

АЛГОРИТМ ОПРЕДЕЛЕНИЯ СХОДСТВА МЕЖДУ ОДНОРОДНЫМИ ОБЛАСТЯМИ НА СТЕРЕОИЗОБРАЖЕНИЯХ

Белорусский государственный университет, Минск, Республика Беларусь

Предложена и разработана модификация алгоритма оконного сопоставления стереоизображений SAD с использованием механизмов адаптации и локализации. Данная модификация решает одну из главных проблем стереозрения – нахождение соответствия областям высокой степени однородности. Ее работоспособность подтверждается примерами обработки наборов стереоизображений.

Введение. С каждым годом увеличивается интерес к робототехнике, появляются новые технологии, позволяющие создавать сложных и multifunctional 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 съемки (часто ее называют картой глубины), что обусловлено наложением объектов и наличием 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, RT, CT (аббревиатуры указывают метрику сходства областей или преобразование, SAD – сумма абсолютных отклонений, SSD – сумма квадратов отклонений, NCC – нормализованная кросс-корреляция, RT – ранговое преобразование, 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 (SAD, SSD, NCC) 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 в настоящее время считается метод на основе нейронных сетей (к примеру, ModuleNet). [3]

В работе предложена модификация метода SAD, основанная на использовании механизмов 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 (равномерности) 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 половины ширины изображения. Для дальнейшей локализации области поиска алгоритм 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

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

где I_1, I_2 – первое и второе изображение стереопары, $\tau \in [-N/2; N/2]$ – компенсация сдвига камеры, \bar{I}_1 и \bar{I}_2 – средние значения яркости пикселей на изображениях.

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

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

где LC – некоторый весовой коэффициент.

Формула (2) подобрана экспериментально, $\operatorname{argmax}(R(\tau))$ является оценкой среднего отклонения объектов на стереопаре, $\max(R(\tau))$ характеризует разброс отклонений. Весовой коэффициент экспериментально подобран равным $LC = 2.2$.

3) Локализовать пороги по СКО. Для этого левое изображение делится на прямоугольные сегменты. Для каждого сегмента вычисляется его СКО, порог же определяется по следующей формуле:

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

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

Формула (3) также была получена экспериментально – квадратный корень необходим для усиления ограничения растяжения областей в сегментах достаточного контраста. Весовой коэффициент экспериментально подобран равным $AC = 2.1$.

4) Найти соответствия для областей изображения слева на правом: для каждой области первого изображения вдоль эпполярной линии (по горизонтали) определить совпадения 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 или не произойдет выход за пределы изображения. Превышение порога с высокой вероятностью означает, что область распространена на объект, которому она не принадлежит, для устранения чего область уменьшается до своей величины на предыдущей итерации.

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

Вычислительная сложность. Пусть изображение стереопары имеет стороны длины N пикселей. Классический метод SAD имеет асимптотическую сложность $O(N^3)$. Сложность разработанного алгоритма не превышает $O(N^4)$, так как увеличенные области максимум могут быть сравнимы по ширине с шириной изображения. Однако на практике, для изображений 400×500 время вычисления модифицированным методом в среднем всего в 3 раза дольше классического алгоритма.

Результаты. В работе рассмотрены две стереопары с наличием однородных областей. Метод был применен со следующими параметрами: количество сегментов – 25, размер окна сканирования – 7. Для оценки эффективности работы классического SAD и разработанной модификации метода SAD рассмотрены критерии: визуальный анализ, время вычислений и оценка качества карты. Оценка качества карты – максимум ее корреляционной функции R_m , на основе корреляции Спирмена с истинной картой. Такая мера наименее чувствительна к

Секция 2. Прикладные проблемы информатики

различиям в шкалировании значений карты. Карта тем точнее, чем ближе мера качества к единице.

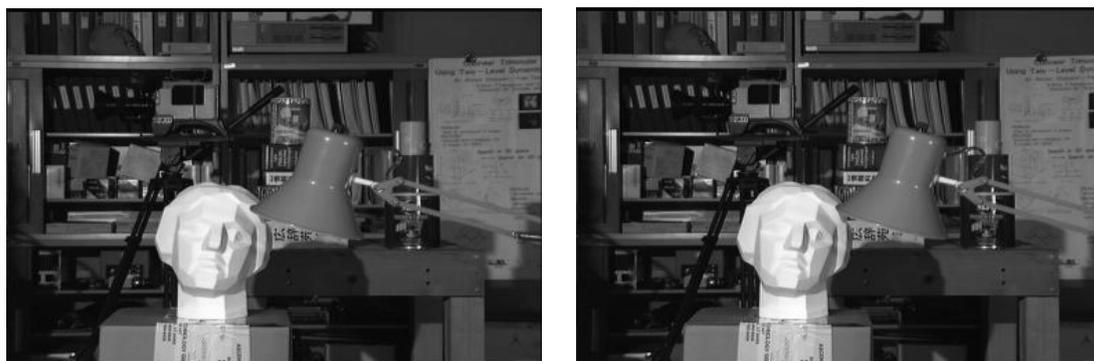


Рисунок 1 – Пример стереопары (левое и правое изображения)

Результаты анализа стереоизображений двумя методами представлены на рисунке 2 и в таблице 1. Модифицированный SAD имеет ряд преимуществ: выше детализация, меньший уровень шума, аккуратная обработка однородных областей (пример – правый верхний угол), в среднем точность увеличивается более, чем на 20%. К недостаткам следует отнести возможные дополнительные искажения (в примере внеслись в виде принятия части стола за часть статуэтки), а также более высокую вычислительную сложность.

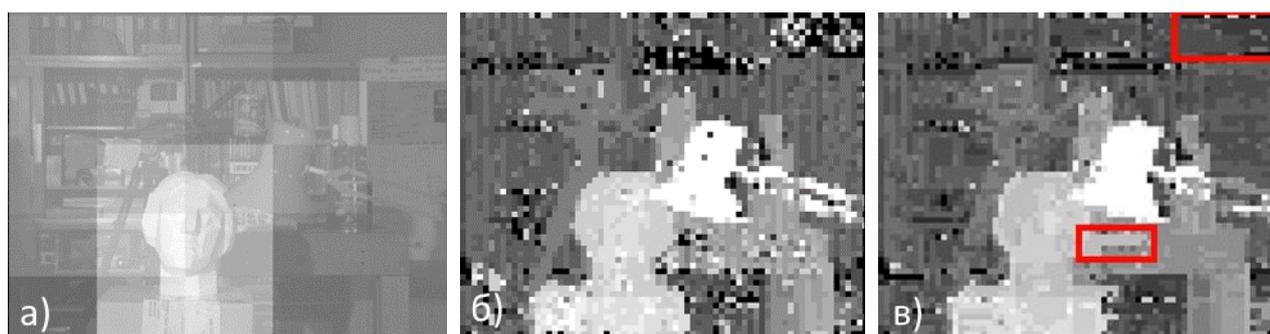


Рисунок 2 – а) Визуализация порогов по сегментам (чем выше порог, тем ярче цвет прямоугольника), б) карта глубины полученная классическим SAD, в) полученная модифицированным методом SAD.

Таблица 1

Сравнительная характеристика результатов работы методов на двух примерах

Метод	Пример 1		Пример 2	
	Время, с	R_m	Время, с	R_m
Классический метод SAD	3.5	0.42	4.8	0.47
Модификация	5.7	0.51	18.3	0.55

Заключение. В работе предложена модификация метода SAD, основанная на использовании механизмов адаптации и локализации. Разработанная модификация имеет следующие преимущества в сравнении с классическим SAD: позволяет анализировать части изображений низкого контраста, увеличивает качество карты (в среднем более, чем на 20%) и ее детализацию.

Список литературы

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