

# Пошук кольорових зображень з використанням методів гістограмних ознак і кластеризації

Башков Є.О., Вовк О.Л., Костюкова Н.С., Тележкін Д.В.  
Донецький національний технічний університет  
shozda@r5.donetsk.ua

## Abstract

*Bashkov E.A., Vovk O.L., Kostyukova N.S., Telezhkin D.V. Color image retrieval by histogram features and clasterization. Two image retrieval methods are compared, their joint using for retrieval effectiveness increasing is analyzed.*

## Постановка проблеми

В сучасних умовах спостерігається постійне зростання об'єму інформації, що підлягає обробці. Кількість зображень, цифрових фото та різних картинок постійно зростає, Інтернет та цифрові бібліотеки дають доступ до грандіозної кількості інформації, тобто можна говорити про тенденцію до накопичення зображень і до створення баз даних зображень. Але зі зростанням кількості накопичених зображень і розміру баз даних ускладнюється пошук зображень, які насправді потрібні. Дуже актуальною є задача контекстного пошуку за вмістом зображення.

Існує багато практичних реалізацій контекстного пошуку зображень в графічних базах даних. Найбільш відомими являються системи QBIC (розробка фірми IBM) [1], WebSeek та VisualSeek (розробка Колумбійського університету, США) [2, 3, 4], Virage [6]. Пошук зображень за їх вмістом також реалізовано на сайті Ермітажу у Санкт-Петербурзі (Росія) [5]. Однак, незважаючи на величезну роботу, проведenu для створення цих систем, вони мають ряд недоліків, які на сьогоднішній день досі не усунено. Основною проблемою є те, що на кожний запит система реагує сортуванням всієї БД за зменшенням подібності зі зразком пошуку.

Таким чином, актуальною являється задача дослідження методів контекстного пошуку зображень в графічних базах даних, та їх удосконалення з метою покращення результатів.

Існуючі системи пошуку зображень характеризуються великою різноманітністю з точки зору їх функцій [7].

В загальному вигляді постановка задачі контекстного пошуку зображень сформульована в [8]. Згідно з формулюванням, база даних складається з певної кількості оцифрованих зображень, і вміст кожного описується спеціальною скалярною або векторною величиною – характеристикою вмісту зображення. Порівняння характеристик

здійснюється шляхом обчислення величини, що характеризує ступінь схожості зображень. Таким чином, задача контекстного пошуку зводиться до обчислення ступенів схожості зображень і сортування зображень бази даних за зменшенням їх схожості зі зразком.

Пошук зображень може бути оснований на ряді характеристик його вмісту [7]:

- на кольорових характеристиках зображення;
- на текстурних ознаках;
- на описах об'єктів, з яких складається зображення.

В якості характеристик кольорового вмісту зображень використовуються точечні та гістограмні оцінки. До перших відносяться середній або переважаючий колір точок зображення, до других – кольорові гістограми та бінарні кольорові вектори.

Кольорова гістограма зображення представляє з себе вектор, кожний елемент якого представляє собою нормовану кількість точок зображення, що мають відповідний колір.

При використанні гістограмних ознак для відображення вмісту зображення контекстний пошук виконується у виді послідовності наступних етапів [7]:

- 1) дискретизація ознаки (Q - quantization) – приведення до базового набору кольорів або яскравостей шляхом заміни вхідних даних найбільш близькими значеннями з базового набору;
- 2) побудова гістограмної ознаки (H – histograming);
- 3) порівняння гістограмних ознак (C – comparison). На даному етапі обчислюється характеристика подібності гістограмних ознак всіх зображень бази даних з ознакою зображення-зразку;
- 4) сортування зображень (S - sorting) за зменшенням подібності зі зразком.

При внесенні нового зображення до бази даних для нього послідовно виконуються етапи Q та H. При пошуку, коли в якості зразка використовується зображення з БД, виконуються етапи C та S. Якщо ж зразком пошуку з'являється

нове зображення, виконуються етапи Q, H, C та S

### **Сучасні системи пошуку зображень на основі їх вмісту**

Серед сучасних систем, які вирішують задачу контекстного пошуку зображень, найбільш успішними реалізаціями являються системи QBIC [1], VisualSeek [2,3], Virage [3].

В системі QBIC(розробка фірми IBM) реалізовані такі функції, як заповнення БД зображень, обчислення характеристик зображення та занесення їх в графічну базу даних, обробка запиту (зокрема, формулювання та виконання запиту, а також перегляд результатів. При необхідності запит уточнюється та виконується повторно). Інтерфейс програми включає завдання параметрів запиту та відображення результатів пошуку в порядку зменшення ступеня їх подібності до зразка.

Для представлення обчислених характеристик використовуються вектори, включаючи колір, текстуру та композицію об'єктів зображення, які виділяються вручну. Текстурні та кольорові ознаки також використовуються для опису змісту всього зображення. Параметрами запиту являється процентний склад кольорів, їх розташування, текстура, композиція зображення та його складових об'єктів.

Слід зазначити, що дана технологія використовується для пошуку зображень в цифровій колекції Ермітажу.

Система VisualSeek (розробка Колумбійського університету, США), що виконує контекстний пошук зображень та відео в WWW, дозволяє виконувати запит за кольором, текстурою та взаємному розташуванню областей зображення. Надалі розробники припускають реалізувати пошук за композицією зображення та вбудованому тексту. У даній системі реалізовано автоматичне виділення областей зображення, обчислення та ефективне представлення характеристик, витягання стислих даних. Користувачеві надана можливість виконувати пошук зображення як по локальним ознакам областей, так і по глобальних ознаках зображення. Для запиту першого типу користувач указує розташування закращених областей шляхом позиціонування області на сітці запиту. Зображення, що повертаються, можуть бути використані для глобального запиту. У складі системи можна виділити такі компоненти, як графічний інтерфейс з користувачем, сервер для обробки запитів, сервер пошуку зображень, архів зображень, підсистема індексування зображень.

В VisualSEEK користувач може будувати візуальний зразок запиту з вибором кольорів, текстур, та завданням форми, розміру та просторового розташування областей.

Технологія Virage побудована навколо

ядра, названого Virage Engine, і працює на рівні об'єктів зображень. До складу ядра входять три функціональні частини для аналізу зображень, порівняння зображень і управління [6].

Існують також і інші реалізації контекстного пошуку.

Слід зазначити, що для урахування кольорової інформації, яка міститься в зображенні, в багатьох випадках використовуються гістограмні ознаки і алгоритми кластеризації для розпізнавання регіонів зображень. Ці підходи використовуються кожен окремо, але цікавим є їх сумісне використання з метою покращення якості пошуку. Саме ці питання і розглядаються в статті.

### **Цілі статті і задачі, що вирішуються**

Метою статті є порівняння можливостей двох алгоритмів пошуку, перший з яких базується на порівнянні гістограмних признаков кольорного вмісту зображень шляхом обчислення коефіцієнта кореляції гістограмних ознак [7], а другий використовує для контекстного пошуку статистичні методи кластеризації [9, 10].

Для досягнення вказаної мети вирішуються наступні задачі:

1. Порівняння пошуку зображень в графічній базі даних методами гістограмних ознак та кластеризації з точки зору точності пошуку, що досягається при використанні вказаних підходів.
2. Визначення меж шумів при виділенні кластерів в зображеннях.
3. Формулювання рекомендацій по сумісному використанню методів.

### **Пошук зображень за їх вмістом з використанням гістограмних ознак і статистичних методів кластеризації**

Спочатку розглянемо, яким чином описується вміст зображення при використанні підходів, що розглядаються в статті. Кожне зображення подається в триколіровому просторі RGB і є набором пікселів. Одновимірною кольорова гістограма відображає частоту, з якою зустрічаються пікселі усіх кольорів. Але така гістограма ніяк не показує їх взаємного місця розташування. Тому є смисл використовувати 2D-кольорову гістограму (далі – гістограму) [7].

2D-кольорова гістограма представляє із себе таку ж гістограму з кількістю пікселів зазначених кольорів, однак інтервалом цієї гістограми являються кольори пікселів-сусідів, а кожен її елемент дорівнює кількості пар точок з відповідними кольорами довкола певної точки. Елементи такої гістограми нормуються, за рахунок чого сума елементів гістограми дорівнює одиниці

Якщо говорити про кластеризацію зображень, то вважатимемо, що кожне зображення  $t_i$  має розмір  $w_i \times h_i$  пікселів. Кожному пікселю  $p_{ijk}$  зображення  $t_i$  відповідає три колірних складових  $p_{ijk} = \{r_{ijk}, g_{ijk}, b_{ijk}\}$ , що відображають інтенсивності червоного, зеленого та синього кольорів відповідно, крім того,  $j, k$  відповідають координатам пікселя зображення. Кожне зображення  $t_i$  має набір характеристик  $C_i = \{c_{i1}, c_{i2}, \dots, c_{iM}\}$ . При обчисленні характеристик зображень  $C_i$  на рівні окремих регіонів для пікселів  $p_{ijk}$  аналізованого зображення  $t_i$  спочатку відбувається їх групування в окремі регіони (кластери), сукупність яких і описує зображення.

Першим кроком алгоритму є об'єднання пікселів в кластери. Для цього використовується бітова маска взаємозв'язків та рангів характеристик кольору центрів кластерів, яка складається з двох груп біт, умовно означених як молодші та старші біти [9].

Молодші біти визначаються шляхом аналізу пар колірних характеристик центрів кластерів на наявність відношень типу: менше, більше та дорівнює. При розгляді тривимірною колірною простору маємо 9 молодших біт.

Для визначення старших біт маски увесь діапазон зміни кожної з колірних компонент розділяється на декілька інтервалів, кожний з яких відповідає рівню колірної компоненти. При введенні трьох рівних інтервалів  $[x1 \dots GL]$ ,  $(GL \dots GH]$ ,  $(GH \dots xh]$ , отримуємо три основних рівні: низький, середній та високий, та загалом маємо 9 старших біт. Принцип формування маски взаємозв'язків та рангів показано [9]. Отже, перший крок алгоритму – об'єднання в кластер пікселів з однаковою маскою.

Далі розглянемо порівняння вмісту зображень.

Порівняння вмісту зображень на основі метода гістограмних ознак виконується шляхом обчислення коефіцієнта їх кореляції. Беручи до уваги властивості нормалізованої 2D-кольорової гістограми, формулу для обчислення коефіцієнту кореляції можна значно спростити [7]. Зображення бази даних сортуються за зменшенням коефіцієнту кореляції їх 2D-колірних гістограм, утворюючи таким чином набір результатів пошуку. До набору результатів включаються лише зображення, при порівнянні яких зі зразком отримано значення коефіцієнту кореляції, що перевищують певне порогове значення.

При використанні статистичних методів кластеризації в основі порівняння зображення  $t_0$  та зображень  $t_i$  бази даних лежить відстань  $d_i$  між характеристиками  $C_i = \{c_{i1}, c_{i2}, \dots, c_{iM}\}$  кожного із зображень  $t_i$  характеристиками зображення зразка  $C_0$ . Для обчислення  $d_i$  найбільш часто

використовується середня Евклідова відстань. Набір відстаней упорядковується за зростанням, утворюючи . результати контекстного пошуку – набір зображень, кількість яких задається користувачем системи пошуку.

Практична реалізація передбачає, що пошук методом гістограмних ознак супроводжується зберіганням гістограм на жорсткий диск, створенням гістограм та порівнянням зображень в базі з зображенням – зразком. Пошук методом кластеризації включає розділення зображення на окремі кластери, з модулем видалення кластерів “шуму”. Також присутні модулі підрахунку матриці відстаней та зберігання кластерів на жорсткий диск.

Для перевірки коректної роботи алгоритмів, а також для отримання кількісних та якісних характеристик пошуку необхідна база зображень, на якій звичайно перевіряються інші реалізації пошуку. З цією метою була обрана база Ванга [11,12], що містить 24-бітні зображення у форматі JPEG, розміром 128x96 пікселів. Однак для того, щоб показати якісні характеристики пошуку, звісно необхідні експертні оцінки, бо саме точка зору людини є останньою метою пошуку зображень. З цією метою база даних була зменшена до 350 зображень, зберігаючи свій кількісний состав в пропорції. Зображення були вибрані з вихідної бази Ванга таким чином, щоб кожна семантична група зображень мала свого “представника” в малій базі зображень.

Основні семантичні групи бази Ванга – це метелики, пейзажі, оцифровані фотографії, тварини, астрономічні фото, текстури; людина, прапори/символіка, будівлі та пам'ятники та інші рисунки.

Більш розвернутий кількісний зміст семантичних груп бази Ванга відображено на рисунку 1

Для реалізації алгоритмів пошуку зображень було вибрано середовище розробки Delphi 7. 2-D кольорова гістограма будувалась розміром 256x256 кольорів. З урахуванням деяких особливостей людського зору застосовувались додаткові уточнення для кольорової гамми. Оскільки людське око найбільш чутливе до кольорів, що відповідають довжині хвилі 550-650nm (зелена частина спектру), а найгірше розпізнаються сині кольори, при побудові гістограми більш важливими кольорами вважались червоний та зелений (для зберігання їх складових було виділено по 3 біти, і 2 біти на синій).

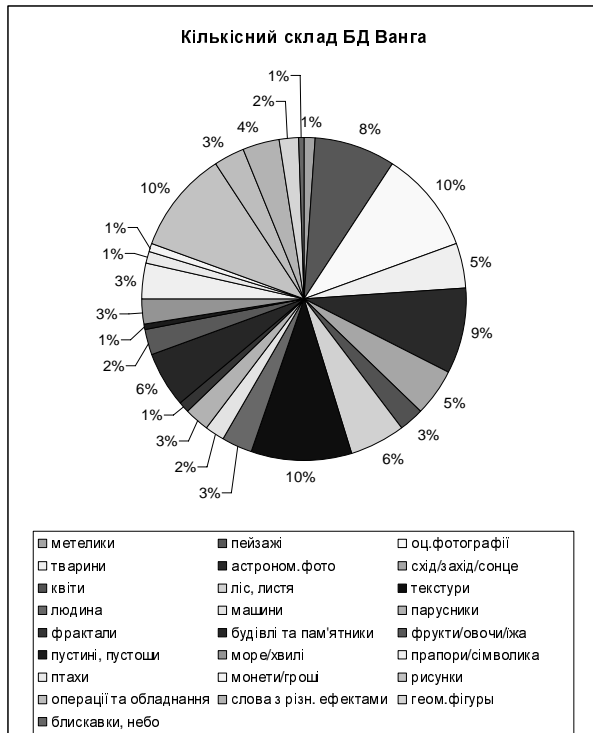


Рисунок 1 – Склад бази Ванга

### Результати, отримані при пошуку зображень

Результати пошуку в різних семантичних групах виявились різними, якість пошуку була обумовлена вмістом зображення – зразка, та вмістом семантичної групи зображень.

При тестуванні роботи алгоритму кластеризації було помічено, що значна кількість запитів при пошуку має велику погрішність. Проаналізувавши алгоритм, було зроблено висновок, що погрішність пов'язана з маленькими кластерами, які можна назвати “шумом”. Суть алгоритму кластеризації, а точніше методика порівняння зображень, полягає в знаходженні мінімуму відстані між кластерами. Але якщо в описі зображення присутні маленькі кластери, що не впливають ні на зміст, ні на семантичну групу зображення, вони продовжують впливати на обчислення мінімумів відстаней. Тому було прийнято рішення не брати до уваги маленькі кластери, які вважаються шумом. В роботі по створенню статистичного ієрархічного агломеративного методу кластеризації [6] було проведено багато експериментів щодо ефективності пошуку методом кластеризації. Згідно з результатами цих експериментів оптимальний розмір кольорового кластеру для найбільш ефективного пошуку лежить у проміжку від 2 до 16.

Була проведена серія практичних тестів алгоритму кластеризації. Виявилось, що, якщо не

приймати до уваги кластери, які містять піксели кількістю менш ніж  $\frac{1}{4}$  від математичного очікування кількості пікселів в кластері, то загальна кількість кластерів у зображенні буде приблизно в рамках “ефективного розбиття” на кластери. Алгоритм було змінено таким чином, щоб кластери, які містять менш ніж  $\frac{1}{4}$  від математичного очікування кількості пікселів у кластері, вважалися шумом. Після цього якість пошуку покращилась. Результати пошуку сортируються за зростанням величини  $d$  – Евклідової відстані між кольоровими характеристиками зображень.

Якість пошуку зображень в графічних базах даних визначається двома основними характеристиками. Перша з них – Precision, або точність пошуку – характеризує долю зображень, схожих на зразок, серед результатів пошуку. Друга характеристика – Recall, або ефективність пошуку – показує, яка доля зображень, схожих на зразок, була включена у склад результатів пошуку [13].

Після серії запитів було побудовано графік вказаних характеристик для методів гістограмних ознак та кластеризації, який приведено на рис.2.

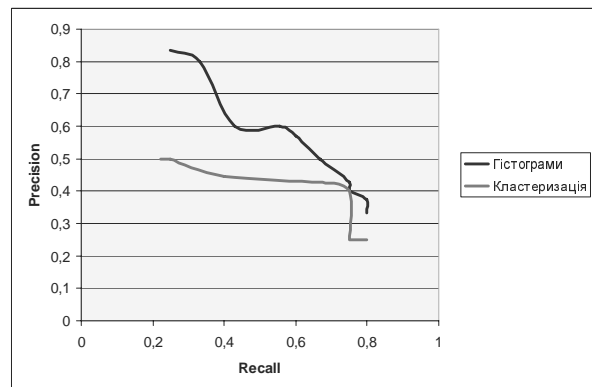


Рисунок 2 – Характеристики пошуку в графічній базі даних

Ці характеристики наявно показують, що в даній реалізації метод гістограмних ознак виявився кращим, ніж метод кластеризації, точніше, його кольорова складова.

Для покращення результатів пошуку можливе сумісне використання двох методів. Тоді один з них має бути приблизним, а інший – точним. Аналізуючи часові характеристики алгоритмів пошуку, можна зробити висновок, що варто проводити приблизний пошук методом гістограмних ознак, а точний – методом кластеризації.

Така точка зору базується на тому, що метод кластеризації працює дуже довго, тоді як метод гістограмних ознак працює швидше і дозволяє визначити приблизну групу зображень, серед яких слід проводити пошук.

## Висновки

В статті були досліджені існуючі методи пошуку зображень в графічних базах даних, вивчені існуючі практичні реалізації систем пошуку зображень в базах даних, їх переваги та недоліки.

При порівнянні можливостей пошуку в графічних базах даних за методами гістограмних ознак та кластеризації встановлено, що якість пошуку методом гістограмних ознак виявилася точнішою за якість пошуку методом кластеризації. Однак обидва методи при пошуку в деяких семантичних групах виявлялися неефективними. Також в роботі проведено практичне дослідження можливостей покращення пошуку методом кластеризації, сформульовано поріг “шумів” в зображенні. Для скорочення часу роботи методу кластеризації ї з метою порівняння з методом гістограмних ознак було взято частину статистичного алгоритму кластеризації, яка порівнює зображення лише за їх кольоровими характеристиками.

## Література

1. IBM Almaden Research Center. Query by Image and Video Content: the QBIC System // IEEE Computer. – September, 1995. – р. 23-31.
2. Smith J. R., Chang S. – F. VisualSEEK: a Content-Based Image/Video Retrieval System: System Report and User's Manual, version 1.0 beta. – December, 1995.
3. Smith J. R., Chang S.-F. Searching for Images and Videos on the World-Wide Web // IEEE Multimedia Magazine. – Summer, 1997. (also Columbia University CU/CTR Technical Report #459-96-25). – Режим доступу: <http://www.ctr.columbia.edu/webseek>
4. Smith J. R., Chang S.-F. VisualSEEK: A Fully Automated Content-Based Image Query System// ACM Multimedia Conference – Boston, MA. – November, 1996. – Режим доступу: <http://www.ctr.columbia.edu/VisualSEEK>
5. Цифровая коллекция Эрмитажа. – Режим доступу: <http://www.hermitagemuseum.org>
6. Bach J R., Filler C ,Gupta A., Namrapur A., Horowitz B., Humphrey R., Jain R. C., Shu C. Virage image search engine: an open framework for image management.// In symposium on electronic imaging: science and technology. Storage & retrieval for image and video databases iv – vol. 2670, IS&T/SPIE, 1996 – р. 76 – 87.
7. Е.А.Башков, Н.С.Костюкова. К оценке эффективности поиска изображений с использованием 2d – цветовых гистограмм. Проблемы управления и информатики, №6, 2006. с.84-89
8. Башков Е.А., Шозда Н.С. Поиск изображений в больших БД с использованием

коэффициента корреляции цветовых гистограмм.// Труды 12-й Международной конференции ГрафиКон'2002. – Нижний Новгород, – 2002. – с. 358–361.

9. Вовк О.Л. Статистичні методи кластеризації для систем контекстного пошуку зображень: Автореферат дисертації на здобуття наукового ступеня кандидата технічних наук: 05.13.06 / Донецький національний університет. – Донецьк, 2006. – 22 с.

10. Башков Е.А., Вовк О.Л. Оценка эффективности нового статистического иерархического агломеративного алгоритма кластеризации для распознавания регионов зображений //Системні дослідження та інформаційні технології. – Інститут прикладного системного аналізу НАН України, Київ. – 2005. – №2 – С.117-130.

11. База Ванга, яка містить 10000 тестових зображень [Електронний ресурс]. – Режим доступу до бази: <http://www-db.stanford.edu/~wangz/image.vary.jpg.tar>

12. База Ванга, яка містить 1000 тестових зображень [Електронний ресурс]. – Режим доступу до бази: <http://wang.ist.psu.edu/~jwang/test1.tar>

13. Форсайт Д., Понс Ж. Компьютерное зрение. Современный подход.: Пер. с англ.–М.:Издательский дом «Вильямс», 2004.–928 с.