORY*) podobne w sensie tekstury do zadanego wzorca. Wykorzystana w klauzuli *WHERE* metoda *SI_FindTexture* wyznacza teksturę nowego wzoru w postaci wartości typu *SI_Texture*. Na rzecz uzyskanej w ten sposób reprezentacji tekstury podanego obrazu, wołana jest następnie metoda *SI_Score*, która wyznacza odległość pod względem tekstury między podanym obrazem a obrazami składowanymi w bazie danych w kolumnie typu *SI_StillImage*.

```
SELECT*
FROM WZORY
WHERE SI_FindTexture(nowyWzor).
SI_Score(Wzor) < 1.2
```

Standard SQL/MM nie porusza kwestii składowania i przeszukiwania semantycznych opisów zawartości obiektów multimedialnych. Wydaje się jednak, że rolę języka opisu semantycznej zawartości multimedii (w tym obrazów) w bazach danych powinien spełniać MPEG-7 [5]. MPEG-7 jest ogólnym standardem języka opisu zawartości obiektów multimedialnych, niezwiązanym z technologiami bazodanowymi. MPEG-7 (oficjalna nazwa standardu to Multimedia Content Description Interface) został opracowany przez grupę MPEG (Moving Picture Experts Group) i jest obowiązującym standardem ISO/IEC.

Specyfikacja MPEG-7 ogranicza się do formatu i zawartości opisów, nie obejmuje natomiast zagadnień związanych z ekstrakcją właściwości i przeszukiwaniem opisów. Opisy MPEG-7 uwzględniają metadane o obiekcie multimedialnym (producent, format zapisu), semantyczne metadane o treści (obiekty, zdarzenia, postacie, role postaci) oraz sygnałowe metadane o treści (kolor, tekstura) [12]. Zawartość opisu MPEG-7 częściowo się więc nakłada na zawartość typów obiektowych SQL/MM, oferując dodatkowo możliwość zapamiętania danych o semantycznej zawartości obiektu. Deskryptory MPEG-7 zapisywane są w formacie XML, a jako język definiowania składni deskryptorów służy XMLSchema, co potencjalnie umożliwia przetwarzanie opisów MPEG-7 w systemach zarządzania bazami danych obsługujących dane XML (np. Oracle9i).

Składowanie i wyszukiwanie obrazów w Oracle9i

Właściwością (ang. feature) serwera Oracle9i umożliwiającą składowanie, odczyt i przetwarzanie danych multimedialnych jest *interMedia* [11]. Oracle *interMedia* obsługuje dane multimedialne składowane jako BLOB i BFILE lub dostępne przez wskazany adres URL. *interMedia* dostarcza typy obiektowe dla poszczególnych rodzajów multimedii (dla obrazów - *ORDImage*), wiążące obiekt multimedialny z metadanymi o jego pochodzeniu i formacie, a także udostępniające metody do operowania na obiektach i ich metadanych. Typ *ORDImage* umożliwia:

- eksport/import danych między bazą danych i lokalizacjami zewnętrznymi;
- automatyczną ekstrakcję i pielęgnację metadanych o formacie (dla obrazów w popularnych formatach graficznych);
- przetwarzanie obrazów: konwersję formatu, kopiowanie, skalowanie i obcinanie.

W oparciu o tradycyjne mechanizmy obiektowo-relacyjne, *interMedia* wspiera obsługę zapytań z kryteriami selekcji odwołującymi się do metadanych, składowanych w atrybutach typu *ORDImage*. Poza zapytaniami o metadane, dla obrazów *interMedia* wspiera również wyszukiwanie ze względu na zawartość (ang. content-based image retrieval).

Wyszukiwanie obrazów ze względu na zawartość - Zasada działania

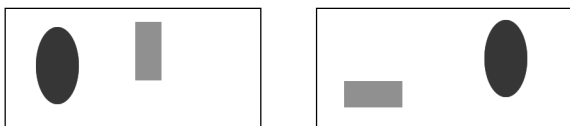
Oracle *interMedia* uwzględnia przy wyszukiwaniu obrazów ze względu na zawartość następujące właściwości:

- kolor - udział poszczególnych kolorów;
- tekstura - „niskopoziomowe” wzory na obrazie, szczególnie te powtarzające się;
- kształt - kształty występujące na obrazie, rozumiane jako regiony o jednym kolorze;

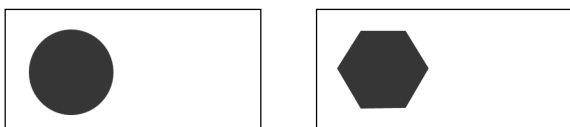
- lokalizacja – reprezentująca położenie kolorów, tekstur i kształtów na obrazie (ma sens tylko w połączeniu z co najmniej jedną z pozostałych właściwości).

Poniższe rysunki przedstawiają przykłady par podobnych obrazów.

- obrazy bardzo podobne pod względem kształtów i udziału kolorów:



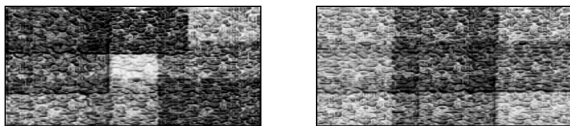
- obrazy podobne pod względem koloru i lokalizacji kolorów:



- obrazy bardzo podobne pod względem kształtów i ich lokalizacji:



- obrazy o podobnej teksturze:



Bazę dla wyszukiwania obrazów podobnych do danego stanowi ekstrakcja własności wizualnych obrazów. Właściwości wizualne obrazów są w interMedia reprezentowane w postaci sygnatur (obiektów typu *ORDImageSignature*). Wszelkie operacje związane z oceną podobieństwa obrazów są wykonywane na sygnaturach, a nie bezpośrednio na obrazach. Podstawą dla wyznaczenia sygnatury jest segmentacja obrazu, czyli podział na regiony o jednokolorowym kolorze. Segmenty opisane są: kolorem, kształtem i teksturą. Sygnatura odzwierciedla własności segmentów i globalne własności obrazu w zakresie koloru, kształtu i tekstury. Sygnatury muszą być jawnie wyznaczone (za pomocą metody *generateSignature* typu *ORDImageSignature*). Następnie można je składować w tabelach tak samo jak obrazy (*ORDImage*) i inne dane.

Przy wyszukiwaniu ze względu na zawartość, użytkownik przypisuje wagi poszczególnym właściwościom obrazów. Waga jest liczbą z przedziału od 0.0 do 1.0 (waga 0.0 dla danej właściwości oznacza, że ma ona nie być brana pod uwagę). Podane przez użytkownika wagi są automatycznie normalizowane, tak by ich suma wynosiła 1.0. Miarą podobieństwa jest ważona suma odległości między obrazami w kontekście poszczególnych właściwości. Wynik jest liczbą zmiennoprzecinkową z zakresu od 0 (dla obrazów identycznych) do 100 (dla obrazów całkowicie różnych).

Do przeprowadzania testów podobieństwa służy operator *IMGSimilar*, który oprócz dwóch sygnatur otrzymuje na

wejście łańcuch znaków z przypisaniem wag do właściwości wizualnych i próg maksymalnej dopuszczalnej odległości (próg ten musi być dobrany empirycznie). Operator *IMGSimilar* może być wykorzystywany w połączeniu z podrzędnym operatorem *IMGScore*, który zwraca wartość funkcji odległości wyznaczoną podczas testu podobieństwa.

Dla poprawy efektywności operacji wyszukiwania obrazów podobnych do danego, Oracle interMedia umożliwia tworzenie specjalnego typu indeksu (*ORDSYS.ORDIMAGEINDEX*) na kolumnach typu *ORDImageSignature*.

Wyszukiwanie obrazów ze względu na zawartość – Przykład

W celu zilustrowania sposobu składowania i wyszukiwania obrazów ze względu na zawartość w Oracle interMedia, rozważmy kolekcję obrazów zawierających wzory tkanin. Poniższe polecenie SQL tworzy tabelę *WZORY* do składowania takiej kolekcji obrazów. Nazwy tkanin będą składowane w kolumnie *Nazwa*, a obrazy reprezentujące wzory - w kolumnie *Obraz*. Ponieważ planowane jest przeszukiwanie kolekcji ze względu na właściwości wizualne wzorów, dodatkowo w tworzonej tabeli zawarta jest kolumna do pamiętania sygnatur obrazów (*Sygn*).

```
CREATE TABLE wzory
(nazwa VARCHAR2(40),
obraz ORDSYS.ORDImage,
sygn ORDSYS.ORDImageSignature);
```

Poniższa instrukcja *INSERT* dodaje do tabeli *WZORY* nowy wiersz (obraz składowany jest jako obiekt *BFILE*):

```
INSERT INTO wzory
VALUES («kratka»,
ORDSYS.ORDImage.init(«file», 'TKANINY', 'kratka.gif'),
ORDSYS.ORDImageSignature.init());
```

Poniższy blok kodu w PL/SQL generuje sygnaturę dla jednego z obrazów składowanych w tabeli *WZORY* i zapisuje ją w bazie danych:

```
DECLARE
ob ORDSYS.ORDImage;
sg ORDSYS.ORDImageSignature;
BEGIN
SELECT obraz, sygn INTO ob, sg FROM wzory
WHERE nazwa = «kratka» FOR UPDATE;
sg.generateSignature(ob);
UPDATE wzory SET sygn = sg WHERE nazwa = 'kratka';
END;
```

Poniższe zapytanie wyszukuje w tabeli *WZORY* wzory podobne do wzoru o nazwie *'kratka'* składowanego również w tej tabeli. Zapytanie wykorzystuje operator *IMGSimilar* odwołujący się do sygnatur składowanych w kolumnie *Sygn*. W teście podobieństwa mają być brane pod uwagę tekstura i kolor obrazów, z większą wagą przypisaną kolorowi. Własnościom kształtu i lokalizacji nie przypisano w zapytaniu wag, w związku z tym mają one domyślną wagę 0 i nie są brane pod uwagę. Wartością progową miary odległości w poniższym zapytaniu jest 5.0.

```

SELECT w.nazwa
FROM wzory w
WHERE ORDSYS.IMGSimilar(
    w.sygn,
    (SELECT v.sygn FROM wzory v WHERE v.nazwa = 'kratka'),
    'texture = „0,3” color = „0,7”, 5.0) = 1

```

W przypadku gdy użytkownika interesuje nie tylko binarny wynik testu podobieństwa, ale również stopień podobieństwa obrazów, należy wykorzystać w zapytaniu pomocniczy operator *IMGSScore*, zwracający wartość miary odległości wyznaczoną w teście podobieństwa przeprowadzonym za pomocą operatora *IMGSimilar*.

```

SELECT w.nazwa, ORDSYS.IMGSScore(303)
FROM wzory w
WHERE ORDSYS.IMGSimilar(
    w.sygn,
    (SELECT v.sygn FROM wzory v WHERE v.nazwa = 'kratka'),
    'texture = „0,3” color = „0,7”, 5.0, 303) = 1

```

„Tajemniczy” parametr operatora *IMGSScore* (w powyższym przykładzie: 303) jest numerem referencyjnym wiążącym w zapytaniu wystąpienie operatora *IMGSScore* z nadrzędnym w stosunku do niego operatorem *IMGSimilar*. Powiązanie to umożliwia wykorzystanie w jednym zapytaniu kilku par operatorów *IMGSimilar* i *IMGSScore*, np. przy wyszukiwaniu obrazów podobnych do kilku zadanych wzorców.

Wyszukiwanie obrazów ze względu na zawartość – Uwagi

Wykorzystując Oracle interMedia do wyszukiwania obrazów należy pamiętać o tym, że komputer ciągle nie dorównuje człowiekowi w zakresie porównywania obrazów. Umysł ludzki znacznie lepiej radzi sobie ze skalowaniem, obrotami i przesłanianiem obiektów. Należy być świadomym tego, na jakich zasadach interMedia ocenia podobieństwo obrazów. Na potrzeby ekstrakcji własności wizualnych obrazy są tymczasowo skalowane do jednego rozmiaru, przez co np. fragment wycięty z obrazu może być uznany za niepodobny do całego obrazu. Dlatego też interMedia radzi sobie dobrze w sytuacjach gdy szukany obiekt zajmuje cały obraz. Jeśli na obrazie jest wiele kształtów, interMedia poradzi sobie lepiej gdy kształtów tych jest niewiele i gdy są to proste kształty, najlepiej w kontrastujących kolorach. Doświadczenia pokazują, że w przypadku złożonych obrazów i fotografii skuteczniejsze od kryterium kształtu są kryteria odwołujące się do rozkładu i lokalizacji kolorów.



Podsumowanie

Wsparcie dla wyszukiwania obrazów w bazach danych wymaga technik automatycznej ekstrakcji własności wizualnych i modeli reprezentacji semantycznych własności obrazów i obiektów na nich przedstawionych. Obecność takich standardów jak SQL/MM i MPEG-7 daje nadzieję na dalszy rozwój narzędzi do wyszukiwania obrazów w bazach danych. Obecna wersja systemu zarządzania bazą danych Oracle – Oracle 9.2 implementuje funkcjonalność SQL/MM (przy składni odbiegającej od standardu) umożliwiając składowanie, przetwarzanie i wyszukiwanie obrazów ze względu na ich własności wizualne. Oracle 9.2 nie oferuje natomiast żadnych modeli do składowania i przeszukiwania metadanych o semantyce obiektów multimedialnych. Możliwość wykorzystania do tego celu opisów w standardzie MPEG-7 sprowadza się do ogólnych mechanizmów przetwarzania i przeszukiwania danych XML dostępnych w Oracle 9.2.

Literatura

- [1] Aslandogan Y.A., Yu C.T.C., Liu C., and Nair K.R.: Design, Implementation and Evaluation of SCORE. Proceedings of the 11th International Conference on Data Engineering, 1995.
- [2] DB2 Universal Database, Image, Audio, and Video Extenders Administration and Programming, Version 8. IBM, 2003.
- [3] Flickner M., Sawhney H., Niblack W., Ashley J., Huang Q., Dom B. et al: Query by image and video content: the qbic system. IEEE Computer, 28(9), 1995.
- [4] ISO/IEC 13249-5:2001, Information Technology – Database Languages – SQL Multimedia and Application Packages – Part 5: Still Image. ISO, 2001.
- [5] ISO/IEC 15938, Information Technology – Multimedia content description interface. ISO, 2002-2003.
- [6] Li J.Z., Özsu M.T., Szafron D., Oria V.: MOQL: A multimedia object query language. Proceedings of the 3rd International Workshop on Multimedia Information Systems, 1997.
- [7] Li W.-S., Candan K.S.: SEMCOG: A Hybrid Object-based Image Database System and Its Modeling, Language, and Query Processing. Proceedings of the 14th International Conference on Data Engineering, 1998.
- [8] Marcus S., Subrahmanian V.S.: Foundations of multimedia database systems. Journal of ACM, 43(3), 1996.
- [9] Melton J., Eisenberg A.: SQL Multimedia and Application Packages (SQL/MM). SIGMOD Record 30(4), 2001.
- [10] Melton J., Eisenberg A.: SQL:1999, formerly known as SQL3. SIGMOD Record 28(1), 1999.
- [11] Oracle9i interMedia User's Guide and Reference, Release 2 (9.2). Oracle, 2002.
- [12] Skarbek W.: MPEG-7. Materiały IX Konferencji PLOUG, 2003.