

ЕТАЛОННО-ЗАЛЕЖНЕ РОЗПІЗНАВАННЯ СИМВОЛІВ У ПРОМИСЛОВИХ СИСТЕМАХ МАШИННОГО ЗОРУ

© Олещук О.В., Попель О.Є., Зацолкін К.В., 2010

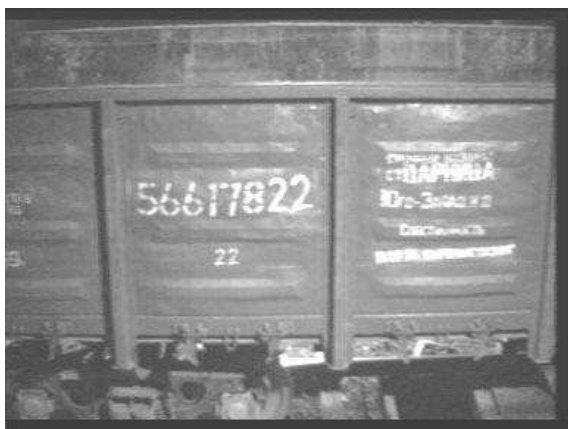
Розглядаються проблеми розпізнавання текстової інформації в промислових системах машинного зору. Пропонується метод розпізнавання, об'єднуючий етап безпосереднього розпізнавання з попередньою обробкою зображень. Оцінено швидкодію стандартного підходу розпізнавання і методу еталонно-залежного розпізнавання. Запропоновані конкретні параметри для ідентифікації символів у предметній області, пов'язаній з розпізнаванням номерів залізничних вагонів.

Ключові слова – обробка зображень, розпізнавання символів, машинний зір.

Problems of recognition of the text information in industrial systems of computer vision are considered. The method of recognition joining a stage of direct recognition with preprocessing of images is offered. The estimation is given speed of the standard approach to recognition and a method of standard-dependent recognition. Concrete parameters to identification of symbols in the data domain connected to recognition of numbers of railway cars are offered.

Keywords – image processing, character recognition, computer vision.

Вступ. В останні десятиріччя системи машинного зору поширені у багатьох сферах промисловості. Не є винятком і транспортна промисловість, де однією з важливих задач є отримання інформації про поточне місцезнаходження та напрямку руху транспортних засобів. Характеристикою, що дає змогу однозначно ідентифікувати транспортний засіб при перевезеннях між підприємствами, є державний номер транспортного засобу (рис. 1а), а при перевезеннях у межах одного підприємства такою характеристикою може бути державний або внутрішній номер (рис. 1, б).



а



б

Рис. 1. Нумерація залізничних вагонів:
а – вагон Укрзалізниці; б – внутрішній вагон підприємства

Сьогодні задача обліку вантажопотоків вирішується шляхом ручного введення даних про транспортні засоби. При цьому людський чинник є однією з перешкод для збільшення швидкості проходження через пункт обліку. Один з перспективних напрямів розв'язання задачі автоматизованої ідентифікації залізничного транспорту [1] заснований на використанні візуальної інфор-

мації про вагони та її обробки з метою локалізації зони розташування номера і його розпізнавання. Ця задача складається з багатьох підзадач [2], серед яких особливо слід виділити розмежування вагонів, пошук та локалізацію написів, розподіл написів на символи та їх розпізнавання. Нижче розглянуто задачу розпізнавання окремих символів.

Аналіз загальноприйнятого підходу до розпізнавання. За загальноприйнятим підходом [3] обирається метрика для оцінювання відстані між еталонним символом і зразком, що підлягає розпізнаванню, обчислюється відстань від зразка до всіх еталонів і ухвалюється рішення про приналежність до того класу, до якого належить символ з числа еталонів, що знаходиться на мінімальній відстані від зразка.

Іншими словами, за загальноприйнятим підходом ставиться завдання розглянути, символ якого класу присутній на зображенні.

Потрібно оцінити витрати обчислювальних ресурсів на реалізацію кожного підходу. У загальному випадку сумарні обчислювальні витрати при обробці зображення стандартними методами визначаються як

$$S_i = N \times M, \quad (1)$$

а об'єм пам'яті V_i для зберігання такої матриці залежить від кількості елементів матриці: $V_i = \Theta(S_i) = \Theta(N \times M)$, де функція Θ – асимптотична форма запису і використовується для спрощення аналізу та запису темпу росту досліджуваної функції [5].

$V_i = \Theta(N \times M)$ тоді і тільки тоді, коли існують додатні константи c_1 та c_2 та n_0 такі, що $c_1(N \times M) \leq V_i \leq c_2(N \times M)$ при $N \geq n_0$. Функції V_i та $(N \times M)$ зростають з однаковою швидкістю при всіх $N \geq n_0$.

Незважаючи на, як здавалося б, обчислювальну ефективність стандартного підходу, слід зазначити, що втрати інформації на етапі попереднього перетворення негативно позначаються на подальших етапах. У ряді випадків фільтрація навіть при неоптимальному виборі параметрів дає змогу отримати спотворений, але впізнаний символ. Щоб його розпізнати, необхідно додавати зразки, що відрізняються від ідеального написання, у множину еталонних символів. Отже, збільшуються обчислювальні витрати на зіставлення розпізнаваного зразка з еталоном. Наступним чинником, що підвищує обчислювальні витрати при стандартному підході, є значне спотворення початкового сигналу при неправильних настройках попередньої обробки, з отриманням символу, що не може бути розпізнаний. У зв'язку з цим необхідно виконувати повторне розпізнавання при іншому наборі значень параметрів фільтрації, обробляти додаткові джерела інформації, наприклад, розглядати сусідні кадри відеофайла, а також застосовувати методи контролю за якістю розпізнавання.

Мета роботи. Метою статті є дослідження еталонно-залежного методу розпізнавання, призначеного для використання у промислових системах розпізнавання як альтернативи стандартним підходам аналізу зображень.

Розпізнавання символів еталонно-залежним методом. У задачах розпізнавання текстової інформації на зображення, що підлягає обробці, зазвичай впливає ряд перешкод, що найяскравіше проявляється у задачах промислового розпізнавання. У таких випадках безпосередньо розпізнаванню передують етап попередньої обробки зображень [4], що безумовно виправдане, коли точно відомо, якого роду перешкоди присутні на конкретному зображенні і який їхній внесок у спотворення. В інших випадках фільтрація може привести до таких змін зображення, що початковий символ стане невпізнаний. Але, щоб оцінити перешкоди, необхідно мати уявлення про ідеальний сигнал, тобто знати, який символ міститься на неспотвореному зображенні, але ця інформація відсутня, поки не ухвалено рішення про приналежність символу до одного з класів $C = \{c_1, c_2, \dots, c_n\}$, а будь-яке ухвалене рішення не може вважатися достовірним, поки не усунуто вплив перешкод.

Нехай є завдання, зворотне дл розпізнавання символів, коли відомо, що ідеальне зображення є еталон e_j , що належить класу c_l . Потрібно дати кількісну оцінку одному або декільком параметрам множини функцій, що спотворюють, $X = \{\chi_1, \chi_2, \dots, \chi_{N_\chi}\}$. Параметри, що представляють інтерес, залежать від виду функції $\chi(m, n)$, де m, n – горизонтальна та вертикальна координати відповідно. Як оцінювані параметри можна трактувати амплітуду, частоту, кількість пікселів, де $\chi(m, n)$ не дорівнює нулю та ін.

У загальному випадку на зображення може впливати декілька спотворюючих функцій, тоді можна записати, що сумарні витрати на спотворення визначаються як

$$K_\chi = \sum_{i=1}^{N_\chi} k_{\chi,i} f_{\chi,i}(P_{\chi,i,1}, P_{\chi,i,2}, \dots, P_{\chi,i,q_{\chi,i}}), \quad (1)$$

де $P_{\chi,i,j}$ – j -й оцінюваний параметр i -ї спотворюючої функції; $k_{\chi,i}$ – ваговий коефіцієнт врахування i -ї спотворюючої функції; $q_{\chi,i}$ – кількість оцінюваних параметрів i -ї спотворюючої функції; $f_{\chi,i}$ – функціональна залежність всіх оцінюваних параметрів, i -ї спотворюючої функції.

Вид функції $f_{\chi,i}$ слід визначати експериментально або аналітично у кожному окремому випадку залежно від впливаючих на зображення спотворюючих функцій та їх вибраних параметрів, що підлягають оцінюванню.

Значення вагових коефіцієнтів $k_{\chi,i}$ вибираються експертом, враховуючи ступінь впливу i -ї спотворюючої функції.

У спрощеному випадку, коли для кожної спотворюючої функції вибирається тільки один визначальний параметр, вираз (1) можна записати у вигляді

$$K_\chi = \sum_{i=1}^{N_\chi} k_{\chi,i} P_{\chi,i}. \quad (2)$$

У цьому випадку на параметри $P_{\chi,i}$ накладаються такі обмеження:

- визначальний параметр повинен бути ненегативною величиною;
- при зростанні впливу i -ї спотворюючої функції значення визначального параметра $P_{\chi,i}$

повинно збільшуватися.

Тоді розв'язок поставленої задачі може бути зведено до пошуку такого набору значень параметрів $P_{\chi,i}$, при яких значення сумарних витрат на спотворення K_χ буде мінімальним. Значення величини K_χ^l визначатиме мінімально необхідні витрати на перетворення спостережуваного зображення до еталону e_l або від цього еталона до спостережуваного зображення.

За пропонованим підходом спочатку висловлюється припущення, що на зображенні був присутній еталонний зразок e_0 , на який діяли перешкоди. Ставиться завдання визначити послідовність впливів на еталон, які в результаті дадуть спостережуване зображення. Витрати на спотворення є аналогом відстані від еталонного зразка до спостережуваного. Потім ті самі дії повторюються для решти всіх еталонів. В отриманому в результаті наборі витрат на спотворення знаходиться мінімальне значення K_χ^l і ухвалюється рішення про приналежність зразка до того класу, до якого належить еталон e_l на підставі того, що витрати на приведення зразка до цього еталону мінімальні.

Оцінка обчислювальних витрат при застосуванні еталонно-залежного розпізнавання. У методі еталонно-залежної обробки початкове зображення безпосередньо бере участь в обробці під кожен еталон окремо.

У зіставленні перетвореного зразка з еталонами, як правило, немає необхідності, оскільки в результаті дії спотворюючих функцій $X = \{\chi_1, \chi_2, \dots, \chi_{N_\chi}\}$ на початкове зображення повинен бути отриманий символ, який збігається з еталонним. Виняток становлять випадки, коли набір спотворюючих функцій є недостатнім для приведення довільного зображення до заданого вигляду. Тоді при ухваленні рішення про приналежність символу до одного з класів повинні одночасно використовуватися результати зіставлення з еталоном і значення сумарних спотворюючих витрат (2). З іншого боку завжди можна розширити множину спотворюючих функцій для отримання перетворень довільного вигляду. Тоді сумарні обчислювальні витрати визначатимуться як

$$Z_e = z_\chi N_e + z_2, \quad (3)$$

де z_χ – обчислювальні витрати на підбір параметрів спотворюючих функцій для одного еталона.

Саме у z_χ зосереджена вся обчислювальна складність методу, а якщо врахувати, що підбір проводиться для кожного еталона, то очевидною стає вимога до набору спотворюючих функцій і способу підбору їх параметрів: спотворюючі функції по можливості повинні бути обчислювально прості, а при підборі параметрів перевагу слід віддавати аналітичним методам з використанням адитивних операцій.

Враховуючи, що підібрана послідовність дій на початкове зображення при заданому наборі спотворюючих функцій завжди дає змогу виключити вплив перешкод, то інформація, присутня на початковому зображенні, використовується з максимальною ефективністю, для отримання бажаної точності розпізнавання за меншої кількості еталонів порівняно із стандартним підходом.

У загальному випадку, еталонно-залежною обробкою можна досягти кращих результатів розпізнавання, але це вимагає великих обчислювальних витрат, тому має сенс розглянути комбінований підхід, коли частина спотворень усувається без прив'язки до еталонів, а частина – за еталонно-залежним методом.

Які з перетворень слід виконувати в прив'язці до еталонів, а які, ґрунтуючись на статистичних даних і додаткових джерелах інформації, залежать від конкретної предметної області і класів впливаючих перешкод? Наприклад, масштабування символів може бути винесене в розряд не залежних від еталону за рахунок визначення середньої висоти символів, що входять в напис, який підлягає розпізнаванню. Аналогічно, має сенс заздалегідь фільтрувати шум з частотою 50 Гц, що виникає за рахунок перешкод за живленням, оскільки за статистикою така перешкода присутня, і дія її з часом майже не змінюється.

Слід зазначити, що, якщо початкове зображення практично не схильне до перешкод, не потрібні попередні його перетворення і як спотворюючу функцію вибрано функцію зіставлення з еталоном, то метод еталонно-залежної обробки збігається із стандартним.

Еталонно-залежне розпізнавання номерів залізничних вагонів. На прикладі написів на бортах вагонів буде розглянуто, яких спотворень зазнає зображення в більшості предметних областей, пов'язаних із відеоспостереженням в умовах природного освітлення.

Треба враховувати, що по можливості слід мінімізувати набір функцій, що спотворюють, але пам'ятати, що вони повинні відображати усі можливі спотворення, яких зазнає розпізнаване зображення від моменту написання текстової інформації до розпізнавання зображених символів.

Наприклад, можливе відхилення від еталонного написання символу за рахунок його повороту (рис. 2, а), а натомість до складу функцій, що спотворюють, вводиться масштабування. З незначною погрешністю масштабування може компенсувати поворот на невеликий кут, але в загальному випадку результат розпізнавання не може вважатися достовірним.

У тривіальному випадку, коли зображення на спостережуваному об'єкті повністю відповідає еталону, всі спотворення вносяться на етапі отримання і оцифровки зображення. За рахунок нерівномірного освітлення (рис. 2, б) з наперед невідомою інтенсивністю, а також неконтрольованої функції гамма-корекції на квантованому зображенні рівні чорного і білого кольорів будуть невідомі, а отже, в загальному випадку не відповідатимуть аналогічним рівням в еталоні.

За рахунок дискретизації за просторовими координатами проводитиметься усереднювання за деякою областю простору. За рахунок цього різко окреслені переходи символ–фон будуть згладжені (рис. 2, в), тобто в прикордонних областях спостерігатимуться пікселі з проміжним значенням яскравості між рівнями чорного і білого. Аналогічний ефект спостерігається за рахунок незначного розфокусування об'єктиву.



Рис. 2. Характерні класи перешкод:

а – поворот символів на невеликий кут; б – нерівномірна освітленість; в – згладжені переходи між символами та фоном; г – темні символи на світлому фоні; д – написи, що давно не оновлювалися

Ще один з видів спотворень – написання темних символів на світлому фоні або світлих на темному (рис. 2, г), що умовно можна назвати ознакою позитив/негатив. Не будучи в звичайному розумінні перешкодою, він все ж таки може бути віднесений до розряду функцій, що спотворюють, оскільки ця ознака може бути відновлена за допомогою еталонно-залежних методів обробки. Єдина його відмінність в тому, що ця функція, що спотворює, буде врахована з нульовим ваговим коефіцієнтом, що означає інваріантність до такого роду спотворень.

Характерною також є наявність написів, які давно не оновлювалися і фарба на яких частково стерлася (рис. 2, д), або забруднених номерів. У подібних відеосигналах значення окремих одиночних пікселів значно відрізнятимуться від еталонів.

Описаний набір перешкод, який характерний для більшості типів зображень, назвемо базовим. Для нього нижче розглянемо можливість аналітичного підбору оптимальних параметрів.

Хай існує зразок $b(m, n)$, що підлягає розпізнаванню і еталони $e_l(m, n)$, де $0 \leq l \leq L-1$ й $l \in Z$, під який необхідно підібрати оптимальні параметри перетворення зразка. Приймається, що еталонний рівень чорного кольору дорівнює одиниці, а еталонний рівень білого – нулю.

Припустимо, що спочатку зразок $b(m, n)$ був еталонним зображенням символу $e_l(m, n)$, вважаємо, що для всіх m та n , де $e_l(m, n) = 1$, піксель розпізнаваного зразка повинен бути чорним, а для m та n , де $e_l(m, n) = 0$, піксель зразка білий. Тоді оптимальне значення рівнів чорного і білого визначатиметься, наприклад, простим усереднюванням.

$$U_b = \frac{\sum_{m=0}^{M-1} \sum_{n=0}^{N-1} b(m, n) e(m, n)}{\sum_{m=0}^{M-1} \sum_{n=0}^{N-1} e(m, n)}, \quad (4)$$

$$U_w = \frac{\sum_{m=0}^{M-1} \sum_{n=0}^{N-1} b(m, n) (1 - e(m, n))}{\sum_{m=0}^{M-1} \sum_{n=0}^{N-1} (1 - e(m, n))}, \quad (5)$$

де U_b – оптимальний рівень чорного; U_w – оптимальний рівень білого.

Як видно, тут замість поняття поріг бінаризації використовуються поняття рівнів яскравості, відповідних чорному і білому кольорам, причому окремим пікселям може відповідати колір, від рівня якого код пікселя розташований на більшій відстані. Подібна ситуація часто існує для одиночних викидів і у випадках, коли цей зразок містить не той символ, під який зараз його підганяють. Тобто відмінність зразка від еталону сприймається як значні відхилення рівня яскравості.

Використовуючи вирази (4)–(5), значення оцінюваних параметрів спотворюючих функцій визначається як

$$P_{\chi,1} = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} |(U_b - b(m, n)) \cdot e_l(m, n) + (U_w - b(m, n)) \cdot (1 - e_l(m, n))|, \quad (6)$$

$$P_{\chi,2} = |U_b - U_w|, \quad (7)$$

$$P_{\chi,3} = \begin{cases} u_p, & \text{при } U_b > U_w \\ u_n, & \text{при } U_b \leq U_w \end{cases}, \quad (8)$$

де $P_{\chi,1}$ – параметр спотворення бінаризації; $P_{\chi,2}$ – параметр спотворення діапазону яскравості; $P_{\chi,3}$ – параметр ознаки позитив/негатив; u_p – значення параметра $P_{\chi,3}$ при позитивному зображенні; u_n – значення параметра $P_{\chi,3}$ при негативному зображенні.

Значення величин u_p та u_n , як правило, вибираються довільно, якщо ваговий коефіцієнт для ознаки позитив/негатив дорівнює нулю. У інших випадках, наприклад, якщо очікуються тільки темні символи на світлому фоні, то значення параметра u_p може дорівнювати нулю, а значення u_n вибирають таким, щоб виконувалася умова $u_n > u_p$. Ваговий коефіцієнт при $P_{\chi,3}$ вибирається чималим, щоб внесок спотворення на інвертування зображення гарантовано переважував інші спотворення.

Значення параметра $P_{\chi,3}$ може бути корисно навіть при нульовому ваговому коефіцієнті, щоб переконатися, що всі символи в написі набули однакового значення ознаки позитив/негатив.

Ваговий коефіцієнт при параметрі спотворення діапазону яскравості вибирають, враховуючи особливості конкретного завдання. Якщо спостереження ведеться в умовах освітленості, що постійно змінюється, та/або співвідношення яскравостей символів і фону значно відрізняються для окремих написів, то цей ваговий коефіцієнт або дорівнює нулю, або близький до нього. У цьому випадку як параметр спотворення бінаризації може використовуватися вираз вигляду

$$P'_{\chi,1} = \frac{P_{\chi,1}}{P_{\chi,2}}, \quad (1)$$

що зробить метод інваріантним до спотворень за діапазоном яскравості.

Обчисливши значення оцінюваних параметрів спотворюючих функцій і визначивши сумарні витрати на спотворення за формулою (2) для кожного еталона, можна ухвалювати рішення про приналежність символу до класу, до якого належить еталон, витрати на приведення до якого мінімальні.

Висновки. У системах промислового розпізнавання, до яких належать вирішувана у цій роботі задача, характерною особливістю є велика зашумленість сигналу, що обробляється. Ця особливість даної предметної області робить малоприсадибними відомі методи обробки сигналів. Насамперед, це стосується методів попередньої фільтрації сигналу. Принципова неможливість усунути більшість перешкод у початковому сигналі потребує розроблення методів розпізнавання, не чутливих до високого рівня перешкод.

Запропонований метод еталонно-залежної обробки дає змогу усунути вплив перешкод заздалегідь відомих класів. Порівняльний аналіз із загальноприйнятою послідовністю дій при розпізнаванні показує, що метод еталонно-залежної обробки за умови правильно встановлених класів перешкод, що впливають на вихідне зображення, та оптимально підібраних вагових коефіцієнтів дозволяє коректно розпізнавати символи при високому рівні перешкод.

1. Олещук О.В. Система распознавания номеров железнодорожных вагонов // Вісник ЧІТІ (Черкаського інженерно-технологічного інституту). – 2004. – № 2. – С. 17–22. 2. Kopitchuk M.B., Oleshchuk O.V., Popel O.E. Formalization of process of identification of railway vehicles in automated control systems of goods traffic. 7th International Conference "Research and Development in Mechanical Industry", RaDMI 2007. – Belgrad, Serbia, 2007. – 16 – 20 sept, p. 838 – 841. 3. Методы компьютерной обработки изображений / Под ред. В.А. Соифера. – 2-е изд., испр. – М.: ФИЗМАТЛИТ, 2003. – 784 с. 4. Копитчук М.Б., Олещук О.В. Предварительная обработка изображений // Вісник ЧІТІ (Черкаського інженерно-технологічного інституту). – 2004. – № 3. – С. 75–80.