

## ІНФОРМАЦІЙНА СИСТЕМА ПОШУКУ ЗОБРАЖЕНЬ В ЕЛЕКТРОННИХ ТЕКСТАХ З ВИКОРИСТАННЯМ ГІБРИДНОГО ПІДХОДУ

Удовенко Сергій<sup>1</sup>, Чала Лариса<sup>2</sup>

<sup>1</sup>Харківський національний економічний університет імені С. Кузнеця

<sup>2</sup> Харківський національний університет радіоелектроніки

### Анотація

*В доповіді розглядається гібридний підхід до створення інформаційної системи автоматичного пошуку зображень в електронних документах. Система передбачає можливість комбінованого аналізу текстових та візуальних характеристик кожного зображення для подальшого пошуку зображень, що відповідають запитам користувачів. Здійснено програмну реалізацію запропонованої системи.*

### Abstract

*The report examines the hybrid approach to creating an information system for automatic image search in electronic documents. The system provides for the combined analysis of the text and visual characteristics of each image to further search for images that meet the user's queries. The program implementation of the proposed system was implemented.*

### Вступ

В даний час отримали розповсюдження цифрові технології, що сприяють стрімкому збільшенню кількості інформації, яка зберігається в базах даних різного тематичного призначення, з метою швидкого доступу до слабо структурованих текстово-графічних електронних документів [1]. Прогрес у методах зберігання даних і технологіях збору зображень дозволив створити великі бази даних з зображеннями, які використовуються в таких сферах, як техніка, торгівля, медицина, розваги, менеджмент, екологічний контроль тощо. Це викликає потребу розробки інформаційних систем для ефективного керування цими базами зображень та оперативного пошуку зображень за запитами користувачів. В доповіді розглядається гібридний підхід до створення такої системи, що передбачала б комбінований аналіз текстових та візуальних змістовних характеристик кожного зображення для подальшого пошуку зображень, що відповідають запитам користувачів пошукової системи.

### Основна частина

Відзначимо, що всі системи пошуку зображень об'єднують необхідність аналізу подібності зображень для подальшого анотування або класифікації. Для реалізації цих завдань зазвичай застосовуються дві основні групи методів: перша з них ґрунтується на текстових метаданих кожного зображення для здійснення пошуку за ключовими словами; друга група ґрунтується на аналізі змістовних характеристик кожного зображення для подальшого пошуку зображень, що відповідають запитам користувачів. Пошук за ключовими словами в більшості випадків забезпечує кращі результати, ніж пошук за змістом в термінах часу відгуку та точності. Однак цей підхід вимагає апріорі наявності анотацій в базах даних зображень, внаслідок чого завдання пошуку є дуже трудомістким і часто суб'єктивним. Пошук за змістовними візуальними характеристиками буває необхідним, коли текстові анотації відсутні або неповні. Крім того, пошук за контентом потенційно може бути більш якісним, навіть якщо текстові анотації попередньо існували, завдяки додатковому інформаційному внеску змісту зображень.

В сучасних системах автоматичного пошуку та аналізу електронних документів часто виникає необхідність оброблення великих обсягів зображень, які можуть з часом

змінюватися. У базі зображень накопичується чимало графічних та текстово-графічних об'єктів, які або вже проіндексовані та проанотовані, або лише знаходяться в черзі на індексацію та анотацію (нещодавні надходження). Оскільки така база не анотована рівномірно, то необхідним є здійснення доступу до графічної інформації через текстові запити. На цей час існують різні методи взаємодії, навігації та пошуку графічних та текстово-графічних об'єктів в базах зображень. Насамперед, це модель короткотермінової взаємодії, що використовується для підвищення точності системи, та модель довготривалої взаємодії, що допомагає пов'язати текстові слова та візуальні характеристики для пошуку зображень за текстом, за візуальним наповненням чи за змішаними характеристиками (текст / зображення). Остання модель пошуку зображень дозволяє ітеративно формувати та уточнювати анотації до зображень.

Визначимо основні особливості гібридної системи пошуку зображень в текстово-графічних документах, яка розглядається у цій доповіді. По-перше, це мультимодальна система пошуку зображень, яка об'єднує різні джерела даних, зокрема, вміст зображення та текст. Ця система дозволяє проводити опитування за зображенням, опитування за ключовим словом або використання гібридних запитів. По-друге, це нова методика зворотного зв'язку, що поєднує класичні методи, які використовуються при пошуку інформації: рух точки запиту і розширення запитів. Ця методика дає хороші результати для інтерактивного пошуку зображень. По-третє, це застосування моделі під назвою «KVR» (Keyword Visual Representation – ключове слово «візуальне представлення») для створення зв'язків між семантичними поняттями та візуальними уявленнями. Завдяки послідовній стратегії навчання ця модель забезпечує зв'язок міжсемантичних понять та візуальних характеристик, що допомагає поліпшити точність анотації на зображення та результати пошуку. Четвертим пунктом є допоміжний алгоритм побудови відповідної бази знань, починаючи з нуля. Слід відзначити, що в запропонованій системі не розділяються етапи побудови анотацій та пошуку. Крім того, також користувач може робити запити відразу після запуску системи, яка буде навчатися в процесі експлуатації.

Система доповнюється інтерфейсом, що дозволяє візуалізувати змішані текстово-графічні запити. Навіть якщо на даний момент використовуються лише окремі типи інформації, тобто текстовий або візуальний, загальна пропонована модель дозволяє її розширити з залученням зовнішніх даних, таких як місцеположення (GPS) та час.

Розглянемо деякі аспекти реалізації запропонованої системи. Як вже було зазначено, база даних зображень запропонованої системи корегується в режимі реального часу. Обсяг зображень цієї бази не є фіксованим, адже інформація про зображення постійно розвивається, в залежності від нових зовнішніх подій, які можуть виникнути в будь-який час. Зазначимо припущення, які є важливими для використання та розвитку нашої системи:

- екземпляри зображень бази не завжди містять повну текстово-графічну інформацію;
- загальна кількість зображень не є фіксованою і обробка нових зображень здійснюється в режимі реального часу;
- база знань системи ґрунтується на анотаціях зображень, доповнених текстом, а також на їх візуальних характеристиках (колір, текстура, форма, просторова диспозиція);
- навчання є інтерактивним і може бути виконане шляхом використання методів навчання з підкріпленням. Взаємодія між користувачами та експертами системи має здійснюватися у простий спосіб для фільтрації не релевантних запитам зображень під час пошуку;
- навчання здійснюється на малих вибірках навчальних даних (до 20 екземплярів).
- система має формувати (аналізувати) анотації зображень в режимі реального часу.

З оглядом на ці гіпотези, в кожний поточний момент можуть розглядатися три типи зображень в пропонованій системі: зображення без анотації; зображення з додатковою

інформацією (наприклад, додатковий опис або довідкові дані); зображення з автоматично доданою інформацією (з «розширеними» анотаціями).

Можливість еволюції бази даних зображень і відповідної бази знань визначила такі сценарії формування запитів:

- сценарій 1: зображення без анотації використовується як запит, спрямований на пошук близьких зображень з бази даних;
- сценарій 2: набір слів в текстово-графічних документах використовується як повноцінний запит;
- сценарій 3: зображення та слова в текстово-графічних документах використовуються як повноцінний запит.

Проблема застосування двох останніх сценаріїв виникає, коли частина бази даних зображень не анотована, що робить цю частину недоступною через текстові запити. У цьому випадку необхідно використовувати зв'язок між текстовими словами та візуальними характеристиками для мультимодального пошуку та анотування зображень, які містять текстову частину.

Мультимодальний пошук зображень – це поєднання візуального пошуку зразків зображень та текстового пошуку за ключовими словами. Візуальна подібність між запитом, пов'язаним з зображенням в запиті, та текстово-графічним зображенням в базі даних оцінюється за величиною скалярного добутку відповідних векторів.

В доповіді запропоновано локальний критерій для визначення текстової подібності між запитом, пов'язаним з зображенням в запиті, та текстово-графічним зображенням в базі даних пошукової системи, що пропонується. Інтегроване значення подібності визначається з застосуванням глобального критерію, який передбачає використання вагових коефіцієнтів для зображень з розширеними анотаціями та для зображень без анотацій.

Структура запропонованої системи передбачає можливість постійної корекції елементів бази зображень з текстовими анотаціями та збільшення її розмірності. До основних блоків системи належать блок «KVR» (для створення зв'язків між семантичними поняттями та візуальними уявленнями), блок «ANOT-IMAGE» (для індексації та анотування текстової частини зображень) та блок навчання з підкріпленням «RL-IMAGE».

На основі запропонованого підходу розроблено програмний модуль інформаційної системи, пошуку зображень за запитом користувачів.

Результатом реалізації запропонованої гібридної схеми може бути формування електронної форми «Індикація можливості плагіату текстово-графічних документів», що створюється для кожного аналізованого об'єкту, з визначенням загальних коефіцієнтів подібності документів, які перевіряються, та побудовою порівняльної гістограми. Запропонована гібридна схема порівняння зображень здійснює перевірку оригінальності текстово-графічного об'єкту, що аналізується, як за джерелами мережі Інтернет, так і за власними базами текстів з зображеннями. Схема передбачає можливість аналізу подібностей та виявлення повного або часткового дублювання зображень.

Отримані результати можуть бути, зокрема, використані для розширення функціональних можливостей існуючих систем виявлення плагіату в наукових текстах, що містять текстово-графічні об'єкти [2].

### Список використаних джерел:

1. Чалая Лариса Ернестовна. Поиск неполных дубликатов в системах анализа цифровых изображений / Л.Э. Чалая, П.Ю. Попаденко // Вісник КрНУ імені Михайла Остроградського. Вип. 5/2014 (88). – 2014. – С.42 – 47.

2. Лізунов Петро Петрович. Автоматичний аналіз подібностей схем та діаграм в електронних текстових документах / П. П. Лізунов, А. О. Білощицький, Л. Е. Чала, С. В. Білощицька, О. Ю. Кучанський, С. Г. Удовенко // Управління розвитком складних систем. – 2016. – № 28. – С. 147 – 156.