

Oracle interMedia na tle standardu SQL/MM i prototypowych systemów multimedialnych baz danych

Marek Wojciechowski, Łukasz Matuszczak

Politechnika Poznańska

Przetwarzanie danych multimedialnych od lat było przedmiotem wielu projektów badawczych i tematem prac naukowych. Obecnie jesteśmy świadkami kształtowania się standardu SQL/MM, który w dużej części poświęcony jest obsłudze danych multimedialnych z poziomu języka SQL. Obsługą obrazów, dźwięku, sekwencji wideo i heterogenicznych danych multimedialnych w ramach systemu zarządzania bazą danych Oracle9i zajmuje się Oracle interMedia. Celem niniejszego artykułu jest ocena możliwości oferowanych przez Oracle interMedia w świetle specyfikacji SQL/MM i funkcjonalności prototypów badawczych.

Informacja o prowadzącym:

dr inż. Marek Wojciechowski – pracownik naukowo-dydaktyczny Instytutu Informatyki Politechniki Poznańskiej. Specjalizuje się w dziedzinach systemów baz danych, eksploracji danych oraz technologii internetowych.

1. Wprowadzenie

Zadaniem współczesnych systemów zarządzania bazami danych jest już nie tylko efektywna obsługa danych o prostej strukturze (danych relacyjnych), ale również danych o strukturze złożonej (danych obiektowych), danych nie w pełni strukturalnych (tzw. semi-strukturalnych) oraz multimedialnych. Dane multimedialne definiuje się jako dane zakodowane w jednym lub wielu mediach, spośród których co najmniej jedno medium nie jest alfanumeryczne. W praktyce, mianem danych multimedialnych określa się obecnie zakodowane w postaci zrozumiałej dla komputera obrazy, sekwencje audio, filmy wideo (sekwencje wideo, zazwyczaj uzupełnione o sekwencję audio i opcjonalnie napisy w postaci tekstowej) oraz heterogeniczne dokumenty multimedialne (np. prezentacje). W porównaniu z innymi rodzajami danych, dane multimedialne stanowią szczególnie trudne wyzwanie dla projektantów systemów informatycznych ze względu na duży rozmiar i brak struktury (postać binarna).

System zarządzania bazą danych Oracle9i wspiera wszystkie wspomniane wyżej rodzaje danych. Oracle9i jest określany jako system obiektowo-relacyjny, czyli oparty o model relacyjny rozszerzony o elementy modelu obiektowego, dostępne przede wszystkim poprzez mechanizm typów strukturalnych (zgodnych ze standardem SQL99 [13]). Do przetwarzania danych semi-strukturalnych (mających postać dokumentów XML) w Oracle9i służy XML DB, z kolei za funkcjonalność dotyczącą składowania i przetwarzania danych multimedialnych odpowiada interMedia [16].

Celem niniejszego artykułu jest przedstawienie i ocena możliwości oferowanych przez Oracle interMedia w zakresie wsparcia dla składowania i przetwarzania danych multimedialnych w bazie danych. Punktem odniesienia dla oceny funkcjonalności interMedia będzie kształtujący się standard SQL/MM [12] oraz rozwiązania zaproponowane w ramach prototypów badawczych. Artykuł odnosi się do najnowszej dostępnej obecnie wersji serwera bazy danych Oracle: Oracle9i Release 2 (9.2).

Struktura artykułu przedstawia się następująco. Rozdział 2 zawiera ogólną charakterystykę danych multimedialnych oraz opis trzech najpowszechniej obecnie wykorzystywanych typów multimedialnych: obrazów, sekwencji audio i filmów wideo. Rozdział 3 stanowi przegląd najciekawszych prototypowych systemów multimedialnych baz danych. Rozdział 4 poświęcony jest standardowi SQL/MM. Rozdział 5 przedstawia funkcjonalność i architekturę Oracle interMedia. Rozdział 6 zawiera szczegółową analizę różnic między interMedia i innymi rozwiązaniami dla multimedialnych baz danych. Rozdział 7 stanowi krótkie podsumowanie.

2. Charakterystyka danych multimedialnych

Dane multimedialne to dane zakodowane w jednym lub wielu mediach, spośród których co najmniej jedno medium nie jest alfanumeryczne. Obecnie, w praktycznych zastosowaniach wykorzystywane są media ukierunkowane na ludzkie zmysły wzroku i słuchu, choć nie można wykluczyć, że w przyszłości przekazy multimedialne zostaną uzupełnione np. o zapach, smak, czy też dotyk (prowadzone są już eksperymenty w tym zakresie). Podstawowa klasyfikacja mediów to podział na media niezmiennie w czasie (dyskretne, statyczne) i zmiennie w czasie (ciągłe, dynamiczne). Do pierwszej grupy należą: teksty, obrazy (fotografie) i grafiki (rysunki, schematy). Druga grupa obejmuje: filmy wideo, sekwencje audio (muzyka, mowa) oraz animacje komputerowe. W praktyce, mianem danych multimedialnych określa się najczęściej obrazy oraz dane audio i wideo.

Wspólną cechą obiektów multimedialnych jest ich duży rozmiar i powszechność stosowania różnych metod kompresji. Celem kompresji jest redukcja rozmiaru danych, co zmniejsza wymagania w zakresie pojemności pamięci masowych i przepustowości kanałów transmisyjnych. Przetwarzanie danych skompresowanych wymaga jednak dużych mocy obliczeniowych, choć typowo

dekompresja jest znacznie mniej czasochłonna niż kompresja. Metody kompresji danych multimedialnych można ogólnie podzielić na metody bez utraty jakości (ang. lossless compression) i z utratą jakości (ang. lossy compression). W przypadku metod bez utraty jakości, po dekompresji otrzymana postać jest identyczna z postacią źródłową poddaną kompresji. Metody z utratą jakości wykorzystują niedoskonałość ludzkich zmysłów. Ceną za lepszy stopień kompresji jest w ich przypadku wprawdzie utrata części informacji, jednak metody te są tak projektowane, aby utrata jakości była niezauważalna lub prawie niezauważalna dla człowieka.

Typowe operacje w zakresie przetwarzania danych multimedialnych to różnego rodzaju operacje edycyjne, konwersje formatów oraz prezentacja danych (kwestia nietrywialna i niezwykle istotna w przypadku danych ciągłych, gdzie wymagana jest synchronizacja i odpowiednia przepustowość kanałów komunikacyjnych dla akceptowalnej jakości przekazu). Podstawowe zadanie, które stawiane jest systemom multimedialnych baz danych, to wyszukiwanie obiektów spełniających kryteria zadane przez użytkownika, czyli innymi słowy przetwarzanie zapytań.

W kontekście multimedialnych baz danych można wyróżnić dwa typy zapytań: zapytania o metadane i wyszukiwanie w oparciu o zawartość (ang. content-based retrieval). Metadane stanowią alfanumeryczny opis obiektu multimedialnego i dotyczą jego parametrów (np. rozmiar, metoda kompresji) lub zawartości (np. autor, wykonawca, tytuł, słowa kluczowe). Obsługa zapytań o metadane, takich jak np. „wyszukaj wszystkie obrazy w formacie GIF o rozmiarze poniżej 30 KB”, czy też „znajdź wszystkie filmy reżysera X, zapisane w formacie AVI”, w zasadzie nie wykracza poza funkcjonalność tradycyjnych systemów relacyjnych. Wyzwanie dla twórców systemów zarządzania multimedialnymi bazami danych stanowi wyszukiwanie ze względu na zawartość, np. „wyszukaj obiekty (obrazy, utwory muzyczne) podobne do danego”, czy też „znajdź obrazy, na których dominuje kolor niebieski”. Jest to silny mechanizm, wymagający złożonych algorytmów analizy zawartości obiektów multimedialnych. Jego atrakcyjność, w porównaniu z selekcją opartą o metadane, wynika głównie z tego, że wprowadzanie metadanych przy dodawaniu kolejnych obiektów do bazy jest uciążliwe, a ponadto nie wszystkie właściwości można łatwo opisać słownie (np. tekstury obrazów). Wyszukiwanie ze względu na zawartość znajduje zastosowanie np. przy wykrywaniu plagiatów, projektowaniu mody, przeszukiwaniu wirtualnych galerii sztuki, czy też projektowaniu wnętrz.

2.1. Obrazy

Obrazy zapisane w postaci cyfrowej najczęściej tworzone są przy użyciu aparatu cyfrowego lub skanera. Mogą być również przechwytywane z sekwencji wideo pochodzących z kamery lub odtwarzanych np. z magnetowidu albo też generowane programowo. Obrazy mają postać dwuwymiarowej tablicy punktów zwanych pikselami. Podstawowe charakterystyki opisujące parametry obrazów zapisanych w postaci cyfrowej to: format, wymiary, liczba kolorów i metoda kompresji. Typowe metadane opisujące zawartość obrazów to: opis (np. opis obiektów przedstawionych na obrazie), data i miejsce wykonania fotografii, dane fotografa, czy też tytuł obrazu i dane o jego autorze (w przypadku fotografii dzieł malarskich)

2.2. Dane audio

Cyfrowe dane audio powstają przy użyciu urządzeń do rejestracji dźwięku, pochodzą z takich źródeł jak mikrofon, ale również mogą być generowane programowo. Urządzenia do cyfrowego zapisu dźwięku dostają na wejściu sygnał analogowy pochodzący z mikrofonu lub zapisany wcześniej na nośnikach magnetycznych i dokonują jego konwersji na zapis cyfrowy o zadanych charakterystykach. Podstawowe charakterystyki opisujące parametry cyfrowych danych audio to: format, typ kodowania, liczba kanałów, częstotliwość próbkowania, rozmiar próbki, metoda kompresji i czas trwania. Typowe metadane opisujące zawartość danych audio to: opis, data nagrania oraz autor, kompozytor lub wykonawca

2.3. Wideo

Cyfrowe dane wideo pochodzą najczęściej z kamery wideo, choć mogą być również generowane programowo. Urządzenia do cyfrowego zapisu wideo mogą otrzymać na wejściu sygnał analogowy pochodzący np. z kamery i dokonać jego konwersji na zapis cyfrowy o zadanych charakterystykach. Podstawowe charakterystyki opisujące parametry cyfrowych danych wideo to: format, typ kodowania, liczba klatek na sekundę, rozmiar klatki, metoda kompresji, liczba kolorów, czas trwania oraz częstotliwość strumienia bitów (ang. bit rate). Typowe metadane opisujące zawartość danych wideo to: opis, data nagrania, producent, reżyser, obsada, itp.

3. Przegląd prototypowych rozwiązań do obsługi danych multimedialnych w bazach danych

Rozwój technologiczny spowodował, że idea przechowywania dużej ilości danych multimedialnych w bazach danych stała się realna. Okazało się jednak, że tradycyjne struktury, a przede wszystkim języki zapytań nie uwzględniają złożonej charakterystyki mediów. Problem ten stał się prawdziwym wyzwaniem dla wielu informatyków, co zaowocowało niezliczoną wręcz liczbą systemów i języków eksperymentalnych. Rozwiązań aspirujących do roli standardu obejmującego całość zagadnienia powstało jednak niewiele. Zdecydowana większość systemów koncentrowała się na jednym typie danych multimedialnych (np. statyczne obrazy) lub na konkretnych zastosowaniach multimedii (np. medycyna). Szczególnie dużą uwagę poświęcono 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.

Przy tworzeniu systemów prototypowych wykorzystano dwa główne podejścia:

- tworzenie nowych języków zapytań od podstaw

Największą wadą jest w tym przypadku brak akceptacji użytkowników, których trudno przekonać do nauki i użytkowania nowego języka. Zwykle jednak języki tego typu w sposób bardziej naturalny są dopasowane do wymagań, które stawia się przed multimedialną bazą danych. Nowe języki są często stosowane w zupełnie nowych, prototypowych systemach zarządzania bazą danych.

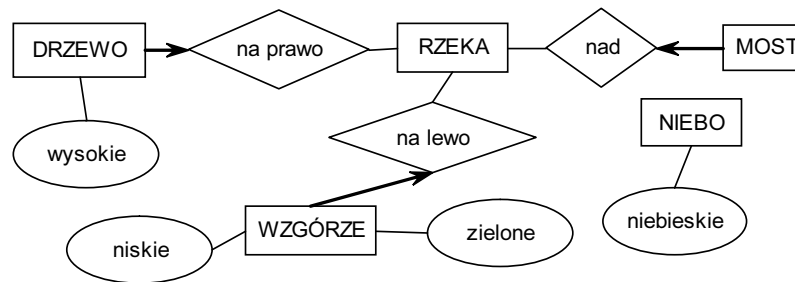
- rozszerzanie języka SQL lub OQL

Największą zaletą jest duża popularność języka SQL. Język ten będąc od dawna standardem dla baz relacyjnych ma dobrze opracowane podstawy teoretyczne. Rozszerzenia języka SQL do zastosowań multimedialnych na ogół opierają się na systemach postrelacyjnych i elementach obiektowych wprowadzonych do języka zapytań. Równie dobrze różnorodność typów mediów oraz mnogość zastosowań uwzględniają systemy typowo obiektowe i podobny do języka SQL język OQL.

Poniżej przedstawiono kilka przykładów systemów prototypowych, które pokazują różnorodne podejścia do kwestii przechowywania i wyszukiwania danych multimedialnych.

3.1. 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.



Rys. 1. Przykładowy diagram E-R reprezentujący zawartość obrazka z krajobrazem

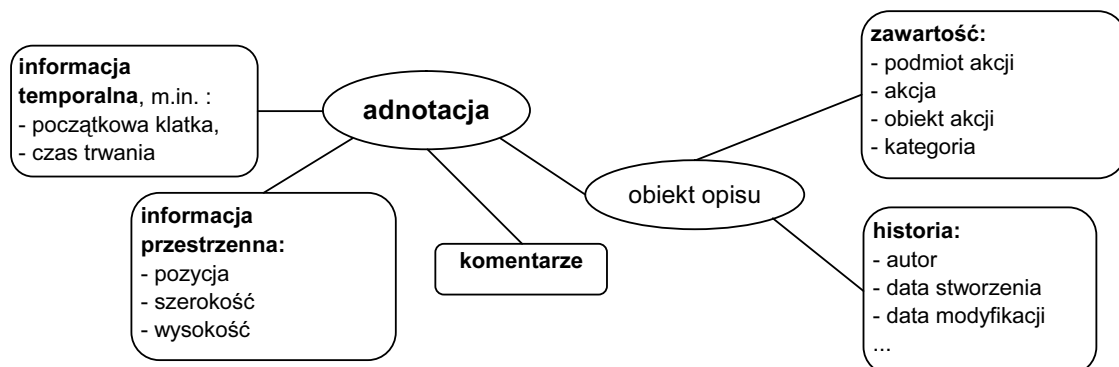
Oprócz związków skierowanych (jak na Rys. 1) mogą istnieć powiązania symetryczne, np. sosna stoi *obok* brzozy. Wyróżniono dwa typy związków. *Akcje* opisują pewne sytuacje obecne na obrazku (pies *goni* kota). *Relacje przestrzenne* określają względne pozycje obiektów na obrazie (*na lewo, pod, z przodu*).

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 następujące elementy:

- Przybliżone dopasowywanie wartości atrybutów oraz akcji. Wykorzystywana jest hierarchia generalizacji dla stosowanych pojęć. Przykładowo poszukiwanie drzewa zakończy się dopasowaniem również dla obrazków z sosną, o ile w hierarchii zdefiniowano, że sosna jest drzewem.
- Reguły dedukcji dla związków przestrzennych, które określają np. przechodniość takich relacji jak *na lewo*.
- Sortowanie wyników wg ściśle określonej miary podobieństwa.
- Użytkownik może iteracyjnie udoskonalać zapytanie (ang. selective relevance feedback). Wyniki zapytania przedstawione są jako pary (obrazek, jego symboliczna reprezentacja). Poprzez przenoszenie pewnych fragmentów wyniku do okna zapytania, można stopniowo dochodzić do pożądanego rezultatu.

3.2. TVQL (Temporal Visual Query Language)

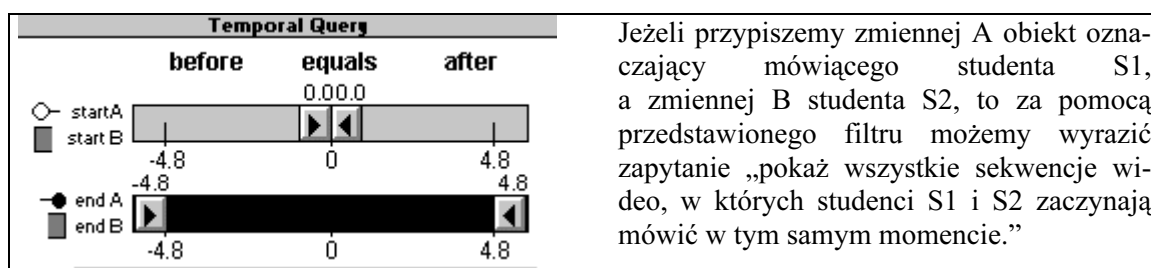
Język TVQL [3] koncentruje się na analizie czasowej danych wideo, wykorzystując przy tym ich cechy przestrzenno-temporalne. Te cechy, jak i pewna informacja semantyczna, są przechowywane w specjalnym modelu adnotacji (Rys. 2). Adnotacje przypisują sekwencji wideo informacje na temat zdarzenia, które jest na niej przedstawione.



Rys. 2. Fragment modelu adnotacji

Obiekt opisu przechowuje semantyczną informację o zdarzeniu, które reprezentuje adnotacja (np. przemawiający prezydent). Wiele adnotacji może wykorzystywać ten sam obiekt opisu, co oznacza występowanie takich samych zdarzeń w różnych sekwencjach wideo.

TVQL wykorzystuje tak zwane *dynamiczne zapytania*, czyli oferuje możliwość łatwej zmiany parametrów zapytania i niemal natychmiastowej obserwacji wpływu tych zmian na rezultat. Głównym zadaniem języka jest *analiza wideo* rozumiana jako odkrywanie czasowych związków między zdarzeniami. Żeby osiągnąć powyższe dwa cele odrzucono tekstowe sposoby formułowania zapytań i stworzono złożony interfejs graficzny. Główną innowacją są różnorodne suwaki (Rys. 3), które pozwalają na zdefiniowanie wszelkich możliwych temporalnych relacji między zdarzeniami.



Rys. 3. Użycie dynamicznych filtrów zapytań do wyrażenia związków czasowych

Okazuje się, że graficzne metody zadawania zapytań zaprezentowane w ramach języka TVQL oferują bardzo duże możliwości. Niestety TVQL nie wspiera innych typów mediów niż wideo. Wadą jest również brak możliwości sformułowania zapytań przybliżonych.

3.3. SMDS (Structured Multimedia Database System)

SMDS [11] 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). Przykładowe zapytanie: „znajdź obrazy, na których widoczny jest biały ford” ma postać:

$$(\exists S)(\text{frametype}(S, \text{image}) \& \text{ford} \in \text{flist}(S) \& \text{color}(\text{ford}, \text{white}, S))$$

Jeżeli zapytanie się powiedzie, rezultat wyświetlany jest na ekranie.

Formułowanie zapytań za pomocą programowania w logice daje niemal nieograniczone możliwości, jest natomiast dość trudne. Dlatego zaproponowano język oparty na składni SQL, który na niższym poziomie wykorzystuje odpowiednie formuły logiczne. Klauzula *SELECT* może zawierać obiekty multimedialne, a także ich atrybuty. Klauzula *FROM* może zawierać pozycje *<typ-medium> <źródło> <M>* oznaczające, że zostaną uwzględnione obiekty medialne tylko określonego typu i z określonego źródła. W ramach klauzuli *WHERE* można stosować konstrukcje *<zmienna> in <funkcja>*, gdzie *<funkcja>* reprezentuje wywołanie jednej z kilku funkcji dodatkowych, m.in. *FindType(Obj)*, która podaje typ danego obiektu, czy *FindFeaturesinObject(Obj)*, której wartością jest zbiór wszystkich cech przypisanych obiektowi medialnemu *Obj*. Odpowiednikiem zapytania przedstawionego wyżej, jest w języku SMDS-SQL poniższe zapytanie:

```

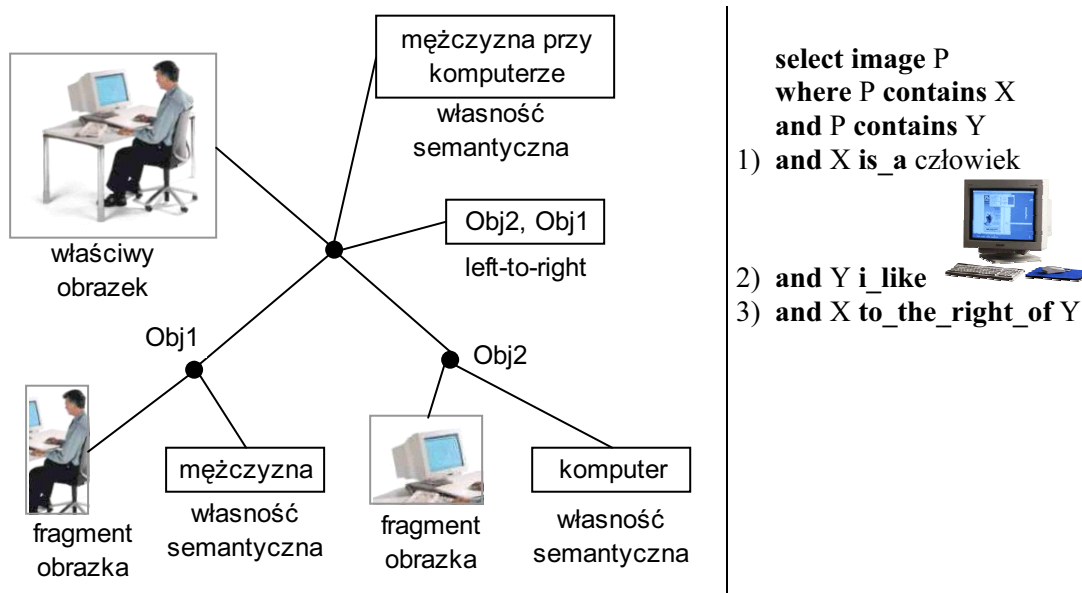
SELECT  M
FROM    smds source1 M
WHERE   FindType(M)=image
AND     FindObjWithFeature(ford)
AND     Color(ford, white, S)

```

3.4. CSQL (Cognition and Semantics-based Query Language)

Język CSQL [10] jest częścią systemu zarządzania bazy 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.



Rys. 4. Hierarchiczny model obrazu i przykład zapytania w języku CSQL

Hierarchia przedstawiona na Rys. 4 pokazuje, że obiekty składowe obrazu posiadają taką samą strukturę wewnętrzną jak obiekt nadrzędny, tzn. zarówno cechy wizualne jak i semantyczne. Mogą również zawierać kolejne podobiekty i relacje przestrzenne między nimi.

Przykładowe zapytanie z Rys. 4 to: „wyszukaj wszystkie obrazy, na których widoczna jest osoba na prawo od obiektu podobnego do obrazu X”. W zapytaniu tym występują trzy różne typy wyrażeń. Podzapytanie 1) odnosi się do semantyki obrazów. Dopasowanie do obrazu z lewej nastąpi,

gdy w bazie terminologicznej mężczyzna jest człowiekiem. Predykat *i_like* w wyrażeniu 2) wykorzystuje mechanizm rozpoznawania obrazów. Szukane są obiekty podobne do monitora wśród wszystkich obrazów w bazie danych. Podzapytanie 3) odnosi się natomiast do relacji przestrzennych, określonych w modelu obrazu.

3.5. MOQL (Multimedia Object Query Language)

Język MOQL [9] jest rozszerzeniem języka OQL opracowanego dla obiektowych baz danych przez ODMG. Celem twórców była reprezentacja dowolnych danych multimedialnych, niezależnie od ich typu i zastosowań. 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. predykat sprawdzający przecinanie się dwóch linii (*intersect*), funkcja obliczająca pole regionu (*area*),
- predykaty i funkcje temporalne – np. sprawdzanie, czy dwa interwały czasowe mają część wspólną (*overlap*),
- predykat *CONTAINS* sprawdza, czy dana multimedialna zawiera w sobie obiekt o ustalonej semantyce.

Składnia zapytań została również poszerzona o specjalną klauzulę *PRESENT*, która daje szerokie możliwości definiowania sposobu prezentacji wyników.

Przykładowe zapytanie wyszukuje pary (obraz, wideo), gdzie klip wideo obejmuje wszystkie samochody widoczne na obrazku. Wyświetla obraz i wideo 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)
```

4. Standard SQL/MM

Ponieważ standard języka SQL nie zawiera konstrukcji do obsługi takich danych jak multimedia, duże obiekty tekstowe, czy też dane przestrzenne, środowiska zajmujące się tworzeniem oprogramowania do przetwarzania tych specjalistycznych typów danych rozpoczęły pracę nad propozycjami rozszerzenia języka SQL o potrzebne im elementy. Organizacje odpowiedzialne za standard SQL zwróciły jednak uwagę na fakt, że ewentualne rozszerzenia SQL dedykowane dla poszczególnych rodzajów danych mogą być niekompatybilne ze sobą. Najłatwiejszym do zauważenia potencjalnym konfliktem był konflikt słów kluczowych. Przykładowo, słowo kluczowe *CONTAINS*, używane jest zarówno w kontekście tekstowych baz danych (do wskazania, że dane słowo lub fraza zawiera się w danym fragmencie tekstu), jak i baz przestrzennych i multimedialnych (do wskazania, że jeden obiekt zawiera w sobie inny obiekt).

Ze względu na wspomniane wyżej problemy zarzucono koncepcję rozszerzania języka SQL w zakresie wsparcia dla baz danych tekstowych, przestrzennych i multimedialnych. Zwyciężyła koncepcja opracowania nowego standardu, obejmującego specyfikacje bibliotek opartych o typy obiektowe SQL99 [13], przeznaczonych do obsługi poszczególnych specjalistycznych rodzajów danych i aplikacji. Nowy standard natychmiast stał się znany pod nazwą SQL/MM („MM” od „MultiMedia”). Dzięki oparciu specyfikacji SQL/MM na obiektowych typach SQL, funkcjonalność bibliotek jest w sposób naturalny dostępna z poziomu poleceń języka SQL, np. poprzez wywołania metod bibliotecznych typów obiektowych w wyrażeniach języka SQL.

Podobnie jak SQL, SQL/MM składa się z wielu części, przy czym części SQL/MM są ze sobą raczej luźno związane w porównaniu ze specyfikacjami SQL. Pełna nazwa standardu SQL/MM

brzmi: SQL Multimedia and Application Packages, co oznacza że jego części niekoniecznie muszą dotyczyć przetwarzania danych multimedialnych, ale również innych zastosowań. W chwili obecnej standard obejmuje cztery zatwierdzone części. Część pierwsza – *Framework* [4] ma charakter ogólny i nie dotyczy żadnego ze specjalistycznych rodzajów danych. Zawiera ona definicje i koncepcje wspólne dla pozostałych, specjalistycznych części standardu. Część pierwsza dotyczy między innymi sposobu, w jaki inne części standardu SQL/MM wykorzystują mechanizm obiektowych typów SQL. Część druga – *Full-Text* [5] definiuje typy danych do składowania dużych obiektów tekstowych, ze szczególnym naciskiem na wsparcie operacji wyszukiwania zadanych wzorców w tekście. Część 3 – *Spatial* [6] definiuje typy danych wspierające aplikacje przetwarzające dane przestrzenne. W obecnej wersji standardu wsparcie ograniczone jest do obiektów 0-wymiarowych (punktów), 1-wymiarowych (linii) i 2-wymiarowych (figur płaskich). Część piąta – *Still Image* [7] definiuje typy danych do obsługi obrazów w bazie danych. Z kolei „najświeższa”, szóstą część standardu SQL/MM – *Data mining* [8], definiuje typy danych do realizacji zadań eksploracji danych. Nie ma w standardzie SQL/MM części czwartej. Miała ona dotyczyć ogólnych operacji matematycznych (*General Purpose Facilities*), ale prace nad nią zarzucono kilka lat temu. Być może w przyszłości pojawią się natomiast kolejne części SQL/MM dotyczące multimediiów, np. danych wideo. Ewentualne poszerzenie SQL/MM o nowe części będzie zależało od zapotrzebowania i odzewu środowisk bazodanowych na istniejące specyfikacje SQL/MM.

Z punktu widzenia niniejszego artykułu, istotna jest część piąta standardu SQL/MM: *Still Image*, gdyż jedynie ona dotyczy multimediiów w rozumieniu Oracle interMedia. Obsługa danych tekstowych, przestrzennych i eksploracja danych nie wchodzi w zakres funkcjonalności Oracle interMedia (poza podstawowym wsparciem dla danych przestrzennych dostępnym w ramach interMedia jako Oracle Locator) Funkcje te są realizowane przez inne elementy systemu zarządzania bazą danych Oracle (w wersji 9i).

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 dwuwymiarowe obrazy 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, konwersję 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
```

5. Funkcjonalność i architektura Oracle interMedia

Oracle interMedia jest właściwością (ang. feature) serwera Oracle9i umożliwiającą składowanie, odczyt i przetwarzanie obrazów, danych audio i wideo oraz heterogenicznych danych multimedialnych. (Częścią Oracle interMedia jest ponadto Oracle Locator oferujący podstawowe wsparcie w zakresie przetwarzania danych geograficznych w aplikacjach internetowych.) Oracle interMedia opiera się na obiektowo-relacyjnych własnościach serwera Oracle9i, wykorzystując wsparcie dla składowania w bazie danych dużych obiektów binarnych (BLOB) i możliwość definiowania obiektowych (strukturalnych) typów danych SQL.

Oracle interMedia nie jest gotową aplikacją do obsługi danych multimedialnych, a jedynie udostępnia w ramach serwera Oracle9i mechanizmy wspierające tworzenie różnych aplikacji multimedialnych. Oracle interMedia oferuje typową funkcjonalność wymaganą w aplikacjach multimedialnych, pozwalając na dodawanie rozszerzeń dla zaspokojenia specyficznych potrzeb aplikacji. Przewidywane przykładowe zastosowania interMedia to:

- internetowe sklepy z muzyką i filmami wideo;
- internetowe galerie obrazów i fotografii;
- bazy danych zapisów rozmów telefonicznych, przemówień, prezentacji;
- biblioteki kursów w postaci filmów wideo.

Funkcjonalność Oracle interMedia jest dostępna głównie poprzez 4 typy obiektowe (zdefiniowane w schemacie *ORDSYS*):

- *ORDAudio* – do składowania i przetwarzania obiektów audio;
- *ORDVideo* – do składowania i przetwarzania obiektów wideo;
- *ORDImage* – do składowania i przetwarzania obrazów;
- *ORDDoc* – do składowania heterogenicznych obiektów multimedialnych.

Typ *ORDDoc* może mieć dwa różne zastosowania. Po pierwsze, może służyć do składowania złożonych dokumentów zapisanych w formacie binarnym specyficznym dla danej aplikacji. Po drugie, może być wykorzystany do składowania danych różnego rodzaju: audio, wideo, obrazów i tekstów w jednej kolumnie tabeli.

Wszystkie cztery wyżej wymienione typy zawierają atrybuty reprezentujące obiekt multimedialny i metadane go opisujące oraz metody do przetwarzania obiektów danego rodzaju. We wszystkich tych typach obiekt multimedialny jest reprezentowany przez atrybut typu *ORDSource*. Typ *ORDSource* wspiera różne sposoby składowania obiektów multimedialnych i oferuje metody dotyczące operacji na źródle danych.

5.1. Składowanie danych multimedialnych

Oracle interMedia umożliwia przetwarzanie danych multimedialnych:

- składowanych lokalnie w bazie danych w postaci dużych obiektów binarnych (BLOB);
- składowanych w lokalnym systemie plików, dostępnych z poziomu bazy danych jako tzw. duże obiekty plikowe (BFILE);
- dostępnych pod danym adresem URL, udostępnianych przez serwer WWW;
- udostępnianych przez specjalistyczne serwery np. dostarczające dane audio lub wideo strumieniowo.

Typ danych *ORDSource* bezpośrednio wspiera obiekty składowane jako BLOB, BFILE i dostępne przez URL. Aby skorzystać z poziomu interMedia z innych źródeł danych takich jak np. specjalizowane serwery wideo, konieczne jest rozszerzenie interMedia poprzez zaimplementowanie stosownych interfejsów.

Zaletą składowania danych obiektu multimedialnego w postaci BLOB w bazie danych jest to, że w takim wypadku zarówno metadane jak i zawartość binarna są objęte przetwarzaniem transakcyjnym i ochroną dostępu. Z kolei składowanie danych binarnych poza bazą danych (np. jako BFILE) umożliwia łatwą migrację do środowiska Oracle interMedia już istniejących poza bazą danych plikowych repozytoriów multimedialnych. W przypadku składowania binariów poza bazą danych, przetwarzaniem transakcyjnym objęte są tylko metadane i „wskaźnik” do lokalizacji binarnej zawartości obiektu.

Jak wspomniano wcześniej, każdy z typów: *ORDAudio*, *ORDDoc*, *ORDImage* i *ORDVideo* zawiera atrybut *ORDSource*. Dzięki takiemu rozwiązaniu, wszystkie kwestie związane ze składowaniem obiektów multimedialnych są obsługiwane przez jeden dedykowany do tego celu typ danych. Ponadto, z punktu widzenia interfejsów typów *ORDAudio*, *ORDDoc*, *ORDImage* i *ORDVideo*, sposób składowania obiektu jest nieistotny – kwestie związane z tym zagadnieniem są „ukryte” w typie *ORDSource*.

Typ *ORDSource* zawiera następujące atrybuty: typ źródła danych, informacja czy dane są dostępne lokalnie w bazie danych, pole BLOB z binarną zawartością obiektu (przy składowaniu lokalnym), adres lokalizacji i nazwa źródła (np. nazwa katalogu i pliku przy składowaniu w systemie plików) oraz data modyfikacji. Metody typu *ORDSource* pozwalają m.in. na odczyt zawartości obiektu oraz operacje eksport/import do migracji danych obiektu między bazą danych i zewnętrznymi lokalizacjami. Należy jednak zwrócić uwagę, że metody typu *ORDSource* nie powinny być bezpośrednio wywoływane z poziomu aplikacji. Aplikacje powinny korzystać z odpowiadających im metod dostępnych w typach *ORDAudio*, *ORDDoc*, *ORDImage* i *ORDVideo*.

5.2. Obiektowe typy danych dla obiektów multimedialnych

Jak wspomniano wcześniej, do reprezentacji obiektów multimedialnych w bazie danych Oracle interMedia oferuje obiektowe typy danych *ORDAudio*, *ORDDoc*, *ORDImage* i *ORDVideo*. Atrybutem występującym we wszystkich czterech typach jest atrybut opisanego wyżej typu *ORDSource*, stanowiący reprezentację danych obiektu multimedialnego. Pozostałe atrybuty typów *ORDAudio*, *ORDDoc*, *ORDImage* i *ORDVideo* zawierają metadane o składowanym obiekcie binarnym. Wszystkie cztery typy zawierają atrybuty dla formatu i typu MIME. Typ *ORDAudio* zawiera dodatkowo atrybuty dla typu kodowania, liczby kanałów, częstotliwości próbkowania, rozmiaru próbki, metody kompresji i czasu trwania. Typ *ORDImage* posiada atrybuty dla wysokości i szerokości obrazu, rozmiaru, formatu i metody kompresji. Z kolei typ *ORDVideo* zawiera atrybuty dla wysokości i szerokości klatki, rozdzielczości obrazu, liczby klatek na sekundę, czasu trwania, liczby klatek, metody kompresji, liczby kolorów i częstotliwości strumienia bitów. Ponadto, typy danych *ORDAudio*, *ORDDoc*, i *ORDVideo* zawierają atrybut komentarza, w którym mogą być umieszczone metadane o treści obiektu multimedialnego.

Oracle interMedia umożliwia składowanie i odczyt multimediiów w dowolnym formacie. Istotną własnością interMedia jest możliwość ekstrakcji metadanych o obiektach multimedialnych zapisanych w wielu popularnych formatach. Do tego celu służy metoda *setProperties* typów *ORDAudio*, *ORDDoc*, *ORDImage* i *ORDVideo*. Dla poszczególnych właściwości dostępne są metody typu „getter” i „setter” do ich odczytu i „ręcznej” modyfikacji. Zgodność zapamiętanych właściwości z rzeczywistą zawartością obiektu multimedialnego można zweryfikować metodą *checkProperties*.

Wspomniana metoda *setProperties* w typach *ORDAudio*, *ORDDoc*, i *ORDVideo*, dla rozpoznawanych przez interMedia formatów, oprócz ustawienia atrybutów reprezentujących podstawowe metadane, opcjonalnie umożliwia również generację rozbudowanego opisu właściwości obiektu w formacie XML, z uwzględnieniem nie tylko parametrów fizycznych obiektu, ale również informacji o zawartości merytorycznej. Oracle interMedia potrafi wydobyć informacje takie jak np. nazwę wykonawcy i tytuł utworu dla danych audio, oczywiście pod warunkiem, że format pliku audio umożliwia składowanie takich danych i że zostały one zapisane przez aplikację, za pomocą której utworzono dany plik. Generowany w postaci dokumentu XML opis jest umieszczany w atrybucie komentarza.

Oprócz wspierania wielu popularnych formatów obrazów, dokumentów, audio i wideo, Oracle interMedia umożliwia rozszerzenie zbioru rozpoznawanych (w zakresie ekstrakcji metadanych) formatów dokumentów, audio i wideo poprzez implementację specjalnych interfejsów. Wsparcie dla „niezrozumiałych” dla interMedia formatów obrazów realizowane jest w prostszy sposób – poprzez umożliwienie jawnego ustawienia wartości atrybutów z metadanymi za pomocą jednego z wariantów metody *setProperties* typu *ORDImage*.

Poza metodami dotyczącymi przetwarzania metadanych, dużą grupę metod typów *ORDAudio*, *ORDDoc*, *ORDImage* i *ORDVideo* stanowią metody do operowania na źródle danych (odczyt, eksport, import, itp.). Są to metody, które wywołują odpowiadające im metody typu *ORDSource* dla atrybutu reprezentującego źródło danych.

Modyfikacje zawartości obiektów multimedialnych są naturalnie wspierane przez interMedia tylko w stosunku do obrazów i jedynie w zakresie prostych operacji takich jak: konwersja formatu i/lub zmiana metody kompresji, kopiowanie, skalowanie i obcinanie obrazu. Istnieje możliwość rozszerzenia interMedia o operacje manipulacji na danych audio i wideo, ale tylko dla formatów zdefiniowanych przez użytkownika, w ramach implementacji interfejsu dla danego formatu.

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

5.3. Wyszukiwanie obrazów ze względu na zawartość

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).

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. Sygnatury muszą być jawnie wyznaczone (za pomocą metody *generateSignature* typu *ORDImageSignature*). Sygnatury obrazów mogą być składowane w tabelach tak samo jak obrazy (*ORDImage*) i inne dane.

Przy wyszukiwaniu ze względu na zawartość, użytkownik przypisuje wagi poszczególnym właściwościom obrazów. 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. Alternatywą dla tych dwóch operatorów jest metoda statyczna *evaluateScore()* typu *ORDImageSignature*, zwracająca tę samą wartość co *IMGScore*.

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

5.4. Korzystanie z interMedia - przykład

W celu zilustrowania sposobu korzystania z typów danych interMedia i wyszukiwania obrazów ze względu na zawartość, 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. Właściwości 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
```

5.5. Architektura Oracle interMedia

Oracle interMedia rozszerza funkcjonalność serwera bazy danych Oracle9i o możliwość składowania i przetwarzania danych multimedialnych. Dane multimedialne mogą w ten sposób być bezpiecznie współdzielone przez wielu użytkowników i udostępniane na skalowalnych platformach. Oracle interMedia opiera się na mechanizmach obiektowo-relacyjnych oferując nowe zło-

żone typy danych dla poszczególnych rodzajów multimediiów. Jednak dla zachowania zgodności z tradycyjnymi, relacyjnymi aplikacjami i dla umożliwienia przetwarzania danych multimedialnych składowanych bezpośrednio jako obiekty BLOB lub BFILE, interMedia oferuje tzw. interfejs relacyjny poprzez statyczne metody typów *ORDAudio*, *ORDDoc*, *ORDImage* i *ORDVideo*. Analiza (parsing) obiektów multimedialnych i przetwarzanie obrazów w bazie danych realizowane jest w środowisku wirtualnej maszyny Java wbudowanej w serwer Oracle9i.

Dostęp do funkcjonalności interMedia jest możliwy z poziomu wielu języków, ale szczególnie silnie wspierane są języki PL/SQL i Java. Skrypty w PL/SQL są podstawowym, obok narzędzia SQL*Loader, mechanizmem do ładowania multimediiów do bazy danych. Wsparcie dla łatwego korzystania z obiektowych typów interMedia w aplikacjach języka Java dostępne jest poprzez Oracle interMedia Java Classes [15], szczególnie zorientowane na przetwarzanie danych multimedialnych z poziomu serwetów i JSP.

Oracle interMedia nie posiada własnych mechanizmów dostarczania zawartości multimedialnej (audio i wideo) w trybie strumieniowym. Dostępny jest natomiast plug-in dla serwera RealServer G2, umożliwiający konfigurację, w której obiekty multimedialne z bazy danych Oracle są dostarczane strumieniowo do aplikacji klienta poprzez serwer RealServer G2.

Funkcjonalność interMedia dostępna jest głównie z poziomu serwera bazy danych. Jednakże z myślą o zewnętrznych w stosunku do bazy danych kolekcjach multimediiów (dostępnych np. na serwerach WWW lub płytach CD), Oracle dostarcza narzędzie Oracle interMedia Annotator [14], które pozwala na ekstrakcję opisu formatu i zawartości merytorycznej w postaci dokumentu XML dla obiektów składowanych poza bazą danych. Taki opis może być następnie załadowany do bazy danych i w niej przetwarzany.

6. Oracle interMedia na tle innych rozwiązań i standardu SQL/MM

Oceniając funkcjonalność Oracle interMedia w świetle możliwości oferowanych przez prototypy badawcze, należy na wstępie zwrócić uwagę na zupełnie inny charakter porównywanych systemów. Oracle jako system komercyjny musi kłaść szczególny nacisk na obsługę wielości formatów danych, z którymi mogą pracować potencjalni użytkownicy. Nie ulega wątpliwości, że w tym zakresie Oracle interMedia spełnia swoje zadanie, rozpoznając wiele popularnych formatów multimediiów z możliwością ekstrakcji metadanych opisujących format i zawartość obiektów multimedialnych. Co więcej, użytkownikom udostępniany jest interfejs umożliwiający dodanie obsługi nowych formatów, gdy zajdzie potrzeba ich analizy w bazie danych.

Twórcy systemów prototypowych głównie skupiali się na obsłudze zaawansowanych zapytań do multimedialnych baz danych, z kryteriami selekcji odwołującymi się do zawartości. Oracle interMedia obsługuje wyszukiwanie ze względu na zawartość tylko dla obrazów, podczas gdy niektóre z systemów prototypowych obsługiwały w tym zakresie również dane wideo. Należy jednak podkreślić, że propozycje metod wyszukiwania ze względu na zawartość w kontekście danych wideo bazują na tekstowych opisach semantycznych sekwencji wideo, poszczególnych scen i klatek. Takie podejście jest w dużym stopniu niepraktyczne, gdyż przygotowanie odpowiednich opisów jest czasochłonne, a nie może być dokonane bez znaczącego udziału człowieka.

Odnosząc się do zgodności interfejsu interMedia ze standardem SQL/MM, po pierwsze należy stwierdzić, że Oracle interMedia jest rozwiązaniem zgodnym ideowo z tym standardem, ze względu na wykorzystanie strukturalnych typów SQL i oparcie składowania multimediiów o duże obiekty binarne (BLOB) lub wskaźniki do źródeł zewnętrznych. Różnice w funkcjonalności i składni między interMedia a specyfikacją SQL/MM można rozpatrywać tylko w kontekście obsługi obrazów, gdyż SQL/MM nie zawiera jeszcze specyfikacji dla danych audio i wideo. Składnia interMedia i SQL/MM: *Still Image* jest odmienna, inne są nazwy typów i metod, ale funkcjonalność jest bardzo zbliżona. Z jednej strony, Oracle nie implementuje wszystkich przewidzianych w standardzie operacji na obrazach, gdyż nie umożliwia obrotów. Z drugiej strony, zbiór właściwości wizualnych branych pod uwagę przy wyszukiwaniu obrazów jest bogatszy w interMedia. Oracle interMedia uzupełnia możliwości SQL/MM o zapytania dotyczące kształtów i ich lokalizacji.

Choć głównym celem niniejszego artykułu jest ocena funkcjonalności interMedia na tle kształtującego się standardu i możliwości systemów prototypowych, nie sposób w tego typu opracowaniach uciec od porównań z konkurencyjnymi komercyjnymi systemami zarządzania bazami danych. Porównywalne i w zasadzie bardzo zbliżone możliwości w zakresie obsługi danych multimedialnych oferuje IBM DB2 (wersja 8.1). W IBM DB2 funkcjonalność związana z przetwarzaniem multimediów jest dostępna w ramach tzw. rozszerzeń: DB2 Image Extender, DB2 Audio Extender i DB2 Video Extender [2]. Funkcjonalność ta jest dostępna poprzez specyficzne typy SQL (ang. distinct types): DB2IMAGE, DB2AUDIO, DB2VIDEO i funkcje SQL na nich działające. Podobnie jak Oracle interMedia, rozszerzenia DB2 umożliwiają składowanie zawartości obiektów multimedialnych w bazie danych (jako BLOB) i systemie plików oraz automatyczne rozpoznawanie typowych formatów z ekstrakcją podstawowych metadanych o formacie obiektu. Wyszukiwanie ze względu na zawartość jest podobnie jak w interMedia ograniczone do obrazów (funkcje QBIC – Query By Image Content). Własności wizualne obrazów, do których mogą odwoływać się zapytania to: średni kolor, histogram kolorów, lokalizacja kolorów i tekstura (tak jak w SQL/MM). W zakresie manipulacji na obrazach DB2 wspiera operację obrotu, niedostępną w interMedia, poza tym zbiór dostępnych operacji jest zbliżony. Podobnie jak w przypadku interMedia, nazwy typów danych i funkcji do obsługi multimediów są w obecnej wersji DB2 inne niż proponowane przez standard SQL/MM. Główne zalety Oracle interMedia w porównaniu z DB2 to rozbudowane możliwości ekstrakcji metadanych o zawartości merytorycznej obiektów multimedialnych, możliwość zakładania indeksu na sygnaturach obrazów dla zwiększenia efektywności wyszukiwania ze względu na zawartość i zgodne z duchem standardu SQL/MM wykorzystanie typów strukturalnych.

7. Podsumowanie

Podsumowując powyższe rozważania, należy przyznać, że Oracle interMedia jest z pewnością atrakcyjnym rozwiązaniem dla użytkowników pracujących z dużymi kolekcjami obiektów multimedialnych. Funkcjonalność oferowana przez interMedia sprowadza się nie tylko do zapewnienia przetwarzania transakcyjnego w stosunku do danych multimedialnych, ale również pozwala na zaawansowane przetwarzanie metadanych dla wszystkich typów multimediów oraz podstawowe metody manipulacyjne i wyszukiwanie ze względu na zawartość dla obrazów.

Możliwości Oracle interMedia są zbliżone, a może nawet nieco przewyższające ofertę głównego konkurenta Oracle na rynku serwerów baz danych – IBM DB2. Problem niezgodności składniowej ze standardem SQL/MM, potencjalnie utrudniającej przenaszalność aplikacji, w chwili obecnej nie ma istotnego znaczenia, gdyż np. w IBM DB2 składnia odbiega od standardu jeszcze bardziej. W przyszłości, w miarę umacniania się pozycji standardu SQL/MM, najprawdopodobniej należy się spodziewać dostosowania składni interMedia do specyfikacji SQL/MM.

8. Bibliografia

- [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] Hibino S., Rundensteiner E.A.: A Visual Multimedia Query Language for Temporal Analysis of Video Data. Nwosu K., Thuraisingham B., Berra P.B. (Eds.): Multimedia Database Systems: Design and Implementation Strategies. Kluwer Academic Publishers, 1996.
- [4] ISO/IEC 13249-1:2002, Information Technology – Database Languages – SQL Multimedia and Application Packages – Part 1: Framework. ISO, 2002.

-
- [5] ISO/IEC 13249-2:2000, Information Technology – Database Languages – SQL Multimedia and Application Packages – Part 2: Full-Text. ISO, 2000.
 - [6] ISO/IEC 13249-3:1999, Information Technology – Database Languages – SQL Multimedia and Application Packages – Part 3: Spatial, ISO. 1999.
 - [7] ISO/IEC 13249-5:2001, Information Technology – Database Languages – SQL Multimedia and Application Packages – Part 5: Still Image. ISO, 2001.
 - [8] ISO/IEC 13249-6:2002, Information Technology – Database Languages – SQL Multimedia and Application Packages – Part 6: Data mining. ISO, 2002.
 - [9] 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.
 - [10] 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.
 - [11] Marcus S., Subrahmanian V.S.: Foundations of multimedia database systems. Journal of ACM, 43(3), 1996.
 - [12] Melton J., Eisenberg A.: SQL Multimedia and Application Packages (SQL/MM). SIGMOD Record 30(4), 2001.
 - [13] Melton J., Eisenberg A.: SQL:1999, formerly known as SQL3. SIGMOD Record 28(1), 1999.
 - [14] Oracle9i interMedia Annotator User's Guide, Release 2 (9.2). Oracle, 2002.
 - [15] Oracle9i interMedia Java Classes User's Guide and Reference, Release 2 (9.2). Oracle, 2002.
 - [16] Oracle9i interMedia User's Guide and Reference, Release 2 (9.2). Oracle, 2002.