

РЕАЛІЗАЦІЯ АЛГОРИТМУ ПОКАДРОВОГО ВИЗНАЧЕННЯ КООРДИНАТ ОБ'ЄКТА У МОНОХРОМНОМУ ВІДЕОПОТОЦІ

© Пуйда В. Я., 2015

Розглянуто реалізацію алгоритму визначення координат об'єкта у монохромному відеопотоці на основі методу аналізу гістограми відеозображення з використанням мікроконтролера з ядром DSP.

Ключові слова: система технічного зору, відеопотік, бінарне зображення, сегментація.

IMPLEMENTATION OF AN ALGORITHM FOR FINDING OBJECT COORDINATES IN MONOCHROME VIDEO

© Puyda V., 2015

We consider a microcontroller implementation of a histogram-based algorithm for finding object coordinates in monochrome video.

Key words: computer vision, monochrome video, image binarization, image segmentation.

Вступ

У системах технічного зору виникає задача визначення координат об'єктів у реальному часі. Існує багато методів розв'язання такої задачі. Вибір методу залежить від специфіки системи технічного зору, складності сцени, заданих технічних параметрів, допустимих апаратних засобів для реалізації.

Аналіз публікацій

Численні дослідження у цьому напрямі проводили багато зарубіжних, наприклад, Nobuyuki Otsu, A. Anjos, H. Shahbazkia, та українських вчених, наприклад, Є. С. Козирев, Я. О. Романюк, Г. В. Славко, Д. Д. Пелешко. Одним з ефективних методів визначення координат об'єкта у відеопотоці є гістограмний метод [1, 2]. Він може використовуватися для визначення координат об'єкта в реальному часі на сцені з мінімальними завадами та з використанням порівняно простих апаратних засобів.

Постановка задачі

Вхідний відеопотік складається з послідовності $(I_n)_{n \in N}$ кадрів розміром 640×480 пікселів. Кожен кадр $I_n = (I_n(i, j))$, $n \in N$, $i \in \{1, \dots, 480\}$, $j \in \{1, \dots, 640\}$ є монохромним зображенням. Глибина кожного пікселя становить 8 бітів: $I_n(i, j) \in \{0, \dots, 255\}$ для всіх n, i, j .

На кожному кадрі I_n кожна точка класифікується в одну з двох груп: точки, які належать об'єкту, і точки, які належать фону: для кожного $n \in N$,

$$\{(i, j)\}_{i=1, \dots, 480, j=1, \dots, 640} = Obj_n \cup Backgr_n, \quad Obj_n \cap Backgr_n = \emptyset. \quad (1)$$

Задача, що ставиться, полягає у відшуванні координат об'єкта на кожному кадрі відеопотоку в реальному часі.



Рис. 1. Кадр з відеопотоку

Розв'язання задачі. В роботі розглянуто апаратну реалізацію методу аналізу гістограми відеозображення [1, 2]. За координати об'єкта приймають центроїд множини Obj_n :

$$(x_{obj,n}, y_{obj,n}) := \frac{1}{\sum_{(i,j) \in Obj_n} 1} \left(\sum_{(i,j) \in Obj_n} j, \sum_{(i,j) \in Obj_n} i \right). \quad (2)$$

Алгоритм пошуку координат об'єкта може виглядати так: для кожного кадру I_n , необхідно виконати такі дії:

1. Бінаризація зображення.

Як правило, гістограма

$$hist_n = (hist_n(k))_{k=0, \dots, 255}, \quad hist_n(k) := \sum_{\{(i,j) | I_n(i,j)=k\}} 1, \quad (3)$$

зображення I_n є бімодальною: одна з мод відповідає фону, інша – об'єкту.

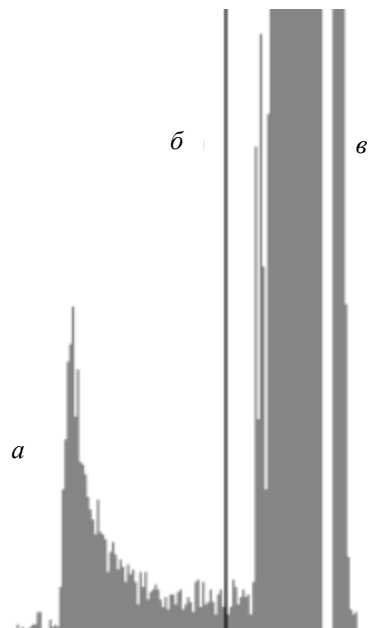


Рис. 2. Гістограма, де мода:
а – відповідає об'єкту; б – поріг; в – відповідає фону

Величина порога для бінаризації може бути адаптивною або сталою. Використання адаптивної величини порога збільшує тривалість обробки кадру. Якщо кількість пікселів, що відповідають об'єкту, співмірна з кількістю пікселів, що відповідають фону, то, якщо вибрати сталу величину порога бінаризації, цю величину можна визначити за допомогою класичного методу Оцу [1].

Якщо кількість пікселів, що відповідають об'єкту, є істотно меншою, ніж кількість пікселів, що відповідають фону, то кращий результат отримують за допомогою балансування гістограми [2]:

– встановити межі пошуку зліва і справа (крайній лівий і крайній правий біни гістограми, відповідно) $left = 0$, $right = 255$;

– вибрати центральний бін гістограми

$$center = \frac{left + right - 1}{2}; \quad (4)$$

– обчислити кількість пікселів (вагу гістограми) зліва і справа від центрального біна

$$w_{left} = \sum_{k=left}^{center-1} hist_n(k), \quad w_{right} = \sum_{k=center+1}^{right} hist_n(k). \quad (5)$$

– якщо вага справа перевищує вагу зліва, то змістити праву межу пошуку вліво на один бін і переобчислити ваги та центральний бін:

$$w_{left} = w_{left} - hist_n(center - 1), \quad w_{right} = w_{right} + hist_n(center) - hist_n(right), \\ right = right - 1, \quad center = center - 1; \quad center = center + 1; \quad (6)$$

– якщо вага зліва перевищує вагу справа, то змістити ліву межу пошуку вправо на один бін і переобчислити ваги та центральний бін:

$$w_{left} = w_{left} + hist_n(center) - hist_n(left), \quad w_{right} = w_{right} - hist_n(center + 1), \\ left = left + 1, \quad center = center + 1; \quad (7)$$

– повторювати попередній крок, доки права межа пошуку не перевищить ліву.

Після завершення процедури величина порога бінаризації відповідає центральному біну: $threshold = center$.



Рис. 3. Бінаризоване зображення об'єкта в кадрі

2. Сегментація зображення.

Якщо тривалість обробки кадру дає змогу виконати цей крок, то необхідно здійснити сегментацію бінаризованого зображення і вибрати сегмент, який відповідає об'єкту.

Якщо бінаризоване зображення містить лише два сегменти, то об'єкту, як правило, відповідає той сегмент, що не має спільних точок з межею зображення. В іншому разі сегмент, який відповідає об'єкту, повинен зафіксувати оператор у момент початку стеження.

На кожному наступному кадрі i_{n+1} об'єкта, як правило, відповідає той сегмент, центроїд якого найближче до центроїда сегмента, що відповідав об'єкту на попередньому кадрі i_n .

3. Обчислити координати центроїда сегмента, що відповідає об'єкту, і прийняти ці координати за координати об'єкта.

Якщо тривалість обробки кадру не дає змоги виконати сегментацію зображення, то за координати об'єкта слід прийняти координати центроїда тих точок бінаризованого зображення, що розміщені по ту саму сторону від порога бінаризації, що й мода, яка відповідає об'єкту:

$$(i, j) \in Obj_n \Leftrightarrow I_n(i, j) < threshold . \quad (8)$$



Рис. 4. Визначений центр зображення об'єкта в кадрі відеопотоку

На рис. 5 показано структуру спецпроцесора для реалізації вказаного методу.

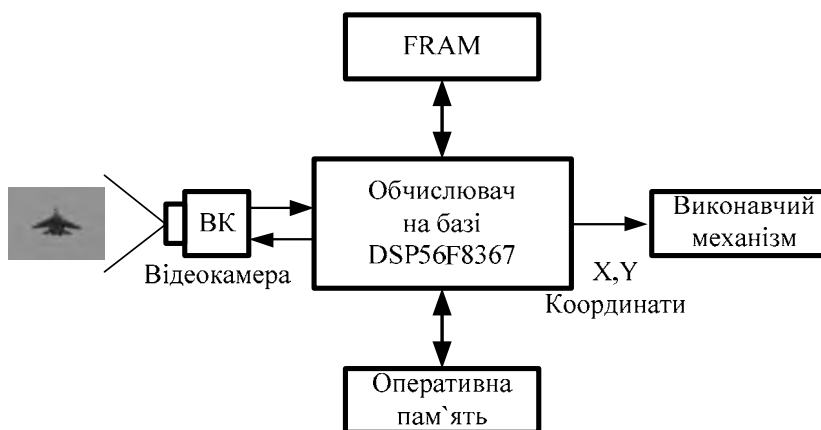


Рис. 5. Структурна схема спецпроцесора визначення координат об'єкта

Специфіка сцени та зображень об'єктів допускає використання монохромної відеокамери. Монохромне зображення забезпечує мінімізацію пам'яті для збереження зображення сцени в процесі реалізації алгоритму, зменшує вимоги до продуктивності обчислювача і в кінцевому результаті забезпечує визначення координат об'єкта в полі зору відеокамери за час, менший від тривалості одного телевізійного кадру в стандарті SECAM та PAL. Для реалізації алгоритму використано мікроконтролер DSP56F8367 з ядром DSP56800E (рис. 6).

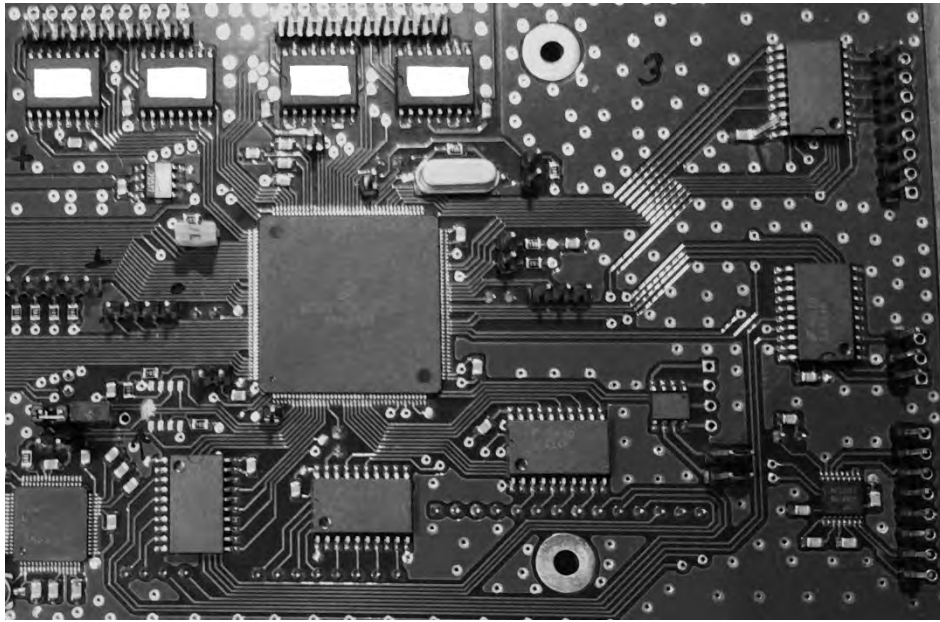


Рис. 6. Фрагмент плати спецпроцесора на базі DSP56F8367

Якщо необхідно застосувати потужніші методи для реалізації процесу ідентифікації та визначити координати ідентифікованого об'єкта на складних сценах, можна використати потужніший мікроконтролер, наприклад, Blackfin Dual Core ADSP-BF6xx, TMS320C6678 р 8-ядерною архітектурою TI Key Stone чи інший подібного класу.

Висновки

Алгоритм визначення координат об'єкта на простій сцені у монохромному відеопотоці на основі методу аналізу гістограми відеозображення може бути реалізований на базі сучасних мікроконтролерів середньої продуктивності з ядром DSP для систем технічного зору різного призначення. Орієнтовний час визначення координат не перевищує тривалості одного телевізійного кадру.

1. Nobuyuki Otsu (1979). *A threshold selection method from gray-level histograms* // *IEEE Trans. Sys., Man., Cyber.* 9. – P. 62–66. 2. Anjos A. and Shahbazkia H. *Bi-Level Image Thresholding – A Fast Method* // *BIOSIGNALS.* – 2008. – Vol. 2. – P. 70–76.