

УДК 681.3.07

А.І.КАРДАШ, С.М.ЛЕВИЦЬКА, А.Т.ДУДИКЕВИЧ

Львівський національний університет імені Івана Франка, Львів

**ЗАДАЧА РОЗПІЗНАВАННЯ ЛЮДСЬКИХ ОБЛИЧ МЕТОДАМИ ШТУЧНОГО ІНТЕЛЕКТУ**

**Анотація.** Об'єктом даного дослідження стало виявлення людських облич на цифрових зображеннях. Ми ознайомились з різними підходами до розв'язання таких проблем, а також реалізували та проаналізували роботу одного з власних підходів.

**Ключові слова:** розпізнавання облич, метод Віоли-Джонса, інтегральне представлення зображень, скануюче вікно, ознаки Гаара, бустинг.

**Аннотация.** Объектом данного исследования стало выявление человеческих лиц на цифровых изображениях. Мы ознакомились с разными подходами к решению таких проблем, а также реализовали и проанализировали работу одного из собственных подходов.

**Ключевые слова:** распознавание лиц, метод Виолы-Джонса, интегральное представление изображений, сканированное окно, признак Гаара, бустинг.

**Abstract.** The object of the face recognition on digital images. The authors have studied different approaches to solving such problems and applied and analyzed the work of one of their own approaches.

**Key words:** face recognition, Viola - Jones method, integral representation of images, scanning window, Haar features, boosting.

**Вступ**

Розпізнавання образів є популярною проблемою сучасної інформатики. Образи можуть бути як візуальні так і звукові чи інші, залежно від типів інформації, з якими має працювати система. Багато задач є дуже складним для комп'ютера (хоча й елементарними для людей з використанням вроджених і набутих умінь). Саме тому часто воно відбувається з використанням засобів штучного інтелекту. Оскільки людині (розробнику системи чи науківцю) важко проаналізувати проблемну область і знайти якісь чіткі правила або риси, які однозначно визначатимуть шуканий об'єкт серед шуму, то в більшості випадків використовується навчання з учителем. Будеться система, яка здатна аналізувати дані з предметної області.

Дуже важливим напрямком у розпізнаванні образів є розпізнавання зображень. Оскільки найбільшу частку інформації людина сприймає через зір, то природно, що інтелектуальна система буде певним чином імітувати вміння людини опрацьовувати власне зображення (або відео). Для того, щоб проаналізувати об'єкт, потрібно спочатку знайти його на зображенні, тобто проглянути зображення та розпізнати (виявити) шукані об'єкти.

Розпізнавання зображень особливо важливе для такого полідисциплінарного напрямку, як взаємодія людини з комп'ютером (Human-computer Interaction), над яким активно працює багато корпорацій, які виготовляють цифрову техніку та розробляють програмне забезпечення. Найбільше розробляються системи з наступними задачами розпізнавання: виявлення об'єктів на зображенні (наприклад пішоходів чи автомобілів на дорозі – очевидне застосування у розробці автоматично керованих автомобілів або при аналізі відео з дорожніх камер), виявлення облич (ця функція присутня майже у всіх сучасних фотоапаратах, допомагає в налаштуванні автофокусу, одна з перших дій при отриманні характеристик для розпізнавання емоцій та міміки лица), розпізнавання жестів (для керування комп'ютером без використання мишки). Все це можна застосувати при створенні людиноподібних роботів-помічників для розробки зручних засобів керування ними [1-3].

У даній роботі розглядається розпізнавання людських облич на зображенні (найчастіше це фотографії). Застосуванням розробленої системи може бути, як робота з фотографіями у складі інших програм для виявлення облич на фотографіях і використання інформації про їх розташування в завданнях, які ставляться перед цими програмами (це може бути ідентифікація облич або виділення лица з тла у відеорозмовах для зменшення обсягу передавання даних), так і виділення облич для подальшого розпізнавання емоцій на них.

**Мета та постановка задачі**

Метою роботи було проаналізувати існуючі підходи до розв'язання цієї проблеми та запрограмувати і проаналізувати роботу одного з найкращих методів, а також створити систему, яка знаходить людські обличчя на зображенні. Було розглянуто метод Віоли-Джонса [2,3]. Для роботи алгоритму використовується інтегральне представлення зображення, яке дозволяє швидко порахувати сумарну яскравість ділянки на зображенні. При цьому воно має таку ж розмірність, як і висхідне, але замість кожного пікселя вказана сума яскравостей всіх пікселів, розташованих ліворуч та вище.

Як прості класифікатори, які використовуються для побудови та розпізнання облич складного класифікатора, взято ознаки Гаара, які описують риси зображення, яке потрібно знайти. Прямокутні ознаки Гаара добре працюють з інтегральним представлення зображень, оскільки для їх обчислення достатньо додати кілька елементів матриці зображення.

Для побудови дієвого класифікатора використовується навчальна вибірка прикладів з обличчями та без них. Для аналізу роботи методу Віоли-Джонса випробувано два типи класифікатора: повільний, по-

будований процедурою AdaBoost та швидкий каскадний, який побудований ітеративним алгоритмом, який використовує AdaBoost для навчання каскадів, з яких складається.

### Методи та алгоритми досліджень

Метод Віоли-Джонса був вперше запропонований у 2001 році і завдяки своїй швидкодії швидко став популярним. У літературі можна знайти багато спроб його удосконалення, проте більшість з них спрямована на модифікацію бустингу, тим не менше, такі основні риси, як інтегральне представлення зображень та використання ознак Гаара залишаються незмінними, хіба що спеціально підбираються конкретні ознаки з наміром показати покращення роботи побудованого з них класифікатора.

Однією з сильних сторін підходу є те, що він може бути застосований до пошуку різних об'єктів на зображеннях [4]. Прямокутні (і не тільки) ознаки можна обчислити дуже швидко, зі складністю  $\Theta(1)$ , використовуючи проміжне представлення зображення в інтегральній формі. Таке зображення у точці  $(x, y)$ , на відміну від вихідного зображення, містить не значення пікселя, а суму всіх пікселів, що розташовані не нижче та не правіше від точки  $(x, y)$ :

$$ii(x, y) = \sum_{x' \leq x, y' \leq y} i(x', y'),$$

тут  $ii(x, y)$  – це значення пікселя в інтегральному представленні, а  $i(x', y')$  – значення пікселя у вихідному зображенні. Можна використовувати наступні рекурентні формули:

$$s(x, y) = s(x, y - 1) + i(x, y),$$

$$ii(x, y) = ii(x - 1, y) + s(x, y).$$

Тут  $s(x, y)$  є кумулятивною сумою стовпчика. Умови  $s(x, -1) = 0$  та  $ii(-1, y) = 0$  роблять обчислення скінченними. Таким чином, інтегральне представлення зображення будується за один прохід через вихідне зображення.

Використовуючи інтегральне представлення, сума будь якої прямокутної області може бути обчислена через 4 звертання до масиву. Очевидно, що різниця між сумами двох прямокутних ділянок обчислена за 8 звертань. А ознаки, які складаються з двох суміжних прямокутників, можуть бути обчислені за 6 звертань до масиву, у випадку трьох суміжних прямокутників – за 8, чотирьох – за 9. Завдяки чому, ознаки Гаара знаходяться швидко.

Математично, інтегральне представлення є подвійним інтегралом функції, яка описує зображення (спочатку по рядках, а потім по стовпцях).

Неважко побачити, що маючи розширення скануючого вікна  $24 \times 24$ , повна множина різних прямокутних ознак є дуже великою: 45396 елементів. Зауважимо що на відміну від базису Гаара, така множина не буде лінійно незалежною. Тут розмірність підпростору буде 576, що набагато менше ніж в [3].

Фройнд і Шапір [5] довели, що навчальна помилка сильного класифікатора прямує до нуля експоненційно по кількості ітерацій. Ще більш важливо, в інших дослідженнях згодом було продемонстровано та обгрунтовано добрі узагальнюючі здатності даного підходу. Ключовим моментом тут є те, що узагальнююча здатність залежить від границі для набору прикладів, а AdaBoost досягає великих границь досить швидко.

На рисунку 1 показано прямокутні ознаки відносно скануючого вікна. Ознаки з двох прямокутників наведені в (A) і (B). Малюнок (C) показує ознаку з трьох прямокутників, і (D) – ознака з чотирьох прямокутників. Справа – решта ознак.

Звичайна процедура AdaBoost може бути легко інтерпретована як жадібний процес вибору ознак. Розглянемо загальну задачу бустингу, в якій великий набір функцій класифікації об'єднуються за допомогою функції зваженого голосування. Проблема полягає в тому, щоб пов'язати великі ваги з добрими функціями класифікації, а малі з поганими. AdaBoost – це агресивний механізм відбору невеликого набору гарних функцій класифікації, які, тим не менш, сильно відрізняються між собою. Проводячи аналогію між слабкими класифікаторами та ознаками, AdaBoost є ефективною процедурою для пошуку з невеликого числа "хороших", які тим не менш, мають значне розмаїття.

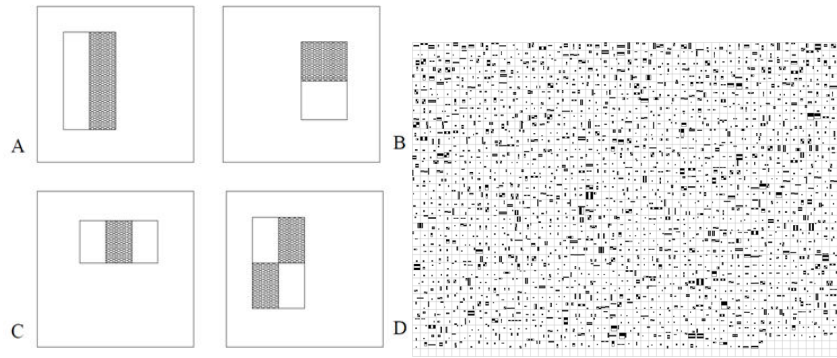


Рисунок 1– Приклади прямокутних ознак

Розглянемо типовий алгоритм навчання із застосуванням бустингу. Будуються гіпотези, кожна з яких використовує одну ознаку. Остаточна гіпотеза є зваженою лінійною комбінацією з гіпотез, де ваги обернено пропорційні помилкам навчання [3, 6].

1. Нехай дано навчальні приклади зображень. Подамо їх парами  $(x_1, y_1), (x_2, y_2) \dots (x_n, y_n)$ , де  $y_i = 0, 1$  для негативних та позитивних прикладів відповідно.
2. Ініціалізуємо ваги наступним чином:  $w_{1,i} = 1/(2m)$  або  $w_{1,i} = 1/(2l)$  для  $y_i = 0, 1$  відповідно. Тут  $m$  та  $l$  – це кількість негативних та позитивних прикладів у навчальній множині.
3. Для кожної гіпотези  $t = 1, \dots, T$  виконати наступне:

- a. Нормалізувати ваги:  $w_{t,i} = w_{t,i} / \sum_{j=1}^n w_{t,j}$ .

- b. Для кожної ознаки  $j$  навчати класифікатор  $h_j$ , який обмежений лише одною ознакою.

Помилка оцінюється для  $w_t$   $e_j = \sum_{i=1}^n w_j |h_j(x_i) - y_i|$ .

- c. Вибрати класифікатор  $h_t$  з найменшим значенням помилки  $e_t$ .

- d. Оновити ваги:  $w_{t+1,i} = w_{t,i} \beta_t^{1-e_i}$ , де  $\beta_t = e_t / (1 - e_t)$ .

4. Остаточний сильний класифікатор матиме вигляд:

$$h(x) = \begin{cases} 1, & \text{якщо } \sum_{t=1}^T a_t h_t(x) \geq \frac{1}{2} \sum_{t=1}^T a_t, \\ 0, & \text{інакше} \end{cases}$$

де  $a_t = \log(1/\beta_t)$ .

### Висновки

Результатом виконаної роботи є створення працюючої системи для розпізнавання обличчя на зображенні з використання методу Віоли-Джонса та з модифікацією методу сканування зображення.

Отримана система працює швидко та може опрацювати близько 15 зображень розміром  $400 \times 300$  пікселів за секунду на звичайному комп'ютері.

Як результат проведеного дослідження з двома способами побудови класифікатора напрашується висновок, що адаптивні каскадні класифікатори є кращими для розпізнавання обличчя, оскільки майже без втрат в ефективності досягається набагато вища швидкість роботи. Також вони є кращими порівняно із

перцептронами, бо в процесі їх роботи задіюються лише ті частини системи, які виконують корисну роботу.

Побудована система може слугувати основою для використання у складніших задачах аналізу обличчя, як перший етап знаходження обличчя, а зважаючи на швидкодію, може працювати з відео.

На перспективу можна, використовуючи такі ж алгоритми побудови класифікаторів, побудувати систему для виділення частин лица та аналізу емоцій.

#### Список літератури

1. Стюарт Рассел. Искусственный интеллект: современный подход, 2-е изд./ Стюарт Рассел, Питер Норвіг; пер.с англ. – М., Вільямс, 2006 - 1408 с.
  2. P. Viola. Rapid Object Detection using a Boosted Cascade of Simple Features,/ P. Viola and M.J. Jones, proceedings IEEE Conf. on Computer Vision and Pattern Recognition (CVPR 2001), 2001.
  3. P. Viola. Robust real-time face detection,/ P. Viola and M.J. Jones, International Journal of Computer Vision, vol. 57, no. 2, 2004., pp.137–154.
  4. Stan Z. Li (Editor). Handbook of Face Recognition. Springer,/ Stan Z. Li (Editor), Anil K. Jain (Editor), 2004.
  5. Yoav Freund. A decision-theoretic generalization of online learning and an application to boosting./ Yoav Freund and Robert E. Schapire. In Computational Learning Theory: Eurocolt '95, pages 23–37. Springer-Verlag, 1995.
  6. Y.Freund. A short introduction to boosting./ Y.Freund, R.E.Schapire. Journal of Japanese Society for Artificial Intelligence 14, 1999 , pp. 771–780.
- Стаття надійшла: 12.02.2013.

#### Відомості про авторів

**Кардаш Андрій Іванович** – кандидат фіз.- мат. наук, доцент кафедри програмування, Львівський національний університет імені Івана Франка.

**Левицька Софія Михайлівна** – старший викладач кафедри програмування, Львівський національний університет імені Івана Франка.

**Дудикевич Анна Теодорівна** – кандидат фіз.- мат. наук, доцент кафедри обчислювальної математики, Львівський національний університет імені Івана Франка.