

УДК 004.932

Методи покращення результатів пошуку зображень в цифрових колекціях

Башков Є.О., Вовк О.Л., Костюкова Н.С.
Донецький національний технічний університет
bashkov@pmi.dgtu.donetsk.ua

Abstract

Bashkov E.O., Vovk O.L., Kostyukova N.S. Results improvement methods for retrieval in digital image collection. This article is devoted to the problem of content-based image retrieval effectiveness increasing and redundancy reducing for search result. Research and experiments show that proposed approaches are appropriate for image retrieval. Using of multisets for relevant images selection allow to increase search result redundancy.

Проблема пошуку зображень в базах даних

Важливою і цікавою з погляду проблематики побудови систем штучного інтелекту є задача пошуку зображень, схожих на заданий зразок, тісно пов'язана з областю аналізу візуальних образів. Така задача зазвичай вирішується людиною інтуїтивно і практично миттєво, але для комп'ютерної реалізації є досить складною. Її складність обумовлена як проблематичністю адекватного формального опису зображення, так і недостатнім вивченням фізіології людського зору і сприйняття візуальної інформації.

В загальному вигляді задача пошуку в колекції (базі даних) зображень, схожих на заданий зразок, формулюється таким чином.

Нехай є цифрова колекція, що містить інформацію про V зображень:

$$S = \{P, F\},$$

де $P = \{Pk, k=1, 2..V\}$ – множина зображень (матриць $M*N$, елементи яких містять дані про кольори пікселів зображення; M і N – ширина і висота зображення відповідно), $F = \{Fk, k=1, 2..V\}$ – множина характеристик вмісту зображень (Fk – скаляр або вектор).

Заданий зразок пошуку $P_{зр}$ – зображення, вміст якого описується характеристикою $F_{зр}$. Для зразка пошуку $P_{зр}$ і будь-якого зображення P_k може бути обчислений ступінь схожості їх вмісту:

$$d_k = f(F_{обр}, F_k), k = 1, 2, \dots, V.$$

Потрібно визначити множину зображень Q , схожих із зразком і розташованих в порядку зменшення цієї схожості:

$$Q = \{P_i, i = 1, 2, \dots | \forall i_1 > i_2 : d_{i_1} < d_{i_2}\}.$$

Таким чином, задача контекстного пошуку зводиться до обчислення значень $dk, k=1, 2..V$, і їх подальшого сортування. При цьому обчислення характеристик F вмісту зображень виконується при занесенні цих зображень в базу даних, де

вони і зберігаються.

Значною проблемою при обробці результатів пошуку зображень є їх надмірність. В даній роботі пропонується модифікація етапу контекстного пошуку, на якому проводиться відбір зображень для пред'явлення користувачу як результати. Суть модифікації полягає у використанні для відбору апарату мультимножин.

Аналіз сучасних досліджень у галузі пошуку зображень за зразком

Сучасні дослідження в області розробки систем пошуку зображень в основному направлені на рішення двох задач: достовірний формальний опис вмісту зображення і ефективне з погляду часу і витрат пам'яті зіставлення цих описів. Існує декілька підходів, орієнтованих на специфічні класи зображень, проте найбільш проблематично рішення даної задачі для випадку довільних зображень. Запити користувачів до колекцій зображень традиційно класифікують по трьох рівнях абстракції [1]: примітивний рівень (пошук за візуальними примітивами: кольором, формою, текстурою), логічний рівень (ідентифікація представлених об'єктів – наприклад, пошук зображень Ейфелевої вежі), абстрактний рівень (облік значущості зображених сцен – пошук зображень, що відображують певні настрої). На першому (примітивному) рівні розробники використовують для опису вмісту кольорні гістограми [2], які описують розподіл кольорів точок зображення, але не дають інформації про розташування цих кольорів. Одним із способів обліку просторової інформації про вміст зображення є виділення ознак окремих областей зображення, причому самі області можуть виділятися різними способами. Найменш складним є ручне виділення областей, коли зображення обробляється людиною і вся необхідна інформація визначається візуально. Зокрема, цей підхід використовували розробники

системи QVIC [2]. Ручне виділення областей і об'єктів є дуже громіздким і для великих наборів зображень вимагає значних витрат часу.

Інший підхід полягає в тому, що області зображення виділяються шляхом розбиття зображення на фіксовані фрагменти [3], але в загальному випадку складно вибрати для кожного зображення масштаб розділення, і, крім того, така операція не є інваріантною щодо зсуву і масштабування.

Третій підхід включає сегментацію зображень – розбиття зображення $I[x, y]$ на гомогенні сегменти R_i , що не перекриваються і утворюють повний набір [4, 5, 6]. Сегментація не є чітко сформульованою задачею для випадку довільних зображень, оскільки поняття «гомогенність» залежить від області використання зображень. Більш того, для довільних зображень немає формального способу оцінки коректності виконання сегментації, а задача сегментації зображення з урахуванням всіх обмежень не має єдиного рішення [4]. Ще одним способом представлення кольорового вмісту зображення є кольорова корелограма [7], яка, на відміну від гістограми, враховує просторовий розподіл кольорів в зображенні. В [7] показано, що найкращою для порівняння кольорних корелограм є їх кон'юнкція.

При використанні гістограмних ознак для обчислення ступеня схожості зображень використовують евклідову відстань, кон'юнкцію гістограм, косинусну або квадратичну відстані [3]. Будь-яка з цих величин сама по собі не відображає ступінь схожості двох зображень і може використовуватись лише при порівнянні з іншими аналогічними величинами. Саме з цієї причини всі практичні реалізації контекстного пошуку зображень виконують сортування всіх зображень з бази даних, що є головним недоліком цих реалізацій.

Пошук зображень за їх описом

В процесі контекстного пошуку зображень послідовно виконуються квантування кольорів або рівнів яскравості, побудова гістограмної ознаки вмісту зображення, порівняння зображень, сортування зображень. Для представлення інформації про кольори в зображенні авторами в [8] пропонується будувати гістограму, аналогічну текстурній, але враховувати співвідношення кольорів пар пікселів, а не тільки складову яскравості (побудована таким чином характеристика називається 2D-кольорою гістограмою). 2D- кольорова гістограма будується як двовимірний масив $S_{max} \times S_{max}$, де S_{max} – число кольорів базового набору, що використався на етапі квантування. Дана характеристика є матрицею, кожний елемент якої містить нормовану кількість пар пікселів з певними кольорами навколо точки. Для порівняння 2D –

кольорних гістограм зображень пропонується обчислювати коефіцієнт їх кореляції, оскільки 2D-кольорова гістограма є випадковим вектором розмірності $S_{max} \times 2$. Враховуючи властивості коефіцієнта кореляції, при формуванні набору результуючих зображень їх слід розміщувати в порядку убутання коефіцієнта.

Оцінка переваг пошуку з використанням 2D – кольорних гістограм виконувалася на основі характеристик точності (Precision) і повноти (Recall), введених в [3]. Дані характеристики, отримані при пошуку за 2D – кольорними гістограмами, порівнювалися з характеристиками пошуку за традиційними кольорними гістограмами (рис. 1).

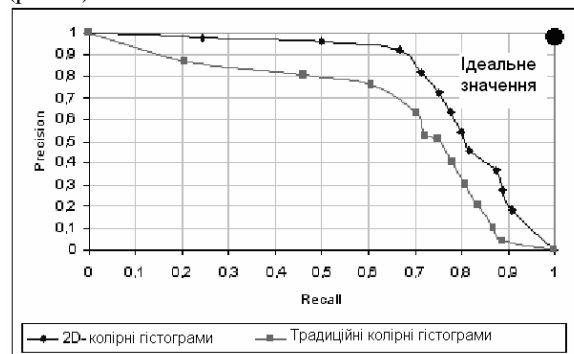


Рисунок 1 – Характеристики якості пошуку

З рисунку видно, що використання 2D – кольорних гістограм дозволяє виконувати пошук більш якісно. Крім того, аналіз зміни цих характеристик, виконаний в [8], показав, що середнє значення характеристики Recall в середньому вна 10% вище при використанні 2D – кольорних гістограм, а середнє значення характеристики Precision в середньому в 2.33 рази перевищує середнє значення цієї ж характеристики для випадку використання традиційних кольорних гістограм.

Пошук зображень з використанням кластеризації

На другому, логічному рівні пошуку в центрі уваги багатьох дослідників лежить задача виділення регіонів зображення. Одним з таких методів є статистичний ієрархічний агломеративний метод кластеризації зображень [9], для якого початковими об'єктами є пікселі, кожен з яких заданий вектором кольорних складових. В ході процедури кластеризації відбувається об'єднання пікселів в окремі групи – кластери (регіони), виходячи з числових значень кольорних характеристик.

Розглянемо метод виділення регіонів зображень, який використовує бітову маску взаємозв'язків і рангів кольорних складових центрів кластерів [10].

Запропоновано визначати маску у вигляді багатовимірного вектора для простору кольорів

RGB, що включає вектора рангів $\bar{s}_R, \bar{s}_G, \bar{s}_B$ і вектора взаємозв'язків $\bar{s}_{GB}, \bar{s}_{RB}, \bar{s}_{RG}$:

$$\bar{S} = \{\bar{s}_R, \bar{s}_G, \bar{s}_B, \bar{s}_{GB}, \bar{s}_{RB}, \bar{s}_{RG}\},$$

причому компоненти векторів можуть приймати тільки два значення: 0 або 1.

Позначимо вектора рангів $\bar{s}_R, \bar{s}_G, \bar{s}_B$ вектором, який можна представити як $\bar{s}_\alpha = (s_{\alpha 1}, s_{\alpha 2}, s_{\alpha 3})$. Компоненти вектора \bar{s}_α визначаються за формулами:

$$s_{\alpha 1} = \begin{cases} 0, \alpha \in [x_l, GH - eps]; \\ 1, \alpha \in [GH - eps, x_h]; \end{cases}$$

$$s_{\alpha 2} = \begin{cases} 0, \alpha \in [x_l, GL - eps) \cup (GH + eps, x_h]; \\ 1, \alpha \in [GL - eps, GH + eps]; \end{cases}$$

$$s_{\alpha 3} = \begin{cases} 0, \alpha \in (GL + eps, x_h]; \\ 1, \alpha \in [x_l, GL + eps]. \end{cases}$$

Тут: $[x_l, x_h]$ – межі зміни числових значень кольорних характеристик (для простору кольорів RGB: $x_l=0, x_h=255$), GL, GH – межі числових значень кольорних характеристик для даних рангів (автором роботи пропонується виділяти три ранги: низький, середній і високий і три, відповідних перерахованим рангам інтервалів: $[x_l, GL]$, $(GL, GH]$, $(GH, x_h]$), eps – параметр методу, введений для можливості декількох рівнів для одного кольорного компоненту і декількох взаємозв'язків для однієї пари компонентів.

Умовно позначимо вектора взаємозв'язків $\bar{s}_{GB}, \bar{s}_{RB}, \bar{s}_{RG}$ вектором $\bar{s}_{\alpha\beta}$, який можна представити у вигляді $\bar{s}_{\alpha\beta} = (s_{\alpha\beta 1}, s_{\alpha\beta 2}, s_{\alpha\beta 3})$.

Компоненти вектора визначаються за формулами:

$$s_{\alpha\beta 1} = \begin{cases} 0, \alpha > \beta; \\ 1, \alpha \leq \beta; \end{cases}$$

$$s_{\alpha\beta 2} = \begin{cases} 0, |\alpha - \beta| > eps; \\ 1, |\alpha - \beta| \leq eps; \end{cases}$$

$$s_{\alpha\beta 3} = \begin{cases} 0, \alpha < \beta; \\ 1, \alpha \geq \beta. \end{cases}$$

Власне метод статистичного ієрархічного агломеративного методу кластеризації зображень полягає в наступному. На першому етапі методу (етапі повного зв'язку) відбувається зменшення

оброблюваного числа кластерів шляхом розподілу в окремі кластери пікселів з однаковими бітовими масками.

Нехай зображення t розміром $[w \times h]$ пікселів визначається як:

$$t = \{p_{jk} = \{r_{jk}, g_{jk}, b_{jk}\} | j \in N, k \in N\}.$$

Тоді формально об'єднання пікселів з однаковими бітовими масками в кластери можна записати в наступному вигляді:

$$\exists m_1, m_2, \dots, m_q : \bar{S}_{j_1 k_1} = \bar{S}_{j_e k_e}$$

$$\forall l, e \leq m_v, l \neq e, v \in [1, q],$$

$$m_v \in [1, w \cdot h], m_v \in N, v \in N.$$

Тут m_v – кількість елементів групи (кластера) з індексом v ; q – кількість кластерів; \bar{S}_{jk} – бітова маска пікселя p_{jk} з координатами (j, k) зображення t ; l, e – індекси пікселів усередині кластера.

Результатом об'єднання пікселів з однаковими бітовими масками буде набір кластерів:

$$A = \{a_1, a_2, \dots, a_q\} : a_v = \cup p_{jk} : \bar{S}_{j_1 k_1} = \bar{S}_{j_e k_e}$$

$$\forall l, e \leq m_v, l \neq e, v \in [1, q], m_v \in [1, w \cdot h],$$

$$m_v \in N, v \in N.$$

На другому етапі методу (етапі одиночного зв'язку) створюємо нові кластери шляхом об'єднання кластерів з мінімальною відстанню. Етап повторюється до тих пір, поки виконується умова "порівняності" кластерів, заснована на бітовій масці взаємозв'язків і рангів [10]. На рис. 2 проведено порівняння пропонованого алгоритму (CIA) з модифікацією методу k-середніх, описаною в [11]. Тут по осі абсцис – ентропія зображень бази даних, по осі ординат – середній час кластеризації зображень для кожної з введених груп. Порівняння проводилося для однакового числа кластерів виділених для кожного з оброблюваних зображень.

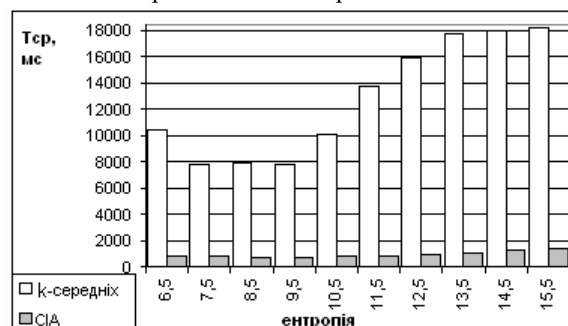


Рисунок 2 - Експериментальна оцінка середнього часу кластеризації зображень CIA і k-means методів

Проведені експерименти з оцінки витрат часу на кластеризацію зображень показали перевагу методу CIA порівняно з методом k-середніх.

Усунення надмірності результатів пошуку за допомогою апарату мультимножин

Початковими даними для пропонованого методу відбору зображень – результатів є відсортований по убутанню близькості до зображення- зразку набір зображень бази даних, які задані характеристиками (колірними, форми, місцеположення, текстури і т.д.). При проведенні експериментів для системи контекстного пошуку в колекції зображень показано [12], що для всіх аналізованих семантичних груп 90% релевантних зображень у відсортованому результуючому наборі знаходяться серед перших 20% результатів запиту.

Далі релевантними зображеннями – результатами пошуку вважаються зображення тематичної групи, до якої належить зображення-зразок.

Хай множина $I = \{I_1, I_2, I_k\}$ – відсортований по убутанню близькості до зображення-зразку набір зображень, множина $R = \{R_1, R_2, R_m\}$ – набір релевантних зображень для певного запиту ($R \in I$), Q – зображення-запит пошуку. Застосування апарату мультимножин для фільтрації нерелевантних зображень дає можливість отримати результуючий набір, кількість зображень в якому p ($p \leq k$).

Докладний опис математичного апарату мультимножин наводиться в [13, 14].

Нехай зображення I_j ($j = \overline{1, n}$) з набору зображень I задається набором характеристик $Q_j = \{q_{j1}, q_{j2}, \dots, q_{jm}\}$, де n – кількість зображень, m – кількість характеристик, за якими проводиться ідентифікація об'єктів – зображень при побудові системи контекстного пошуку.

При побудові запиту до бази даних користувач задає зображення – шаблон пошуку X і кількість зображень – результатів K , що пред'являються користувачу системи контекстного пошуку зображень після обробки запиту.

Для заданого користувачем зображення обчислюється набір характеристик $Q_x = \{q_{x1}, q_{x2}, \dots, q_{xm}\}$ і відбувається сортування зображень цифрової колекції I по убутанню близькості до зображення зразку X . Перші K зображень отриманого набору і є результати.

Окрім кількісних та якісних оцінок множини Q , методика припускає використання двох параметрів r_a і r_b , які показують, чи є зображення I_l ($l = \overline{1, k}$) релевантним (r_a) або нерелевантним (r_b) для зображення запиту X ($r_a + r_b = m$):

- r_{al} – кількість ознак, згідно яким зображення I_l можна вважати релевантним для зображення пошуку X ;

- r_{bl} – кількість ознак, згідно яким зображення I_l можна вважати нерелевантним для зображення пошуку X .

Задачу визначення релевантності зображення I_l ($l = \overline{1, k}$) для зображення пошуку X відповідно до характеристики q_{ls} ($s = \overline{1, m}$), пропонується вирішувати відповідно до наступного правила:

$$r_{als} = \begin{cases} 1, & \text{якщо } (q_{ls} - q_{xs})^2 \leq eps, \\ 0, & \text{інакше.} \end{cases}$$
$$r_{al} = \sum_{s=1}^m r_{als}$$

Параметр eps підбирається експериментально залежно від особливостей характеристик, що розраховуються при індексації зображень, і специфіки бази даних зображень.

Розділення зображень на релевантні X_a і нерелевантні X_b для зображення запиту виконується за наступним принципом: зображення I_l відноситься до класу X_a , якщо, інакше зображення відноситься до класу X_b , для $l \in \{1, k\}$.

Як основний критерій якості контекстного пошуку використовується оцінка характеристики точності, розглянута в [12]. Як результати пошуку розглядалися 10 найближчих до шаблону пошуку зображень. Для оцінки механізму пошуку, модифікованого за допомогою апарату мультимножини, задані наступні числові значення параметрів: $eps=0.1$, $k=20$. Експерименти проводилися з використанням системи контекстного пошуку кластеризованих зображень, описаної в [12], з наступними параметрами: характеристики форми враховувати як другорядні, характеристики місцеположення враховувати, характеристики текстури не враховувати. Як колекція зображень для тестування вибрана база зображень групи дослідників Ванга [15].

Результати експериментів за оцінкою точності контекстного пошуку, які приведені на рис.3, показують, що застосування модифікації механізму пошуку за допомогою апарату мультимножини дає вигоду при оцінці характеристики точності в середньому на 3,24%.

Висновки

Проведені авторами роботи дослідження й експерименти показали доцільність використання запропонованих модифікацій алгоритмів і можливість підвищення ефективності пошуку зображень в базах даних шляхом запровадження цих модифікацій. Використання апарату мультимножин на етапі відбору релевантних зображень дає можливість суттєво зменшити надмірність результуючого набору зображень.

Література

1. Chen C.H., Pau L.F., Wang P.S.P. The Handbook of Pattern Recognition and Computer Vision (2nd Edition). – World Scientific Publishing Co., 1998. – 1004 p.
2. Flickner, M.; Sawhney, H.; Niblack, W.; Ashley, J.; Qian Huang; Dom, B.; Gorkani, M.; Hafner, J.; Lee, D.; Petkovic, D.; Steele, D.; Yanker, P. Query by Image and Video Content: the QBIC System // IEEE Computer. – September, 1995. – p. 23-31.
3. Smith J. R. Integrated Spatial and Feature Image Systems: Retrieval, Analysis and Compression: Submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Graduate School of Arts and Sciences. – Columbia University, 1997. <http://disney.ctr.columbia.edu/jrsthesis/node1.html>
4. Stricker M., Dimai A. Color indexing with weak spatial constraints. In Symposium on Electronic Imaging: Science and Technology//Storage & Retrieval for Image and Video Databases. – 1996. – № 4, vol. 2670. – p. 29– 41.
5. Chua T.-S., Lim S.-K., Pung H.-K. Content-based retrieval of segmented images //In Proceedings of ACM International Conference on Multimedia. – Oct., 1994.
6. Hsu W., Chua T. S., Pung H. K.. An integrated color-spatial approach to content-based image retrieval.//Proceedings of ACM International Conference on Multimedia. – November, 1995 – p.305 – 311.
7. Huang J, Kumar S.R., Mitra M., Wei-Jing Zhu, Zabih R. Image Indexing using color correlograms// Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition. – 1997. – p. 762-768
8. Башков Е.А., Костюкова Н.С. К оценке эффективности поиска изображений с использованием 2d – цветовых гистограмм. Проблемы управления и информатики, №6, 2006. с.84-89
9. Башков Е.А., Вовк О.Л. Оценка эффективности нового статистического иерархического агрегативного алгоритма кластеризации для распознавания регионов изображений // Системні дослідження та інформаційні технології. – Інститут прикладного системного аналізу НАН України, Київ. – 2005. – №2. – С. 117-130.
10. Вовк О.Л. Новый подход к выделению визуально подобных цветов изображений // Проблемы управления и информатики. – 2006. – №6. – С. 100-105.
11. Wang J. Z., Li J. Wiederhold G. SIMPLiCity: Semantics-Sensitive Integrated Matching for Picture Libraries // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 2001. – vol. 23, №9. – P. 947-963.
12. Вовк О.Л. Применение процедуры кластеризации для содержательного поиска изображений // Матеріали I Міжнародної науково-технічної конференції “Моделювання та комп’ютерна графіка”. – ДНТУ, Донецьк. – 2005. – С. 55-58.
13. Петровский А.Б. Упорядочение и классификация объектов с противоречивыми признаками // Новости искусственного интеллекта. – 2003. – №4. – 17 с.
14. Петровский А.Б. Пространства множеств и мультимножеств. – Москва: Едиториал УРСС, 2003. – 248 с.
15. Колекція зображень групи дослідників Ванга [Electronic resource]. – Mode of access: <http://wang.ist.psu.edu/~jwang/test1.tar>

Надійшла до редакції 30.03.2010