

Метод виявлення текстових областей в відеопослідовності з використанням вейвлет-перетворення

А.О. Ніколенко, О.Ю. Бабилунга, Т.Т.К. Нгуєн, В.М. Зайковський

Одеський національний політехнічний університет

anatolyn@ukr.net, babilunga@mail.ru, ktien@ukr.net, vzzzone@gmail.com

Анотація

Досліджено процес виявлення текстових областей на неоднорідному фоні в відеопослідовності. Розглянута двоетапна схема обробки даних відеореєстрації. Запропоновано метод виявлення текстових областей на зображенні на основі безперервного вейвлет-перетворення з ітераційною обробкою зображення на різних масштабах. Досліджено результати реалізації методу виявлення текстових областей на зображеннях, видані рекомендації по автоматичному вибору масштабу вейвлет-перетворення.

1. Вступ

В останні роки, завдяки активному розвитку комп'ютерних засобів обробки інформації та можливостей передачі даних через локальні та глобальну мережі визначилась тенденція постійного зростання обсягів інформації з мультимедійним контентом, що необхідно оперативно обробляти. При розробці інформаційних систем різного призначення, систем пошуку даних в мультимедійних колекціях вирішується задача автоматичного пошуку візуальної інформації, яка відповідає визначеним семантичним ознакам. Одним з різновидів таких даних є текстова інформація (наприклад, номери залізничних вагонів, автомобілів, контейнерів, написи на дорожніх знаках чи інших об'єктах), яка присутня в даних, отриманих за допомогою технічних засобів відеореєстрації, та є джерелом семантичних ознак об'єктів різних рівнів.

При проведенні аналізу (розпізнавання) текстової інформації у відеопослідовностях однією з важливих процедур, яка виконується на першому етапі обробки, є виявлення текстових областей (ТО). Більшість із запропонованих в наукових джерелах методів виявлення ТО на зображеннях використовують як ознаки текстових написів інформацію про колір, перепади інтенсивності і текстурні особливості зображень. Щоб описати ТО і відрізнити їх від фону зображення, розроблено методи і алгоритми виявлення текстових написів, які базуються на евристичних правилах.

Багато з існуючих евристичних методів засновано на однорідності кольору або інтенсивності символів ТО. Вони виявляють області символів в зображенні та групують їх в слова і рядки тексту на основі геометричних обмежень. Ці методи відомі як методи аналізу зв'язаних компонент [1-3]. Вони показують задовільні результати виявлення ТО на зображеннях з простим (однорідним) фоном при відомому кольорі тексту, що не відповідає випадку обробки відеозображень. Крім того, колір у відеозображеннях часто спотворюється в результаті стиснення даних.

Деякі інші евристичні методи виявляють ТО на зображеннях на підставі інформації про перепади інтенсивності. В [4] – вихідне зображення обробляється горизонтальним диференціальним фільтром розміром 3×3 пікселя з подальшою бінаризацією з заданим порогом. В [5] запропоновано метод на основі карти контурів, яка згладжується за допомогою спеціальних фільтрів і піддається морфологічній обробці. В [6] застосовано обчислення значення локального порога для контурного зображення, отриманого за допомогою оператора Собеля. В [7] використовується карта контурів, яка одержана за допомогою детектора контурів Канні, морфологічні операції та аналіз проекцій виявлених областей на осі координат. В [8] замість використання карти контурів для виявлення ТО пропонується використання карти переходів, яка згенерована на підставі зміни інтенсивності і насиченості пікселей зображення. Для перевірки приналежності пікселя зображення до ТО застосовано евристичне правило з використанням різних локальних бінарних образів.

Розглянуті евристичні методи виявлення ТО довели свою ефективність для випадків обробки зображень, на які накладено додаткові обмеження стосовно високої контрастності і відносно однорідного фону. Недоліком цих методів, що значно обмежує їх використання, є необхідність попередньої експериментальної оцінки багатьох параметрів.

Уникнути цього недоліку при розпізнаванні ТО на зображеннях дозволяє використання методу машинного навчання на реальних даних. Однак застосування цього методу приводить до зростання обчислювальної складності алгоритмів обробки зображень. Крім того, останнім часом були запропоновані деякі гібридні підходи до вирішення цієї проблеми, які поєднують переваги застосування евристичних методів пошуку текстових областей (на попередньому етапі обробки зображення) та методу машинного навчання для остаточного прийняття рішення про відповідність визначеної області до класу текстових областей (верифікації). Вищесказане дозволяє зробити висновок, що незважаючи на існування багатьох методів і алгоритмів виявлення текстових областей на зображеннях, до теперішнього часу немає остаточного вирішення цього завдання, тому тема даного дослідження є актуальною.

Метою даної роботи є розробка методу виявлення текстових областей на відеозображеннях з неоднорідним фоном без додаткових обмежень на місце розташування, розмір, колір тексту напису на основі безперервного вейвлет-перетворення.

2. Загальний підхід до виявлення текстових областей на зображеннях

Відеопослідовність розглядається як набір кадрів (окремих зображень). Зображення, що обробляється, може мати ряд особливостей: низьку роздільну здатність, різний розмір шрифтів тексту на одному зображенні, деформацію написів (поворот, розташування напису по дузі, зміну розміру символів від початку до кінця напису), слабкий контраст між текстом і фоном, неоднорідний фон з різкими перепадами інтенсивності, відблиски джерел світла на поверхнях об'єктів, що ускладнює виявлення текстових областей.

Відомо, що для читабельності тексту окремі символи напису повинні мати достатньо чіткі контури. Оскільки, текстові рядки мають «сильні» вертикальні контури, які розташовані в горизонтальному напрямку з високою щільністю, можливо визначити текстову послідовність (в рядку зображення) як квазіперіодичну послідовність імпульсів. Саме ця особливість напису, зазвичай використовується при вирішенні задачі виявлення текстових областей на зображенні. Взагалі, знаходження подібних характерних фрагментів зображення зводиться до виділення, по-перше, границь окремих символів, по-друге, границь ТО в цілому та передачі координат визначеної області для подальшого аналізу (ідентифікації

та класифікації). Тобто результати виконання процедури виявлення ТО безпосередньо впливають на якість роботи системи обробки та розпізнавання зображень в цілому. Якщо на даному етапі був отриманий недостатньо якісний результат, то при подальшому аналізі семантичного складу зображення будуть також з'являтися некоректні результати. Саме це і визначає важливість розробки більш універсального підходу до вирішення цього завдання.

Як правило, для виявлення ТО на зображеннях використовується підхід на базі градієнтних методів обробки значень інтенсивності зображення [7,9], які дозволяють визначити контури окремих символів тексту. На рис. 1 наведено приклад використання детектору контурів Канні (рис. 1, а, б) та просторової фільтрації зображення з заданою квадратною маскою (рис. 1, в) для отримання контурного препарату вихідного зображення. Виявлення ТО на зображеннях на базі градієнтних методів успішно застосовується для демонстраційних зображень, що характеризуються високою контрастністю та мають однорідний фон, та дає суттєво гірші результати на реальних зображеннях зі слабким контрастом між областями інтересу, при наявності завад або неоднорідності фону зображення.

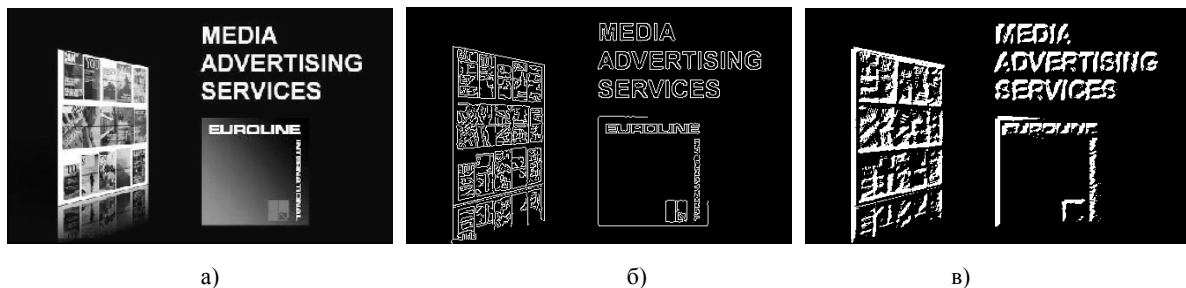


Рисунок 1: Приклад виявлення текстових областей на зображенні з використанням градієнтних методів:
а) початкове зображення, б) результат обробки зображення детектором контурів Канні,
в) контурний препарат після застосування процедури просторової фільтрації

В даній роботі запропонована схема виявлення текстових областей на зображенні з використанням вейвлет-перетворення, що дозволяє обробляти зображення як набір одновимірних сигналів (рядків і стовпців) або проводити аналіз у двовимірному просторі. В даному дослідженні обрана реалізація безперервного вейвлет-перетворення для проведення мультимасштабного аналізу зображення.

3. Виявлення текстових областей на зображенні з використанням вейвлет-перетворення

Відома двоетапна схема виявлення текстових областей в відеопослідовності, що заснована на градієнтних методах обробки зображення, і алгоритми її реалізації [9].

На першому (евристичному) етапі виявлення ТО в відеопослідовності використовуються алгоритми виявлення ТО і текстових рядків (стовпців), які засновані на аналізі перепадів інтенсивності в локальних областях зображення. В результаті чого здійснюється відбір

текстових областей-кандидатів (ТО-К), в яких може бути текстова інформація.

На другому етапі обробки виконується верифікація виявлених ТО-К на відповідність класу текстових послідовностей. Для цього використовуються методи теорії розпізнавання образів, теорії нейронних мереж, нечіткої логіки та інші.

Пропонується удосконалення цієї схеми за рахунок використання на першому етапі її роботи безперервного вейвлет-перетворення зображення, що дозволяє зробити процедуру виявлення ТО на зображенні більш універсальною, тобто незалежною від розміру шрифту і напрямку розташування напису на зображенні, а також фону зображення, що дуже важливо в випадку аналізу відеопослідовності.

Один з перспективних підходів до визначення ТО на зображенні ґрунтується на результатах мультимасштабного контурного аналізу зображення, який може проводитися з різною детальністю в залежності від вимог до інформації, яку необхідно виявити. Для мультимасштабного контурного аналізу звичайно використовують безперервне вейвлет-

перетворення (ВП) [10] з дійсними симетричними непарними вейвлет-функціями, які мають компактний або ефективний носій, зокрема, вейвлети Гауса, базисні функції гіперболічного вейвлет-перетворення та інші [10-12]. Методика автоматичного вибору масштабу на основі аналізу інтегральної функції розподілу енергії за масштабами вейвлет-перетворення (ІРЕМ) запропонована в [12]. Масштаб вибирається як абсциса точки останнього мінімуму (або точки останнього

перегину) кривої ІРЕМ $|E| = f(s)$, де s – масштаб ВП. Якщо мінімум або перегин відсутні в розглянутому діапазоні масштабів, то в цьому рядку немає текстової інформації, і він виключається з подальшого аналізу.

Приклад визначення масштабу вейвлет-перетворення по ІРЕМ та результат виявлення текстових областей з використанням даної методики представлений на рис. 2.

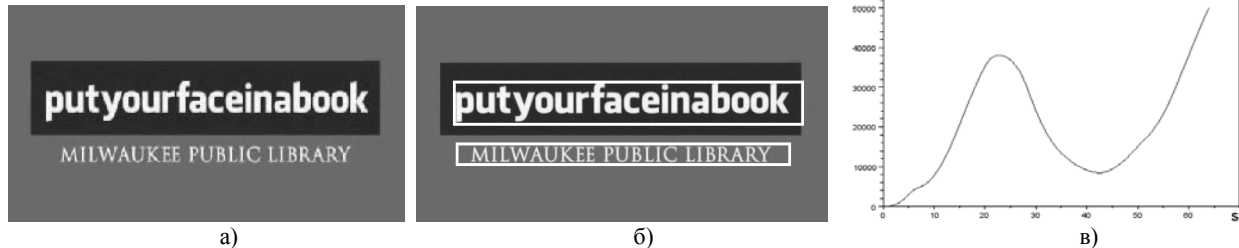


Рисунок 2: Приклад виявлення текстових областей з використанням вейвлет-перетворення Хаара: а) вхідне зображення, б) результат виявлення текстових областей; в) графік ІРЕМ рядку зображення з текстом

Метод виявлення текстових областей на зображенні на основі безперервного вейвлет-перетворення, що використовується на першому етапі виявлення ТО (рис. 3), представлений наступним алгоритмом:

Крок 1. Завантаження вхідного відеозображення.

Крок 2. Перетворення кольорового зображення в градації «сірого».

Крок 3. Нормалізація отриманого зображення з метою розширення діапазону інтенсивності до максимуму.

Крок 4. Попередня обробка зображення: фільтрація зображення від шуму, посилення контрасту або відсікання апріорно непотрібних ділянок зображення для зменшення обсягу оброблюваної інформації).

Крок 5. Вейвлет-перетворення зображення з автоматично вибраним або заданим масштабом:

Крок 5.1. Формування маски розміром $a \times b$, шляхом множення двох векторів (довжиною a і b) значень коефіцієнтів вейвлет-функцій вибраного масштабу ВП.

Крок 5.2. Просторова фільтрація зображення з отриманою маскою.

Крок 5.3. Визначення модулю інтенсивності результату обробки зображення.

Крок 6. Бінаризація зображення по порозу P ($P \sim 0,2$).

Крок 7. Морфологічна обробка зображення: виконання операцій дилатації і розмикання, з метою визначення зв'язаних областей – ТО-К.

Крок 8. Гістограмний аналіз виявлених ТО-К. Побудова проєкцій інтенсивності кожної області на горизонтальну та вертикальну осі, в результаті чого попередньо знайдені текстові області поділяються на декілька менших областей.

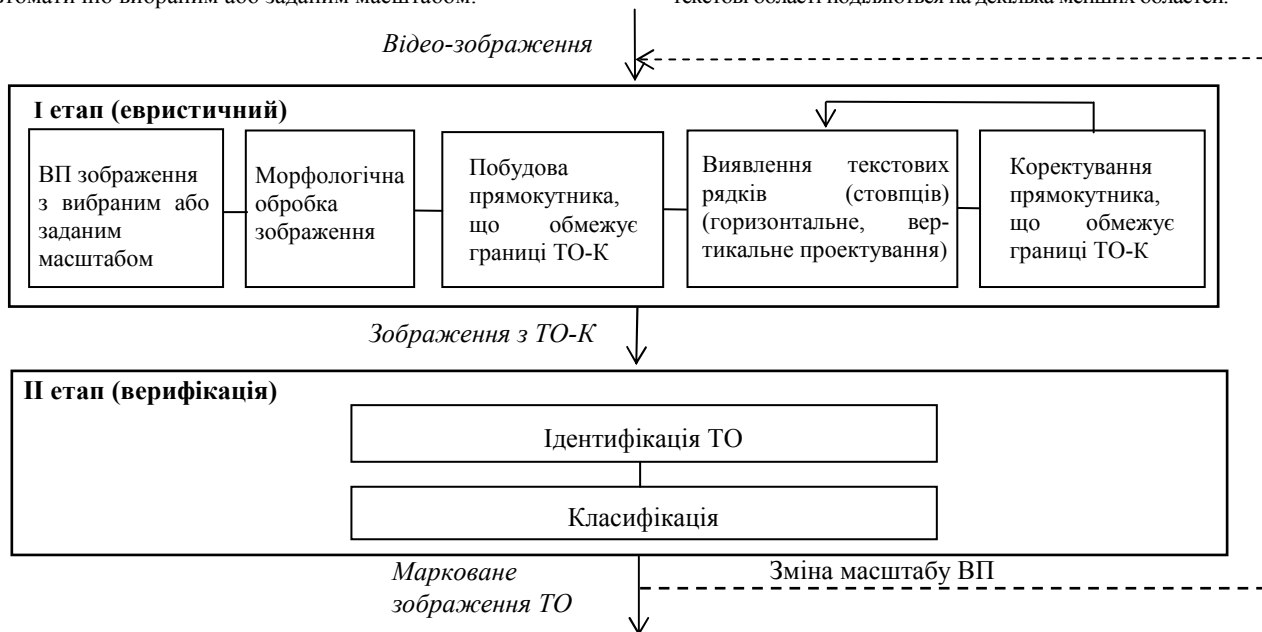


Рисунок 3: Двоетапна схема виявлення текстових областей з використанням вейвлет-перетворення

Крок 9. Аналіз виявлених текстових областей зображення з урахуванням апіорно отриманої інформації про їх властивості.

Крок 10. Маркування областей, які пройшли перевірку і визначені як текстові області, для подальшого аналізу.

Для проведення комп'ютерного експерименту і оцінки якості методу виявлення ТО з використанням безперервного вейвлет-перетворення була виконана програмна реалізація запропонованого алгоритму. Для

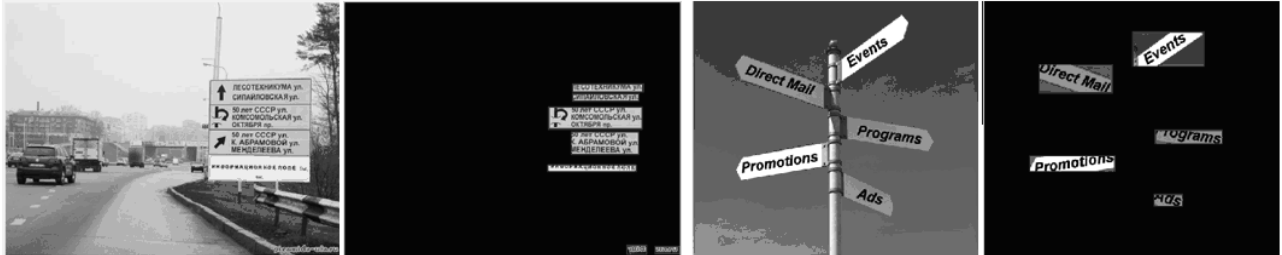


Рисунок 4: Приклади роботи комп'ютерної реалізації методу виявлення текстових областей на відео-зображеннях

4. Висновки

В результаті аналізу існуючих методів виявлення ТО на зображеннях встановлено, що більшість з них базується на евристичних правилах і потребує апіорної оцінки властивостей зображень та параметрів розташування напису, що не дозволяє виявляти ТО на зображеннях з неоднорідним фоном, що швидко змінюється.

Для підвищення універсальності підходу до виявлення ТО на відеозображеннях запропоновано вдосконалення двоступінної схеми виявлення текстових областей у відеопослідовності шляхом введення ітераційної обробки зображення з різними масштабами вейвлет-перетворення.

Розроблено метод виявлення ТО на зображеннях на основі безперервного вейвлет-перетворення з заданим або автоматично вибраним масштабом ВП, що дозволило підвищити точність виявлення ТО на зображеннях з неоднорідним фоном на етапі отримання границь ТО та сприяло підвищенню достовірності роботи системи розпізнавання текстової інформації в цілому.

Комп'ютерна реалізація розробленого методу виявлення ТО показала в середньому на 15 % більше успішних результатів виявлення ТО, ніж градієнтні методи виявлення ТО на зображеннях.

Запропонований метод виділення ТО на відеозображеннях з неоднорідним фоном може бути використано при розробці інформаційних систем різного призначення, систем комп'ютерної обробки і розпізнавання зображень, систем пошуку даних.

5. Література

- [1] Воскресенский М.Е. Метод параметрической оптимизации процесса принятия решений в системах распознавания текстовых меток на видеоизображениях / М.Е. Воскресенский // *Компьютерная оптика.*—2009.— том 33.— №2.— С 202-209.
- [2] Lienhart R. Automatic text segmentation and text recognition for video indexing / R. Lienhart,

тестування були вибрані реальні зображення, на яких присутні текстові області. Зображення мають різні розміри, ступінь освітленості, текстові області на них розташовані довільно на неоднорідному фоні.

На рис. 4 показані результати роботи розробленої програми. Виявлені текстові області визначені на зображенні прямокутниками означеного кольору. Координати областей передаються на другий етап обробки для проведення процедури верифікації ТО-К на приналежність класу текстових послідовностей.

- W. Effelsberg / *Springer Multimedia Systems.* — 2000. — № 8. — P. 69–81.
- [3] Sobottka K. Identification of text on colored book and journal covers / K. Sobottka, H. Bunke, H. Kronenberg // *International Conference on Document Analysis and Recognition.* —1999. — P. 57–63.
- [4] Video OCR for digital news archives / [T. Sato, T. Kanade, E. Hughes, M. Smith] // *IEEE Workshop on Content-Based Access of Image and Video Databases.* — 1998. — P. 52–60.
- [5] A video text detection and recognition system, multimedia and expo / [J. Xi, X.-S. Hua, X.-R. Chen, L. Wenyan, H. Zhang] // *IEEE International Conference.* —2001.—P. 873–876.
- [6] Cai M. A new approach for video text detection / M. Cai, J. Song, M.R. Lyu // *IEEE International Conference on Image Processing.*—2002.— P. 117–120.
- [7] Anthimopoulos M. Multiresolution text detection in video frames / M. Anthimopoulos, B. Gatos, I. Pratikakis // *International Conference on Computer Vision Theory and Applications.* — 2007.— P. 161–166.
- [8] Kim W. A new approach for overlay text detection and extraction from complex video scene / W. Kim, C. Kim // *Image Processing, IEEE Transactions.* — 2009. — № 18 (2).—P. 401– 411.
- [9] Anthimopoulos M. A two-stage scheme for text detection in video images / M. Anthimopoulos, B.Gatos, I.Pratikakis // *Image and Vision Computing.* —2010.— № 28.—P. 1413–1426.
- [10] Малла С. Вейвлеты в обработке сигналов / С. Малла. — М.: Мир, 2005. — 671 с.
- [11] Антошук С.Г. Анализ базисных функций вейвлет-преобразования при многомасштабном контурном представлении изображений / С.Г. Антошук, А.А. Николенко, Е.В. Ткаченко // *Электромашиннобуд. та електрооблад.* — 2009. — Вип. 72. — С. 15 – 19.
- [12] Использование вейвлет-преобразования при локализации последовательностей символов / [С.Г. Антошук, А.А. Николенко, Е.В. Ткаченко, О.Ю. Бабилунга] // *Искусственный интеллект.* — 2009. — № 4. — С. 23 – 29.