

8. Олексів М., Пуйда В. Первинне опрацювання зображень літаків з метою їх ефективного відділення від фону // *Advanced computer systems and networks: design and application: Матеріали 4-ї Міжнар. конф. ACSN-2009. 9–11 листопада 2009 р. – Львів, Україна, 2009. – С. 191–193.*

УДК 004.932:656.223

О.В. Олещук, О.Є. Попель, М.Б. Копитчук
Одеський національний політехнічний університет

ФОРМУВАННЯ ЗОБРАЖЕНЬ У СИСТЕМАХ ПРОМИСЛОВОГО РОЗПІЗНАВАННЯ

© Олещук О.В., Попель О.Є., Копитчук М.Б., 2010

Розглянуто процес вимірювання зображень. Проаналізовано особливості зображень, отриманих за фіксованого та змінного, тривалого та нетривалого часу експозиції. Визначено, який режим відеозйомки є найпридатнішим для конкретних задач. Наведено математичні залежності, що відповідають зображенню спостережуваного об'єкта, характерного для систем промислового розпізнавання.

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

Process of image measurement is considered. Features of images, which are received with fixed and variable, long-term and short-term exposure time, is analyzed. The most suitable modes are defined for particular tasks. Mathematical dependencies are given, which correspond to observed object that is typical for industrial system of recognition.

Keywords: image processing, character recognition, computer vision.

Вступ. Як початкові дані для систем обробки зображень та розпізнавання образів можуть виступати як одиночні зображення, так і їх послідовність. Більшість задач промислового характеру, пов'язаних із вимірюванням параметрів об'єктів, має справу з об'єктами, що змінюються у часі. Тому для систем промислового розпізнавання є сенс конкретизувати таке поняття, як "зображення об'єкта". Зазвичай воно трактується як статичне зображення, що характеризує спостережуваний об'єкт тільки у певний момент часу t .

Огляд літературних джерел. Для комп'ютерної обробки зображень необхідно насамперед формалізувати саме поняття "зображення" і забезпечити можливість перетворення зображення у числову форму.

Нехай є об'єкт, освітлений джерелом світла. На деякій відстані від об'єкта розподіл енергії джерела світлового випромінювання, відображеного об'єктом за просторовими координатами x, y і за довжинами хвиль I , описується функцією

$$C' = C'(x, y, I). \quad (1)$$

Область значень функції (1) розташована в діапазоні

$$0 \leq C'(x, y, I) \leq C_{\max},$$

де C_{\max} – максимальна яскравість зображення, обмежена граничною величиною світлочутливості реєструючих середовищ.

Геометричні розміри обмежені характеристиками формуючої системи і розмірами фото-реєструючого середовища. Вважатиметься, що усі зображення обмежені прямокутною областю:

$$x_{\min} \leq x \leq x_{\max} \quad y_{\min} \leq y \leq y_{\max},$$

де $x_{\min}, x_{\max}, y_{\min}, y_{\max}$ – мінімальні і максимальні границі по координатах x, y відповідно.

Людський зір і відеодатчики мають спектральну чутливість, що описується функцією $s(I)$, більша частина енергії якої зосереджена у деякому діапазоні (I_{\min}, I_{\max}) , відповідно за межами вказаного діапазону значення функції $s(I)$ нехтовно малі. Свого максимального значення функція спектральної чутливості досягає приблизно у середині цього діапазону і спадає ближче до його країв.

Як у разі спостереження об'єкта людиною, так і під час використання відеодатчика, спостережуване зображення є результатом усереднювання функції (1) по діапазону хвиль з ваговою функцією $s(I)$ і описується виразом

$$f(x, y) = \int_{I_{\min}}^{I_{\max}} C(x, y, I) s(I) dI . \quad (2)$$

Функція (2) зазвичай і називається зображенням. У цій роботі функція (2), щоб відрізнити її від інших видів зображення, називатиметься безперервним статичним зображенням. Отже, безперервне статичне зображення – це обмежена функція двох просторових змінних, яка задана на обмеженій прямокутній області.

У цьому випадку інтеграція ведеться по усьому діапазону довжин хвиль, внаслідок чого виходить монохромне зображення, яке відображає тільки характеристики яскравості спостережуваного об'єкта і не містить інформацію про колір. Незважаючи на втрату колірної інформації, монохромні зображення містять досить даних для аналізу зображених об'єктів, вони відрізняються простотою обробки і дешевизною апаратної реалізації [1]. Крім того, моделі і методи, запропоновані для монохромних зображень, легко розширюються для кольорових зображень шляхом розгляду кожного колірної діапазону окремо [2], тому далі увага буде зосереджена виключно на монохромних зображеннях.

Безпосередньо для комп'ютерної обробки функція (2) піддається дискретизації формуванням послідовності дискретних відліків, які після амплітудного квантування можуть бути збережені в пам'яті або оброблені. Символічно цей процес може бути записаний у вигляді

$$f(x, y) \rightarrow g(m, n) , \quad (3)$$

де

$$x = m\Delta_x + x_{\min} , y = n\Delta_y + y_{\min} ,$$

m, n – цілі числа у діапазоні;

$$0 \leq m \leq M - 1, \quad 0 \leq n \leq N - 1$$

за умови, що функція (2) дискретизується еквідистантним растром:

$$\Delta_x = \frac{x_{\max} - x_{\min}}{M} , \quad \Delta_y = \frac{y_{\max} - y_{\min}}{N} ,$$

де M, N – кількість відліків по координатах x, y відповідно.

Амплітудне квантування функції (2) виконується з урахуванням гамма-корегування, а також, як правило, з автоматичним регулюванням яскравості і контрастності. Одержана в результаті перетворення (3) яскравість

$$g = g(m, n)$$

називатиметься дискретним статичним зображенням.

Корегування амплітудних характеристик зазвичай виконується як для пристроїв вводу зображень, так і для пристроїв виводу [3]. Оскільки ця робота досліджує формування зображень, призначених для подальшої автоматичної обробки, то інтерес являє тільки корегування характеристик пристроїв вводу.

Під час вимірювання зображення $g(m, n)$ одержують зображення вигляду

$$G'(m, n) = a(g(m, n)) ,$$

де a – функція автоматичного регулювання рівня (АРР).

Реальні пристрої вводу зображень мають, як правило, нелінійну характеристику передачі рівнів яскравості.

Постановка задачі. Для розроблення методів фільтрації та розпізнавання зображень необхідно ввести поняття динамічного зображення, що розширює поняття статичного зображення і дає змогу врахувати зміну стану об'єкта у часі. Потрібно визначити особливості формування зображень залежно від конфігурації програмно-апаратних засобів та виявити обмеження, що накладає використання кожного способу адаптації вимірювальних засобів до умов освітлення і характеристик руху спостережуваних об'єктів.

Базові математичні моделі. Безперервним динамічним зображенням об'єкта називатиметься яскравість як функція часу і двох просторових координат

$$f = f(x, y, t), \quad (4)$$

де t – час.

Необхідно врахувати, що спостереження за об'єктом ведеться від деякого початкового моменту часу і має обмежену тривалість, тоді

$$t_{\min} \leq t \leq t_{\max},$$

де t_{\min}, t_{\max} – час початку і закінчення спостереження відповідно.

За аналогією з дискретним статичним зображенням вводиться поняття дискретного динамічного зображення (ДДЗ), одержаного в результаті перетворення

$$f(x, y, t) \rightarrow g_t(m, n, t), \quad (5)$$

де t – час, що пройшов з моменту початку спостереження.

Тоді

$$t = t + t_{\min}.$$

За рівномірного відліку часових інтервалів з кроком Δ_t отримано

$$t = k\Delta_t. \quad (6)$$

де k – ціле число в діапазоні

$$0 \leq k \leq K - 1, \quad (7)$$

де K – кількість кадрів зображення.

Очевидно, що

$$K = \frac{t_{\max} - t_{\min}}{\Delta_t}.$$

Тоді ДДЗ вважається функція яскравості

$$g_t = g_t(m, n, t), \quad (8)$$

отримана в результаті перетворення (5), або функція яскравості

$$g_k = g_k(m, n, k) \quad (9)$$

отримана в результаті аналогічного перетворення:

$$f(x, y, t) \rightarrow g_k(m, n, k).$$

Використовуючи (6), ДДЗ визначається як послідовність дискретних статичних зображень, отриманих за умови (7).

Особливості формування динамічного зображення. Розглядаючи докладніше дискретизацію зображень за часом, для простоти береться рівномірний відлік часових інтервалів, тоді вимірювання кожного кадру зображення починається у момент часу t згідно р (6), але оскільки будь-яке вимірювання має ненульову тривалість, то результатом вимірювань буде не функція вигляду (8), а функція

$$G_t(m, n, t) = a_t(P(m, n, t), a_{t-\Delta_t}), \quad (10)$$

де a_t – функція АРР у момент часу t ; $a_{t-\Delta_t}$ – функція АРР у момент часу $t - \Delta_t$;

$$P(m, n, t) = \int_{n\Delta_y + y_{\min}}^{(n+1)\Delta_y + y_{\min}} \int_{m\Delta_x + x_{\min}}^{(m+1)\Delta_x + x_{\min}} F(x, y, t + t_{\min}) dx dy; \quad (11)$$

$$F(x, y, t) = \int_t^{t+d_t} f(x, y, T) dT, \quad (12)$$

де d_t – час експозиції, тобто тривалість вимірювання; T – інтервал часу від моменту t до $t + d_t$.

Функція (12) є інтегралом функції яскравості в інтервалі часу, що відповідає тривалості вимірювання одного кадру, а функція (11) – інтеграл функції (12) по просторових координатах відповідно до розмірів одного пікселя.

АРР включає у себе автоматичний підбір яскравості, контрастності і гамма-корегування залежно від інтенсивності освітлення в точці спостереження. Функція АРР є безперервною, монотонно зростаючою функцією по $P(m, n, t)$. Особливості її поведінки залежать від використовуваних апаратних і програмних засобів. Якщо a_t залежить тільки від яскравості g_t у момент часу T і не залежить від функції АРР в попередні моменти часу, це означає, що усі параметри жорстко встановлені вручну до початку часу спостереження, тобто автоматичне регулювання не використовується. Функція (10) матиме особливості залежно від значення d_t .

Модель динамічного зображення за постійного тривалого часу експозиції. За часу експозиції, зрівняного із швидкістю зміни яскравості в точці спостереження (рис. 1) яскравість усереднюється по доволі тривалому часовому інтервалу.

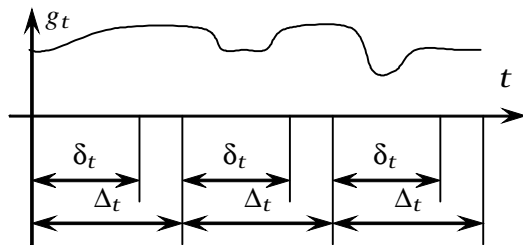


Рис. 1. Формування зображення за тривалого часу експозиції

Яскравість в точці спостереження прямо пропорційна до інтенсивності освітлення і відбивної здатності спостережуваних об'єктів, тобто

$$f(x, y, t) = R(x, y, t)L(x, y, t), \quad (13)$$

де $R(x, y, t)$ – відбивна здатність об'єкта в точці з координатами x, y у момент часу t ; $L(x, y, t)$ – освітленість в точці з координатами x, y в момент часу t .

Аналогічно до виразу (13) можна записати:

$$g_k(m, n, k) = g_t(m, n, t) = R_t(m, n, t)L_t(m, n, t), \quad (14)$$

де $R_t(m, n, t)$ – квантована відбивна здатність об'єкта за дискретними просторовими координатами m, n у момент часу t ; $L_t(m, n, t)$ – квантована освітленість в точці з дискретними координатами m, n у момент часу t .

Характеристики інтенсивності освітлення істотно залежать від типу освітлення: природного, штучного або комбінованого. У разі застосування штучного освітлення змінами інтенсивності в часі можна знехтувати. При використанні природного світла час істотної зміни освітленості, як правило, значно перевищує величину Δ_t , тому зміни інтенсивності освітлення можуть бути компенсовані системою АРР. За допущення, що відбивна здатність об'єктів не змінюється з часом, буде одержано, що найістотніший вплив на зміну функції g_t надасть зміна просторового розташування спостережуваних об'єктів.

Розглянемо ситуацію, коли присутній єдиний об'єкт спостереження, який в межах часу експозиції переміщається прямолінійно і рівномірно уздовж горизонтальної осі, перпендикулярної до напрямку "погляду" відеодатчиків (рис. 2). Подібний взаємний

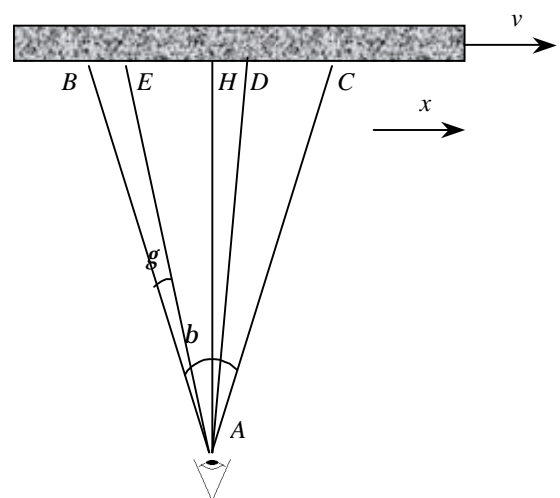


Рис. 2. Переміщення об'єкта, що спостерігається, під час експозиції

рух та розташування датчиків та спостережуваного об'єкта є характерним для багатьох систем машинного зору: для розпізнавання деталей на конвеєрі, а також під час зчитування номерів на залізничних вагонах [4].

За час d_t точка B об'єкта спостереження переміститься в точку E . Прийmemo, що відстань AN від точки спостереження A до спостережуваного об'єкта велика порівняно з ділянкою BC , що потрапляє у поле зору, тобто кут b доволі малий. Тоді буде справедливим твердження: кутова швидкість g переміщення променя спостереження кожної частини об'єкта спостереження не залежить від просторового розташування цієї частини.

Для того, щоб припущення про рівність кутових швидкостей відповідало дійсності, кут огляду b повинен бути достатньо малий або відмінність кутових швидкостей була компенсована оптичними або алгоритмічними методами.

Для об'єкта, що задовольняє вищеописаним умовам, вираз (11) з урахуванням (12) набуде вигляду:

$$P(m, n, t) = \int_{n\Delta_y + y_{\min}}^{(n+1)\Delta_y + y_{\min}} \int_{m\Delta_x + x_{\min}}^{(m+1)\Delta_x + x_{\min}} \int_t^{t+d_t} f(x + vT, y, t_{\min}) dT dx dy, \quad (15)$$

де v – швидкість руху об'єкта.

Причому абсолютне значення величини v дорівнює модулю вектора швидкості, а знак цієї величини буде додатним, якщо напрямок руху об'єкта збіжиться з напрямком горизонтальної осі, та від'ємним, якщо напрямок руху протилежний до напрямку горизонтальної осі.

В такий спосіб зміни виразів x і vT мають аналогічний вплив на зміну значення результуючої функції $P(m, n, t)$, отже, щоб вплив часу експозиції на виміряне значення для кожного пікселя був неістотним, необхідно, щоб виконувалася умова

$$v((t + d_t) - t) \ll ((m + 1)\Delta_x + x_{\min}) - (m\Delta_x + x_{\min}),$$

одержана з урахуванням меж інтегрування у виразі (15). Звідси

$$\begin{aligned} vd_t &\ll \Delta_x; \\ d_t &\ll \frac{\Delta_x}{v}. \end{aligned} \quad (16)$$

З виразу (16) можна зробити висновок, що максимально допустимий час експозиції прямо пропорційний до розміру пікселя й зворотно пропорційний до часу руху об'єкта.

Висновок 1. Якщо швидкість руху об'єкта і час експозиції фіксовані, то існує граничне значення роздільної здатності, вище за яке не досягається істотного поліпшення якості зображення.

На рис. 3 показаний процес вимірювання одного пікселя зображення протягом часу експозиції за рівномірного руху об'єкта уздовж горизонтальної осі.

Інтегрування за часом і координатою x призводить до формування прямокутної області, показаної жирною лінією. Її внутрішня частина є зміряним значенням пікселя. З часом вимірюваний піксель зміщується уздовж горизонтальної осі, утворюючи паралелограм. В результаті зміряне значення пікселя збігається з істинним тільки в області B . Область A містить частину пікселя, що знаходиться зліва від вимірюваного, а область C – частина вимірюваного пікселя.

Помилка вимірювання залежатиме від співвідношення правильної і помилкової інформації, тобто від співвідношення площ B та A . Порівняння рис. 3, а з рис. 3, б дає змогу зробити висновок, що за однакової швидкості руху (графічно швидкості відповідає кут нахилу паралелограма) достовірнішим буде вимірювання пікселя за допомогою відеодатчиків з низькою роздільною здатністю. Для збільшення достовірності вимірювань датчиками високої роздільної здатності буде необхідно підсумовувати сусідні пікселі.

Існує математична залежність, яка дає змогу відновити початковий сигнал за "змазування" об'єкта [3], проте за значного зрушення спостережуваного об'єкта задача переходить в розряд некоректно поставлених, отже, не дає можливості одержати точний і достовірний результат.

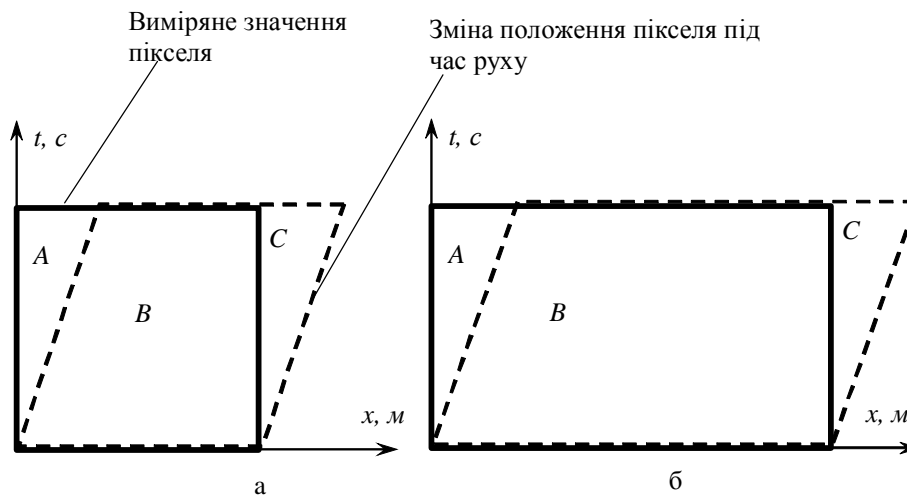


Рис. 3. Формування пікселя зображення рухомого об'єкта:
 а – висока роздільна здатність відеодатчиків; б – низька роздільна здатність

Крайнім випадком є ситуація, коли $d_t = \Delta_t$. При цьому вимірювання вхідного сигналу виконується у кожен момент часу, отже, це є єдиний прийнятний режим для відстеження короткочасних змін сигналу.

Модель динамічного зображення за постійного нетривалого часу експозиції. Рис. 4 відображає ситуацію, коли час експозиції малий порівняно із швидкістю зміни яскравості в точці спостереження, тобто виконується умова (16). У цьому випадку заштриховану область на рис. 4, що відповідає виміряному значенню яскравості в певний момент часу, можна вважати прямокутною, тобто значення функції (15) буде прямо пропорційно до значення d_t .

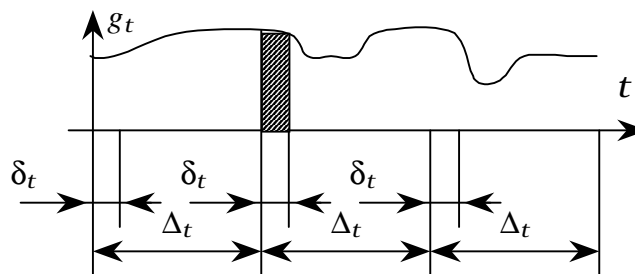


Рис. 4. Формування зображення за нетривалого часу експозиції

Одержаний масштабний коефіцієнт за функції (15) буде врахований системою АРР і компенсується коефіцієнтом посилення яскравості, зворотно пропорційним до величини d_t .

Висновок 1. Якщо виконується умова (16), то d_t істотно не впливає на значення виразу (10).

Із зменшенням часу експозиції збільшується проміжок часу, коли не проводиться вимірювання сигналу.

Отже, режим роботи з нетривалим часом експозиції може бути рекомендований тільки для об'єктів, що знаходяться у полі зору відеодатчиків доволі тривалий час, перевищуючи величину Δ_t . Вибір конкретного значення d_t з допустимого діапазону може бути зроблений з урахуванням особливостей освітлення і системи АРР.

Модель динамічного зображення за змінного часу експозиції. Вплив часу експозиції на діапазон яскравості, сприйманий системою вимірювання зображення, став основою для систем

АРР, що використовують автоматичне регулювання швидкості затвора для обліку освітленості. За такого підходу час експозиції для кожного кадру варіюватиметься в межах $0 < d_t \leq \Delta_t$ (рис. 5).

Безперечною перевагою такого підходу є швидкість автокорегування діапазону яскравості на відміну від програмних або електромеханічних способів, проте використання змінного часу експозиції об'єднує в собі недоліки розглянутих вище установок діапазонів часу експозиції, а саме: можуть бути тривалі паузи, коли сигнал не вимірюється, і може вироблятися усереднювання сигналу по тривалому інтервалу часу те, що призводить до ефекту "змазування", причому наперед невідоме, коли існуватиме будь-який недолік.

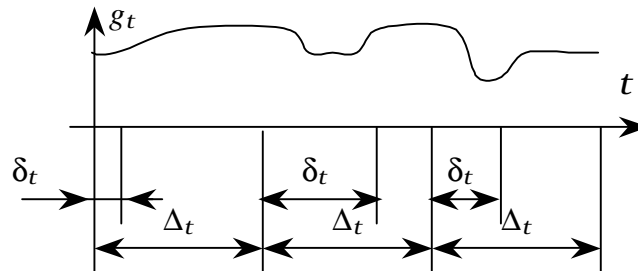


Рис. 5. Формування зображення за змінного часу експозиції

Отже, режим вимірювання із змінним часом експозиції малопридатний для отримання зображень рухомих об'єктів.

Висновок. Залежно від особливостей освітлення, швидкості руху спостережуваного об'єкта та швидкості можливих його змін слід використовувати різні підходи до дискретизації зображень за часом.

Здебільшого потрібно формувати зображення за постійного нетривалого часу експозиції. Такий підхід, особливо, якщо умови освітлення не дають змоги враховувати корегування, яке вноситься системою АРР, дає можливість істотно спростити методи обробки зображень, оскільки припустимо буде використовувати як математичні моделі функцію (8), а не узагальненішу функцію виду (10).

Пропоновані математичні моделі були використані у технології ідентифікації залізничних вагонів [5] і покладені в основу методів розпізнавання текстової інформації на зображеннях [6].

1. Демьяновски В. Библия охранного телевидения. – М.: Изд-во ИСС, 2002. – 323 с.
2. Фурман Я.А., Юрьев А.Н., Яншин В.В. Цифровые методы обработки и распознавания бинарных изображений. – Красноярск: Изд-во Краснояр. ун-та, 1992. – 248 с.
3. Методы компьютерной обработки изображений / Под ред. В.А. Сойфера – 2-е изд., испр. – М.: ФИЗМАТЛИТ, 2003. – 784 с.
4. Олещук О.В. Система распознавания номеров железнодорожных вагонов // Вісник ЧІТІ (Черкаського інженерно-технологічного інституту). – 2004. – № 2. – С. 17–22.
5. Kopytchuk N., Oleshchuk O. & Popel O.: Formalization of process of identification of railway vehicles in automated control systems of goods traffic. In: Proceedings of 7th International Conference ²Research and Development in Mechanical Industry – RaDMI 2007², Beograd, Serbia, 16–20 September 2007. Edited by Predrag Dašić. Trstenik: High Technical Mechanical School, 2007. – P. 838–841.
6. Oleshchuk O., Popel O. & Kopytchuk N.: Standard-dependent symbol recognition. In: Proceedings of 9th International Conference ²Research and Development in Mechanical Industry – RaDMI 2009², Vrnjačka Banja, Serbia, 16–19 September 2009. High Technical Mechanical School, 2009.