

БИНАРИЗАЦИЯ ИЗОБРАЖЕНИЯ В СИСТЕМАХ ТЕХНИЧЕСКОГО ЗРЕНИЯ

Формирование изображений включает следующие подэтапы: настройка системы, в том числе подбор освещения рабочей сцены; аналоговые преобразования видеосигнала; ввод изображения в компьютер или другое решающее устройство; предварительная обработка оцифрованного изображения и сжатия информации в целях облегчения дальнейших этапов технического зрения [1].

Настройка СТЗ. Алгоритмы автоматизированной настройки системы очень сильно зависят от используемой аппаратуры, а также от конкретного вида рассматриваемой сцены.

В ряде современных систем применяется управляемая подстройка интенсивности источников света и/или их пространственной направленности, которая осуществляется в контуре обратной связи по специальным программам до удовлетворения выбранных критериев зрительного восприятия [2].

Автоматизация настройки видеосистемы предусматривает также наведение видеосенсора на нужный объект – выбор поля зрения, фокусировку и диафрагмирование объекта. При рассмотрении сцен с изолированными объектами определяется такое расположение камеры и такие параметры оптического тракта, при которых объекты полностью попадают в заданную рамку; в других случаях эту рамку можно наложить на нужный фрагмент сцены, «привязывая» её к некоторым характерным точкам.

Аналоговые преобразования. Алгоритмы аналоговых преобразований видеосигнала на стадии формирования изображения в СТЗ, как правило, реализуются аппаратными средствами и служат для обеспечения требуемых электрических характеристик сигнала.

Ввод изображения в компьютер. В общем случае под изображением рабочей сцены будем понимать двумерную функцию $g(x, y)$, где g – яркость соответствующей точки сцены (x, y) . Вместо исходного изображения $g(x, y)$ в память компьютера вводится массив значений $G(i, j)$, задаваемый на целочисленной решетке $i = \overline{0, N_x}, j = \overline{0, N_y}$, по-

казывающей всю область $X \times Y$ или её некоторую подобласть. Значение G в узловой точке представляет усредненную яркость g элемента изображения в окрестности этой точки.

После квантования функция $G(i, j)$ принимает целочисленные значения: $G(i, j) = k$ при $T_{k-1} \leq G(i, j) < T_k$, $k \in [0, k_{\max}]$, где T_k – значение k -го порогового уровня. В случае $k_{\max} = 1$ оцифрованное изображение называется бинарным (двухградационным), в случае $k_{\max} > 1$ – полутоновым (многоградационным).

Для автоматического выбора порога бинаризации T чаще всего строят гистограмму яркости (показывающая, какое число n пикселей картинной плоскости имеют данную яркость), после чего ищут минимум гистограммы на впадине между двумя основными её пиками, отвечающими яркостям фона и объектов (рис. а).

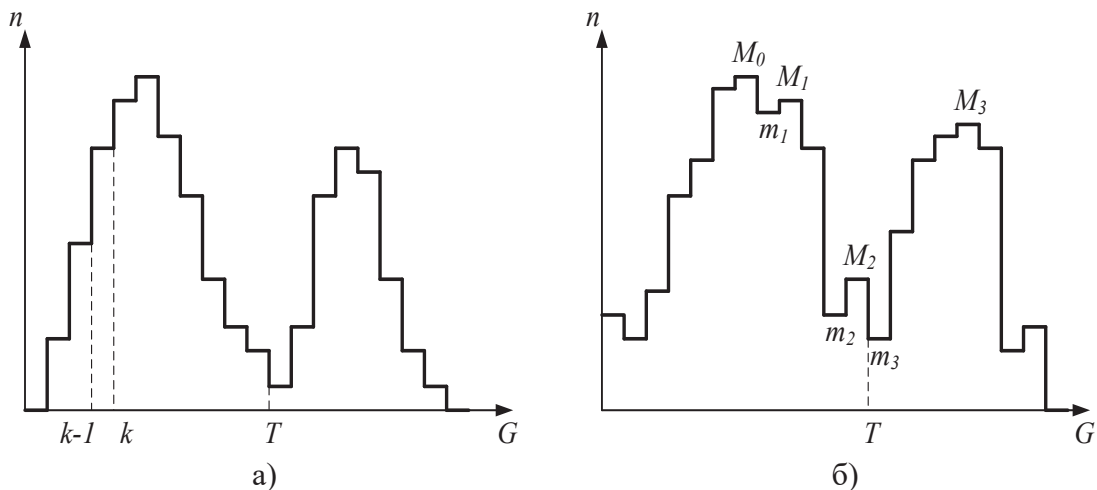


Рисунок – Определение порога бинаризации по гистограммам яркости

Однако вследствие неидеальной контрастности реальная гистограмма редко бывает бимодальной: наряду с двумя основными пиками она имеет множество дополнительных отражающих, наличие теней и бликов, шумов и неоднородной чувствительности видеосенсора (рис. б).

Бинаризация изображения с глобальным порогом осуществляется по формуле [3]:

$$G(i, j) = \begin{cases} 1, & G(i, j) \geq T \wedge |G(i, j)| \leq \varepsilon; \\ 0. & \end{cases} \quad (1)$$

где ε – дополнительный порог по градиенту.

Бинаризация изображения с локальным порогом осуществляется по формуле:

$$G(i, j) = \begin{cases} 1, & G(i, j) > \bar{G}(i, j) - D(i, j); \\ 0. & \end{cases} \quad (2)$$

где $D(i, j) = 0,5[G(i, j) - G(i-1, j) + G(i, j) - G(i+1, j)]$, а $\bar{G}(i, j)$ – среднее значение яркости на выбранном окне.

Если разбить всё поле зрения на подобласти, в которых легко выбрать локальные пороги, то далее можно провести пространственную интерполяцию порога бинаризации через центры таких подобластей. Это даёт переменный (динамический) порог для всего поля зрения в целом.

Предварительная обработка изображения. Эта группа операций направлена на облегчение дальнейших этапов технического зрения и проводится в целях коррекции различных искажений, улучшения контрастности, удаления шумов и т. п..

Традиционным методом фильтрации изображений в целях выделения полезной информации из «зашумленной» картины является использование Фурье-преобразования, которое основано на хорошо исследованном теоретическом аппарате и позволяет применять современные средства, как цифровой (например, быстрое Фурье-преобразование), так и аналоговой реализации.

Дискретное Фурье-преобразование, реализуемое программным путем, дает спектр $\Gamma(l, m)$ функция яркости $G(j, k)$ в соответствии с формулой:

$$\Gamma(l, m) = \frac{1}{n_x n_y} \sum_{j=0}^{n_x-1} \sum_{k=0}^{n_y-1} G(j, k) \exp \left[-2\pi i \left(j \frac{l}{n_x} + k \frac{m}{n_y} \right) \right]. \quad (3)$$

Обозначив Фурье-преобразование F , определим обратное образование F^{-1} и получаем

$$G(j, k) = \sum_{l=0}^{n_x-1} \sum_{m=0}^{n_y-1} \Gamma(l, m) \exp \left[2\pi i \left(j \frac{l}{n_x} + k \frac{m}{n_y} \right) \right]. \quad (4)$$

Введем функцию $H(l, m)$, задающую линейный пространственный фильтр. Произведение спектра $\Gamma(l, m)$ и этой функции определит Фурье-образ на выходе фильтра $\Gamma_0(l, m) = \Gamma(l, m) \times H(l, m)$, откуда получаем выходную функцию яркости следующего вида:

$$G_0(i, j) = F^{-1} \{ \Gamma_0(l, m) \} = F^{-1} \{ \Gamma_0(l, m) H(l, m) \} = G(i, j) * h(i, j), \quad (5)$$

где $h(i, j) = F^{-1} \{ H(l, m) \}$, а символом $*$ обозначена операция свёртки.

Поскольку программная реализация алгоритмов, базирующихся на Фурье-преобразованиях, требует сравнительно больших вычислительных затрат, ниже представлено более простое преобразование. Так, преобразование Адамара имеет спектр (при $n_x = n_y = n$)

$$\Gamma_1(l, m) = \frac{1}{n} \sum_{l=0}^{n-1} \sum_{k=0}^{n-1} G(j, k) (-1)^{\alpha_{jk}(l, m)}, \quad (6)$$

где $\alpha_{jk}(l, m) = \sum_{v=0}^{n-1} [\beta_v(j)\beta_v(l) + \beta_v(k)\beta_v(m)]$, а коэффициенты $\beta_v(q)$ равны либо 0, либо 1 в соответствии со значением v -го разряда числа q , представленного в двоичной системе счисления.

Перечисленные выше алгоритмы фильтрации и сглаживания, устраняя высокочастотные шумы, одновременно приводят к потере мелких деталей изображения и к размыванию границ объектов, что напоминает эффект расфокусировки.

Стремясь ослабить это нежелательное побочное явление, обращаемся к выборочным усреднениям, пользуясь следующими методами:

1) Провести усреднение не по всем элементам выбранного окна, а лишь по тем, которые лежат вблизи границы между фоном объектом; как модификация такого метода применяется усреднение по элементам, находящимся по ту сторону от границы, что и рассматриваемый пиксель;

2) При усреднении учитывают только те пиксели выбранного окна, яркость которых отличается от яркости рассматриваемого элемента не более чем на заранее заданное значение;

3) В выбранном окне выделяют подмножество элементов, лежащих по разные стороны от рассматриваемого пикселя и дающих минимальный разброс значений яркости, и усреднение проводят только по этому подмножеству.

ЛИТЕРАТУРА

1. Мошкин, В.И. Техническое зрение роботов / А.А. Петров, В.С. Титов, Ю.Г. Якушенков. – Москва: Под общ. ред. Ю.Г. Якушенкова. – М.: Машиностроение, 1990. – 272 с.
2. Shirai Y. Robot vision // Robotics. 1986. V. 2. N. 3. P. 175-203.
3. Келли, Р.Б. Три алгоритма технического зрения для задачи взятия деталей из бункеров / Р.Б. Келли, Э.А.Ш. Мартиш, Дж.Р. Берк, Ж.Д. Дессимос: Пер. с англ. ТИИЭР. 1983. Т.71. № 7. С. 23-44.