

М. Медиковський, М. Чаплагін
Національний університет “Львівська політехніка”,
кафедра автоматизованих систем управління

ЕФЕКТИВНІСТЬ ПРОГРАМНОГО ЗАБЕЗПЕЧЕННЯ ДЛЯ СЕМАНТИЧНОГО ПОШУКУ ДАНИХ ПРО ГРАФІЧНІ ОБ’ЄКТИ

© Медиковський М., Чаплагін М., 2008

Досліджено ефективність розроблених алгоритмів та програмного забезпечення для семантичного пошуку даних про графічні об’єкти. Проаналізовано результати експериментальних досліджень. Розглянуто можливі способи підвищення ефективності і точності пошуку.

Developed algorithms and programs efficiency for semantic search of data about graphic objects examined. Analysis of experimental researches was executed. Possible way to increase search efficiency and precise was considered.

Вступ

Розвиток і запровадження інформаційних технологій стримуються значними затратами часу та інших ресурсів систем для розміщення, пошуку і відображення даних. Той факт, що людина-оператор більшу частину інформації отримує через зоровий апарат, зумовлює переважне використання графічних об’єктів у інтерфейсі машина-людина. Для прийняття рішень оператору потрібні лише ті графічні об’єкти, які забезпечують розв’язання поставленої задачі, що вимагає організувати пошук потрібних графічних об’єктів серед всієї множини у сховищі даних. Порівняння заданого фрагмента з отриманим результатом пошуку проводять на боці користувача з використанням графічного подання даних, що зумовлює додаткові затрати часу. Особливі ускладнення виникають при опрацюванні даних про графічні об’єкти, оскільки кожне зображення можна охарактеризувати за різними критеріями. Крім того, постійне зростання обсягів інформації за незмінних параметрів інсталюваної автоматизованої системи управління потоками даних зумовлює зниження її ефективності.

Вирішення завдання зменшення потоків даних про графічні об’єкти здійснювалося переважно шляхом розроблення форматів кодування та компресії, особливо ефективні з яких використовують методи статистичного ущільнення (компресії без втрат) та врахування психовізуальних і психофізичних властивостей сприйняття мультимедійної інформації людиною [1]. Проте за реалізацією це швидше екстенсивний, ніж інтенсивний шлях вирішення проблеми зменшення потоків даних між окремими вузлами системи.

Розглянуто підходи до можливості підвищення ефективності пошуку даних про графічні об’єкти на основі їх семантики. Матеріал, викладений у статті, ґрунтується на результатах раніше проведених досліджень [2–4].

Постановка задачі

Аналіз відомих результатів показує, що значного спрощення розв’язання окремих задач зберігання та пошуку даних, зокрема про графічні об’єкти, можна досягти шляхом їх структуризації в сховищах з використанням для цього особливим чином створеного семантичного опису. Такі методи практично не застосовуються для організації пошуку даних через недостатність системних досліджень способів опису та технологій застосування.

Тому перспективним шляхом розв'язання вказаної задачі є застосування методів інтелектуалізації систем зберігання та пошуку графічних даних з використанням логіко-математичного опису їхньої семантики, що забезпечує удосконалення методів зберігання, зменшення обсягів передавання даних і тим самим підвищення ефективності автоматизованих систем.

Для оцінювання процесу пошуку та визначення можливостей покращання результатів проведено такий експеримент: вибрано два типи, але різнопланові рисунки «River» та «Lena», з яких виділено графічні примітиви існуючими програмними засобами (рис. 1–2). За розробленими в [3] методами та алгоритмами визначаємо початкову семантичну значущість графічних примітивів (табл. 1). Порівняння даних про ці рисунки з використанням розроблених у [4] семантичних фільтрів показує, що за алгоритмами пошуку вони будуть визначені не подібними.

Надалі з одного чи обох рисунків вилучається частина примітивів з таких міркувань, щоби рисунки за наборами примітивів могли бути визначені першим семантичним фільтром як подібні. Причому, через вилучення частини примітивів змінюється семантична значущість тих, що залишаються (через зміну їх взаємного розташування).

Метою дослідження є визначення кількості примітивів, які можуть бути вилучені, а також їхньої ваги у описі об'єкта з тим, щоби описи були визначені як подібні всіма трьома семантичними фільтрами.

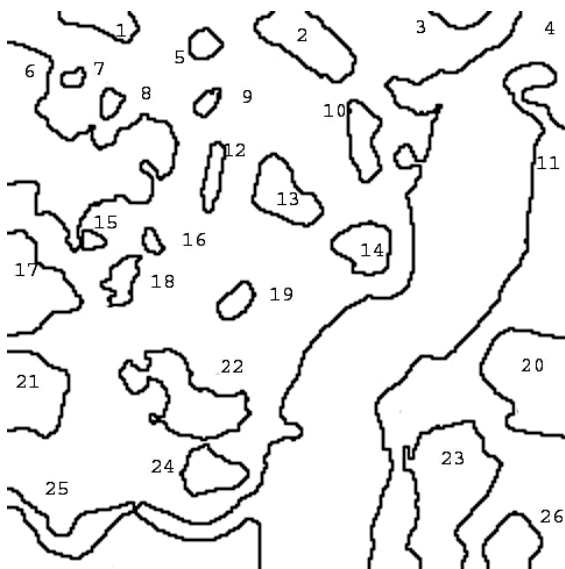


Рис. 1. Графічні примітиви тестового зображення «River»

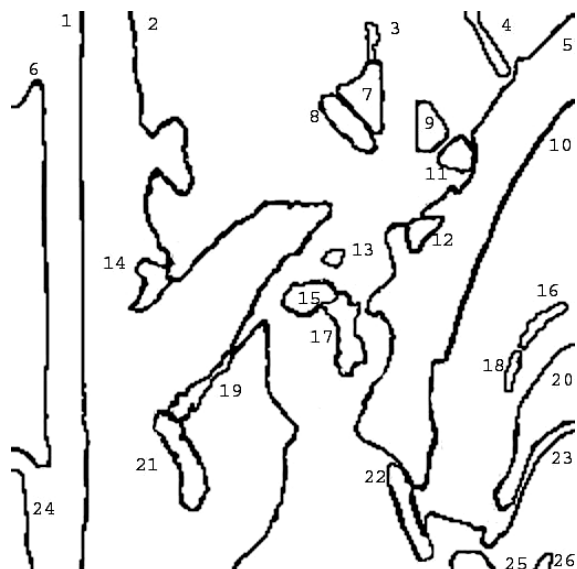


Рис. 2. Графічні примітиви тестового зображення «Lena»

Таблиця 1

Семантична значущість графічних примітивів

№ з/п	Тип графічного примітива рис. «River»	Семантична значущість графічного примітива рис. «River»	Тип графічного примітива рис. «Lena»	Семантична значущість графічного примітива рис. «Lena»
1	2	3	4	5
1	curve	3	line	5
2	curve	3	curve	4
3	curve	3	fuzzy polygon	3
4	curve	3	curve	4
5	fuzzy circle	3,5	curve	4
6	curve	3	curve	4
7	fuzzy polygon	3	fuzzy polygon	3
8	fuzzy polygon	3	fuzzy polygon	3
9	fuzzy polygon	3	fuzzy triangle	4
10	fuzzy polygon	3	curve	4

1	2	3	4	5
11	Curve	3	fuzzy polygon	3
12	fuzzy polygon	3	fuzzy polygon	3
13	fuzzy polygon	3	fuzzy circle	5
14	fuzzy polygon	3	fuzzy polygon	3
15	fuzzy triangle	4	fuzzy polygon	3
16	fuzzy polygon	3	fuzzy polygon	3
17	curve	3,5	fuzzy polygon	3
18	fuzzy polygon	3	fuzzy polygon	3
19	fuzzy polygon	3	fuzzy polygon	3
20	curve	3	curve	4
21	curve	3	fuzzy polygon	3
22	fuzzy polygon	3	fuzzy polygon	3
23	curve	3	curve	4
24	fuzzy polygon	3	curve	4
25	curve	3	curve	4
26	curve	3	curve	4

Результати досліджень

Процедура дослідження полягає у тому, що запит та об'єкт піддавалися семантичному аналізу і визначалося, чи будуть об'єкт та запит визнані як семантично близькі з урахуванням вилучених примітивів і яка їхня загальна семантична значущість. За межу збігу семантичної значущості запиту та семантичної значущості об'єкта прийнято 5% від семантичної значущості об'єкта. Результати експерименту наведено у табл. 2.

Таблиця 2

Результати експерименту з визначення граничного рівня семантичної суперечності

№ з/п	Запит	Який примітив запиту відкинуто	Об'єкт	Який примітив об'єкта відкинуто	Відповідає запиту	На якому фільтрі відкинуто	Семантична суперечність на цьому фільтрі
1	2	3	4	5	6	7	8
1	Lena	жоден	River	жоден	ні	1	(перевищення кількості примітивів першого типу у запита над об'єктом)
2	Lena	1 line	River	жоден	ні	1	(перевищення кількості примітивів першого типу у запита над об'єктом)
3	Lena	1 line 3 fuzzy polygon	River	жоден	ні	2	5
4	Lena	1 line 10 curve 24 curve	River	жоден	ні	1	(перевищення кількості примітивів першого типу у запита над об'єктом)
5	Lena	1 line 3 fuzzy polygon 9 fuzzy triangle	River	жоден	так	---	на 2-му фільтрі суперечність 0, на 3-му фільтрі суперечність 1
6	Lena	1 line 3 fuzzy polygon 12 fuzzy polygon 21 fuzzy polygon	River	жоден	так	---	на 2-му фільтрі суперечність 0, на 3-му фільтрі суперечність 0

1	2	3	4	5	6	7	8
7	Lena	1 line 3 fuzzy polygon 10 curve 24 curve	River	жоден	так	---	на 2-му фільтрі суперечність 2, на 3-му фільтрі суперечність 2
8	Lena	1 line 3 fuzzy polygon	River	7 fuzzy polygon	ні	2	5
9	Lena	1 line 3 fuzzy polygon 10 curve 24 curve	River	3 curve	ні	2	13
10	Lena	1 line 3 fuzzy polygon 12 fuzzy polygon 21 fuzzy polygon	River	7 fuzzy polygon	ні	2	21
11	Lena	1 line 3 fuzzy polygon	River	7 fuzzy polygon 12 fuzzy polygon 19 fuzzy polygon	ні	3	значення суперечності на 2-му фільтрі 4; на третьому 28
12	River	жоден	Lena	жоден	ні	2	14
13	River	1 curve	Lena	жоден	ні	3	значення суперечності на 2-му фільтрі 4; на третьому 12
14	River	8 fuzzy polygon	Lena	жоден	ні	3	значення суперечності на 2-му фільтрі 4; на третьому 27
15	River	1 curve 8 fuzzy polygon	Lena	жоден	ні	2	20
16	River	1 curve 2 curve	Lena	жоден	ні	3	значення суперечності на 2-му фільтрі 3; на третьому 8
17	River	1 curve 20 curve	Lena	1 line	так	---	значення суперечності на 2-му фільтрі 4; на 3-му фільтрі 3
18	River	1 curve 26 curve	Lena	1 line	так	---	значення суперечності на 2-му фільтрі 2; на 3-му фільтрі 3
19	River	1 curve 21 curve 24 fuzzy polygon	Lena	1 line 21 fuzzy polygon	так	---	значення суперечності на 2-му фільтрі 4; на 3-му фільтрі 3

Аналіз експериментальних даних показує, що для того, щоб запит, сформований з семантичного опису об'єкта «Lena», був визнаний семантично близьким до графічного об'єкта «River», він повинен втратити від 3-х графічних примітивів і більше (експерименти №№ 5, 6 та 7). Водночас, втрата примітивів об'єктом «River», коли він є запитом, у більшості випадків зумовлює зростання значення семантичної суперечності між цими об'єктами (можна порівняти значення семантичної суперечності експериментів №15 та №14 з експериментом №12). Зменшення загального значення семантичної суперечності відбулося лише у експерименті №13. У більшості випадків елементи

відфільтровуються вже на другому фільтрі. В експериментах № 3 та №8 величина семантичної суперечності на другому фільтрі наблизилася впритул до встановленого експериментального граничного значення. І навіть у випадку проходження цих фільтрів значення семантичної суперечності на третьому фільтрі становили би понад 10.

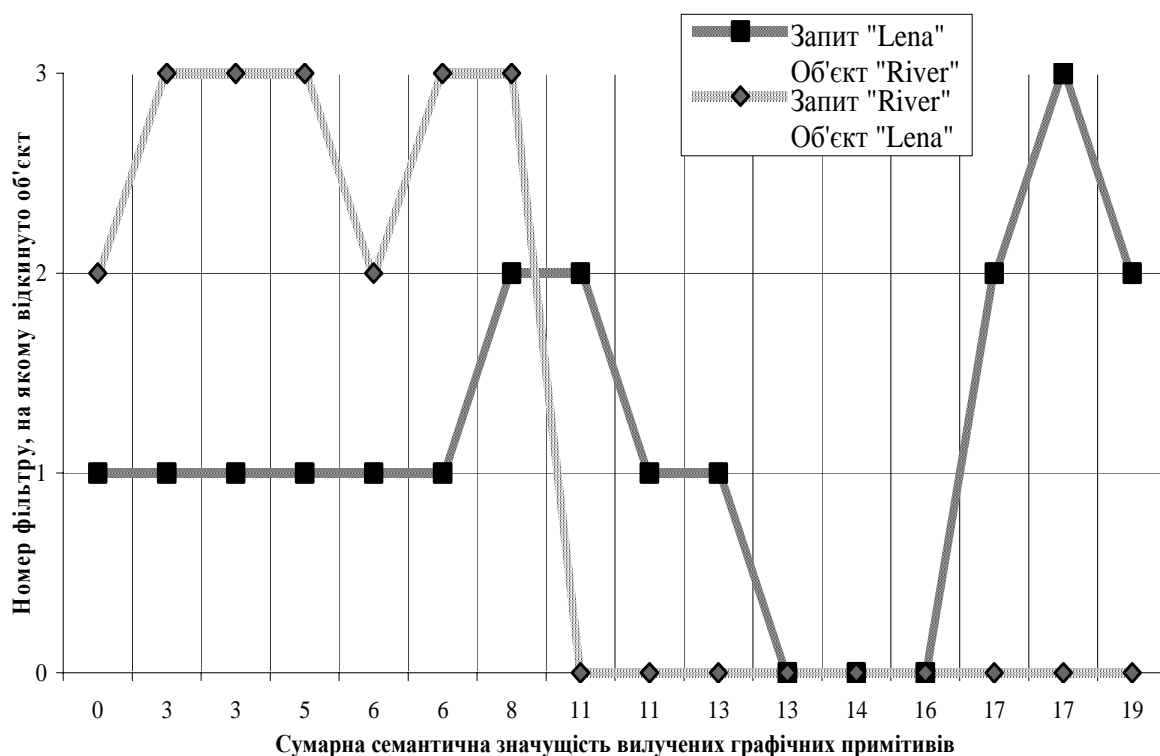


Рис. 3. Залежність подібності від сумарної семантичної значущості вилучених графічних примітивів

Висновки

Прийнята експериментальна межа подібності у 5% для відсіювання об'єктів, що не відповідають запиту, є ефективною і достатньою. Проте будь-який вплив у початковій фазі аналізу зображення, який може спотворити або вилучити певні контури з графічного об'єкта, має істотне значення для результатів пошуку, що вимагає детального дослідження властивостей алгоритмів встановлення контурів зображення для визначення або розроблення такого, що даватиме стабільний результат з мінімальним ризиком спотворення кінцевого набору графічних примітивів. Крім того, доцільно провести дослідження для визначення додаткових значущих показників та додаткових фільтрів, які б давали змогу збільшити точність пошуку та зменшити вплив спотворення чи відсутності певних примітивів на об'єкті. Такими показниками можуть бути колір, яскравість або введення додаткових параметрів взаємного розташування примітивів для врахування їх перетинів, вкладень тощо.

1. Ватолин Д., Ратушняк А., Смирнов М., Юкин В. Методы сжатия данных. Устройство архиваторов, сжатие изображений и видео. – М.: Диалог-МИФИ, 2002. – 384 с. 2. Медиковський М.О., Чаплагін М.П., Метод розроблення словника для побудови семантичного опису графічного об'єкта // Вісник Нац. ун-ту "Львівська політехніка". – 2007. – № 598. – С.149–153. 3. Тимченко О.В., Чаплагін М.П. Оцінювання вмісту графічного об'єкту на основі його семантики та її використання для організації пошуку графічних об'єктів в АСУ потоками даних // Зб. наук. пр. ІПМЕ НАН України. – 2007. – Вип. 39. – С.12–16. 4. Тимченко М.О., Чаплагін М.П. Розроблення семантичних фільтрів для організації пошуку графічних об'єктів в автоматизованих системах управління потоками даних // Тез. доп. XXVI Науково-практичної конференції „Моделювання”. ІПМЕ ім. Г.Є. Пухова. – К., 2007. – С.75–76.