

МОДИФИКАЦИЯ АЛГОРИТМА КОРРЕЛЯЦИОННОГО АНАЛИЗА СТЕРЕОИЗОБРАЖЕНИЙ SAD ДЛЯ НАХОЖДЕНИЯ СООТВЕТСТВИЙ ОБЛАСТЯМ ВЫСОКОЙ ОДНОРОДНОСТИ

К. Д. Колесников

kolesnikovkirilld@gmail.com;

Научные руководители — Н. Н. Яцков, кандидат физико-математических наук, доцент; О. Ф. Ковалев, старший преподаватель

В работе предложена модификация алгоритма оконного сопоставления стереоизображений SAD с использованием процедур адаптации и локализации. Данная модификация решает одну из главных проблем стереозрения – нахождение соответствия между областями высокой степени однородности. Работоспособность подтверждается на примерах обработки наборов стереоизображений.

Ключевые слова: стереозрение; карта глубины; адаптивный алгоритм; однородные области; оконное сопоставление стереоизображений.

ВВЕДЕНИЕ

С каждым годом увеличивается интерес к робототехнике, появляются новые технологии, позволяющие создавать сложных и мультифункциональных роботов. Одним из простейших способов практической реализации зрения роботов является использование цифровых камер. Однако стандартные цифровые камеры производят двумерные изображения, не позволяющие определить расстояние от точки съёмки до объектов на изображении. Для определения расстояний используются методы кибернетики, а именно – стереозрения. В общем случае сложно получить точную матрицу расстояний объектов до точки съёмки (часто ее называют картой глубины), что обусловлено наличием окклюзий, отсветов, а также однородных областей.

Для построения карты глубины используются локальные и глобальные методы. Локальные строят карту по частям, глобальные – оптимизируют карту целиком. Достоинствами локальных алгоритмов является скорость, а глобальных – точность. Локальные подразделяются на корреляционные – SAD, SSD, NCC, CT (аббревиатуры указывают на используемую метрику сходства областей или преобразование, SAD – сумма абсолютных отклонений, SSD – сумма квадратов отклонений, NCC – нормализованная кросс-корреляция, CT – преобразование переписи) и на методы, учитывающие особенности изображений, среди них – HCD (алгоритм детекции углов Харриса), SIFT (масштабно-инвариантное преобразование особенностей), SURF (ускоренный алгоритм надежного

нахождения особенностей). В классических локальных методах не предусмотрен механизм работы с областями высокой однородности [1].

Из вышеперечисленных методов самым простым в программной реализации, надежным и вычислительно эффективным является метод SAD[2]. Наиболее точным, но также и высоко затратным считается метод на основе нейронных сетей (к примеру, Module Net) [3].

В работе предложена модификация метода SAD, основанная на использовании вычислительных процедур адаптации и локализации, позволяющая повысить точность сопоставления однородных областей.

МЕТОДОЛОГИЯ

Идея разработанного алгоритма состоит в адаптивном растяжении областей, для которых ищется соответствие, при их невысокой контрастности. Для оценки порога сохранения информативности в процедуре растяжения применен механизм локализации. В качестве меры контрастности (однородности) области и значения порога используется средне-квадратичное отклонение (СКО) распределения яркости пикселей на изображении. Для повышения точности и увеличения быстродействия предполагается, что стереопара (далее рассматривается как изображение слева и справа) горизонтально ректифицирована – это позволяет ограничить область поиска соответствий до, как минимум, четверти ширины изображения. Для дальнейшей локализации области поиска определяется аргумент максимума корреляционной функции стереопары, с некоторым весовым коэффициентом он принимается за максимальное отклонение объектов на первом изображении от их аналогов на втором.

Алгоритм:

- 1) Вычислить корреляционную функцию изображений стереопары
- 2)

$$R(\tau) = \frac{\sum_{x,y \in I} (I_1(x,y) - \bar{I}_1)(I_2(x-\tau,y) - \bar{I}_2)}{\sqrt{\sum_{x,y \in I} (I_1(x,y) - \bar{I}_1)^2 \sum_{x,y \in I} (I_2(x-\tau,y) - \bar{I}_2)^2}}, \quad (1)$$

где I_1, I_2 – первое и второе изображение стереопары, $\tau \in [-N/4; N/4]$ – компенсация сдвига камеры, N – ширина изображения, $\bar{I}_1$ и $\bar{I}_2$ – средние значения яркости пикселей на изображениях.

Определить максимум $\max(R(\tau))$ и аргумент $\operatorname{argmax}(R(\tau))$.
Максимальное отклонение принимается равным:

$$MD = \frac{EC}{\max(R(\tau))} \cdot \operatorname{argmax}(R(\tau)), \quad (2)$$

где EC – некоторый весовой коэффициент.

В формуле (2) $\operatorname{argmax}(R(\tau))$ является оценкой среднего отклонения объектов стереопары, $\max(R(\tau))$ характеризует разброс отклонений. Весовой коэффициент подобран равным $EC = 2.2$.

3) Локализовать пороги по СКО. Для этого первое (левое) изображение разделяется на прямоугольные сегменты. Для сегментов вычисляются их СКО, пороги определяются по формуле:

4)

$$AT_{i,j} = AC \cdot \sqrt{СКО_{i,j}}, \quad (3)$$

где AC – весовой коэффициент, $СКО_{i,j}$ – значение СКО сегмента $S_{i,j}$.

Квадратный корень в формуле (3) необходим для усиления ограничения растяжения областей в сегментах достаточного контраста. Весовой коэффициент подобран равным $AC = 2.1$.

5) Найти соответствия для областей левого изображения на правом (рис. 1). Для каждой области первого изображения вдоль эпполярной линии (по горизонтали) определить наиболее схожую во втором, ограничивая область поиска рассчитанным максимальным отклонением MD . Модификация корректирует области, которым нужно найти соответствия. После получения базовой области определяется, какому сегменту $S_{i,j}$ она принадлежит, запоминается порог $AT_{i,j}$. Далее область увеличивается влево пока ее контраст не превысит порог, после чего область уменьшается до своей величины на предпоследней итерации.

6) Вычислить карту глубины. Рассчитать по парам соответствующих областей относительные расстояния, заполнить карту глубины.

РЕЗУЛЬТАТЫ

В работе рассмотрены три стереопары с наличием однородных областей. Метод был применен со следующими параметрами: количество сегментов – 25, размеры окна сканирования – 7, 5 и 3; значения карты рассчитывались только для неперекрывающихся областей. Для оценки эффективности работы классического и разработанной модификации метода SAD используется максимум корреляционной функции Пирсона между сгенерированной картой и подлинной (R_M).

Результаты анализа стереоизображений двумя методами представлены на рис. 2 и в таблице. Модифицированный *SAD* имеет ряд преимуществ: выше детализация, аккуратная обработка однородных областей (пример – правый верхний угол), а качество карты увеличивается более, чем на 10%. К недостаткам следует отнести возможные дополнительные искажения (в примере часть стола была соотнесена статуэтке), а также относительно высокую вычислительную сложность.

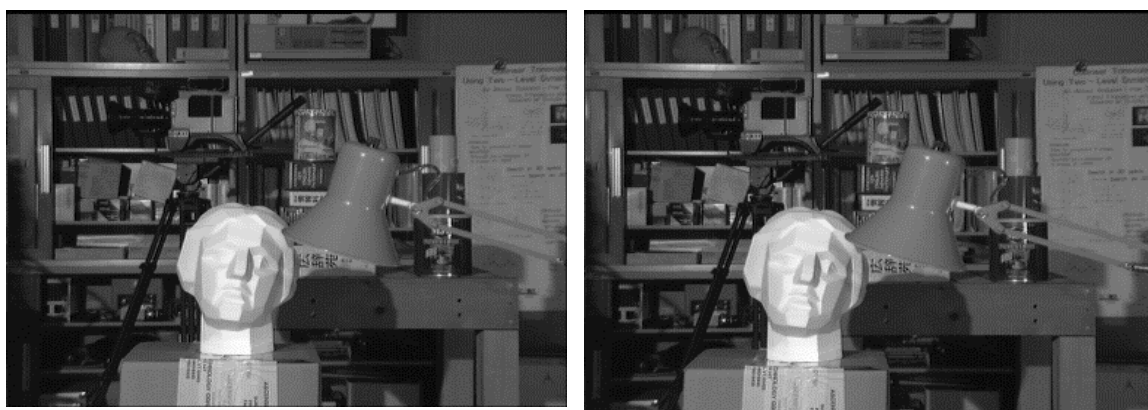


Рис. 1. Пример стереопары (левое и правое изображения)



Рис. 2. а) Визуализация порогов по сегментам (чем выше порог, тем ярче цвет прямоугольника); б) карта глубины полученная классическим *SAD*, в) полученная модифицированным методом *SAD*

Сравнительная характеристика результатов работы методов

Метод	Пример 1		Пример 2		Пример 3	
	Время, с	R_M	Время, с	R_M	Время, с	R_M
Классический метод	3.5	0.40	3.2	0.49	3.7	0.42
Модификация	5.7	0.47	14	0.55	22.1	0.49

ЗАКЛЮЧЕНИЕ

В работе предложена модификация метода *SAD*. Разработанная модификация имеет следующие преимущества в сравнении с классическим *SAD*: позволяет анализировать части изображений низкого

контраста, увеличивает качество карты глубины (при небольших окнах более, чем на 10%) и ее детализацию.

Библиографические ссылки

1. Multi-object detection and tracking by stereo vision / L. Cai [et al] // Pattern Recognition. 2010. Vol. 43. Issue 12. P. 4028-4041.
2. *Liu Y. Local and Global Stereo Methods / Y. Liu, J. Aggarwal* // Elsevier eBooks. 2005. P. 297–308.
3. ModuleNet: A Convolutional Neural Network for Stereo Vision/ O. Renteria [et al]//Pattern Recognition. 2020. P. 219-228.