

# МЕТОД ОПТИМИЗАЦИИ ЭНЕРГИИ С ПОМОЩЬЮ МЕХАНИЗМА ПЕРЕДАЧИ СООБЩЕНИЙ В ЗАДАЧАХ СТЕРЕОЗРЕНИЯ

*А.В. Аргутин*

## ENERGY OPTIMIZATION METHOD BASED ON MESSAGE PASSING MECHANISM APPLIED FOR STEREO-VISION PROBLEMS

*A. V. Argutin*

Рассматривается алгоритм оптимизации функции энергии вдоль сканлайна, его особенности и способ реализации с использованием механизма передачи сообщений. В проведенном обзоре рассмотрен метод поиска минимума функции энергии, а также метод поиска аргумента, при котором достигается это значение.

*Ключевые слова:* стереозрение, алгоритмы на графах, минимизация энергии, передача сообщений.

In the article the scanline energy optimization algorithm, its characteristics and implementation method with the help of message passing mechanism is considered. The method of searching for energy function minimum and the method of searching for the argument at which this minimum occurs have been observed in the review.

*Keywords:* stereo-vision, graph algorithms, energy minimization, message passing.

### Введение

Большинство современных научных работ по стереозрению для построения математической модели используют два предположения о структуре наблюдаемой сцены [1]:

- 1) предположение о *кусочной гладкости*;
- 2) предположение о *сегментированности*.

Первое утверждает, что соседние пиксели имеют схожие значения *диспаритета* (разницы между положением объекта на левом и правом изображениях), поскольку объекты сцены являются кусочно-гладкими. Второе предположение гласит о том, что однотонным участкам сцены соответствуют плоские поверхности в 3D.

### Общая оценка методов оптимизации энергии

Согласно терминологии [2] распространенным подходом к решению задачи стереозрения является метод *глобальной оптимизации*, основанный на поиске аргумента, при котором некая функция энергии принимает минимальное значение. В наиболее общем случае функция энергии может быть записана следующим образом:

$$E(d) = E_{data}(d) + \lambda E_{smooth}(d), \quad (1)$$

где  $E_{data}(d)$  есть функция, описывающая штраф, накладываемый на пиксели левого и правого изо-

бражения, приводимые в соответствие текущим значением диспаритета  $d$ , в зависимости от того, насколько они различаются по значению цвета, интенсивности, либо в зависимости от степени их соответствия более сложным критериям [3];  $E_{smooth}(d)$  – функция, описывающая штраф, накладываемый на пару пикселей, приводимых в соответствие текущим значением диспаритета  $d$ , если для них нарушается условие непрерывности функции диспаритета.

Поскольку минимизация функции (1) в общем случае является NP-сложной задачей, существует подход, значительно упрощающий минимизацию функции энергии и переводящий функцию  $E_{smooth}(d)$  из пространства 2D в пространство 1D. Другими словами, данный подход учитывает кусочную непрерывность функции диспаритета только в горизонтальном направлении. Такие методы оптимизации функции энергии в литературе по стереозрению [4–6] принято называть методами «Оптимизации вдоль *сканлайна*» (сканлайном называется пара строк с одним индексом, полученных со стереопары). Функцию энергии в таком случае можно записать отдельно для каждого сканлайна в отличие от единой функции энергии для изображения со стереопары в 2D-методах глобальной оптимизации:

Аргутин Александр Вячеславович – аспирант кафедры ЭВМ, Южно-Уральский государственный университет; alex.argutin@gmail.com

Argutin Aleksandr Vyacheslavovich – Post-Graduate Student of Electronic Computer Department of South Ural State University; alex.argutin@gmail.com

$$E(d, y_0) = E_{data}(d, y_0) + \lambda E_{smooth}(d, y_0), \quad (2)$$

где  $y_0$  – индекс сканлайна, для которого осуществляется оптимизация.



Рис. 1. Горизонтальные артефакты на карте глубины

Положительными качествами таких методов являются простота реализации, быстродействие и эффективность, однако общим недостатком является наличие горизонтальных артефактов, представляющих собой шум, ошибку работы алгоритма на результирующей карте глубины (рис. 1), вызванную упрощением (2) изначальной функции энергии (1).

### Оптимизация энергии с помощью механизма передачи сообщений

В данной статье рассматривается один из методов оптимизации вдоль сканлайна, использующий механизм передачи сообщений. Вид функций  $E_{data}(d)$  и  $E_{smooth}(d)$  не рассматривается в данной статье, однако имеет смысл сообщить, что в качестве функции  $E_{data}(d)$  может применяться один из существующих методов формирования цены сопоставления, например SAD, SSD, NAC, VT [3].

Функция  $E_{smooth}(d)$  может иметь следующий вид:

$$E_{smooth}(d_i, d_{i+1}) = \text{sign}(d_i - d_{i-1}), \quad (3)$$

$$E_{smooth}(d_i, d_{i+1}) = \frac{(d_i - d_{i-1})^2}{K}. \quad (4)$$

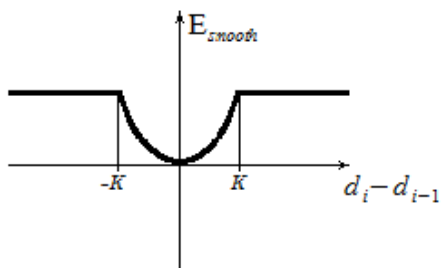


Рис. 2. Вид функции  $E_{smooth}(d)$

Следует отметить, что в случае (3) любое изменение (в том числе плавное) величины диспаритета будет восприниматься алгоритмом как резкий скачок. Это означает, что изменение величины диспаритета на 1 облагается штрафом за нарушение условия непрерывности функции диспаритета в той же мере, что и изменение величины диспаритета на 10. В каких-то случаях такое условие может быть слишком строгим, поэтому имеет смысл применять функцию вида (4) (рис. 2).

Принцип действия одного из методов оптимизации вдоль сканлайна описывается ниже. Для данного метода удобной математической сущностью для представления сканлайна является граф, вершины которого – потенциально возможные соответствия между пикселем на изображении со стереопары и величиной диспаритета. При такой формулировке задача стереосопоставления может быть преобразована к задаче поиска кратчайшего пути на графе, в котором длина пути определяется функцией энергии в целом и ценой сопоставления  $E_{data}(d)$  совместно с критерием кусочной непрерывности функции диспаритета, определяемым функцией  $E_{smooth}(d)$  в частности (рис. 3).

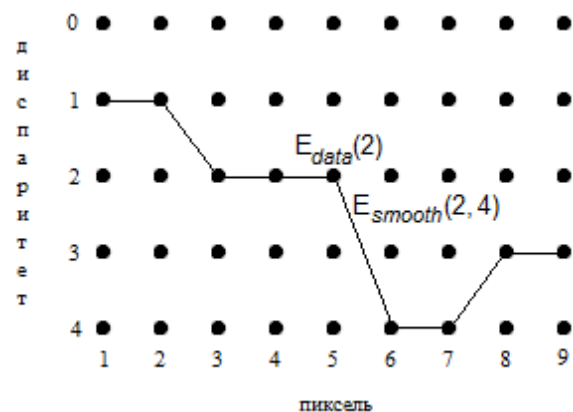
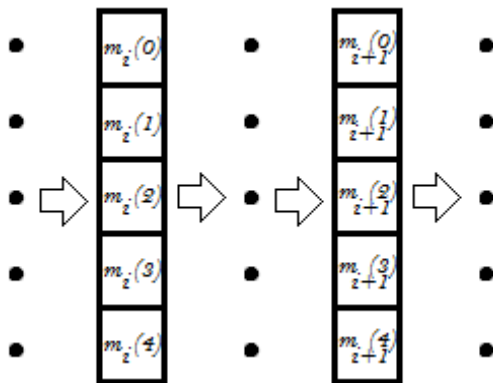


Рис. 3. Поиск кратчайшего пути на графе

Величина функции энергии для всего сканлайна может быть записана следующим образом:

$$E(d) = E_{data}(d_1) + E_{smooth}(d_1, d_2) + E_{data}(d_2) + E_{smooth}(d_2, d_3) + \dots + E_{data}(d_N). \quad (5)$$

В данном случае  $E_{data}(d_i)$  является длиной пути, добавляемой текущей вершиной графа, а  $E_{smooth}(d_i, d_{i+1})$  – длиной пути, добавляемой текущим ребром графа. Такая запись функции энергии позволяет сделать предположение, что поиск кратчайшего пути на данном графе может быть реализован с помощью механизма *передачи сообщений*. В таком случае *сообщение* является информацией о длинах пути по графу от начала сканлайна до текущего пикселя для всех возможных значений величины диспаритета, передаваемой от пикселя к пикселю (рис. 4).



**Рис. 4. Поиск кратчайшего пути с помощью механизма передачи сообщений**

Сообщение, передаваемое от пикселя  $i-1$  к пикселю  $i$ , может быть записано следующим образом:

$$D$$

$$m_i^L(k) = \min_{j=1} (m_{i-1}(j) + E_{data}(j) + E_{smooth}(k, j)). \quad (6)$$

Для того чтобы минимизировать функцию энергии  $E(d)$ , достаточно применить механизм передачи сообщений в одном направлении для графа сканлайна, на котором производится минимизация. Сообщение, переданное в последний пиксель, будет хранить в себе  $D$  (по количеству допустимых значений диспаратета) кратчайших длин путей для всех анализируемых величин диспаратета. Минимальное значение, передаваемое в этом сообщении, является кратчайшей длиной пути по графу и одновременно минимальным значением функции энергии для всего сканлайна. В дальнейшем эта величина будет называться  $E_{min}$ .

Однако минимизация функции энергии не является решением задачи стереозрения, необходимо знать аргумент функции энергии, при котором функция принимает минимальное значение, другими словами, необходимо вычислить сам минимальный путь, а не только его длину. В [7] для решения данной проблемы предлагается произвести процесс передачи сообщений в двух направлениях. Тогда к уравнению (6) можно добавить уравнение, выражающее сообщение, приходящее в пиксель  $i$  из пикселя  $i+1$ :

$$D$$

$$m_i^R(k) = \min_{j=1} (m_{i+1}(j) + E_{data}(j) + E_{smooth}(k, j)). \quad (7)$$

В таком случае формула (6) выражает кратчайший путь до текущего пикселя от левого края сканлайна, а формула (7) – кратчайший путь до текущего пикселя от правого края сканлайна. При сложении этих двух величин со значением  $E_{data}(x_i, k)$  для данного пикселя получается функция энергии в целом:

$$E(d) = m_i^L(d) + E_{data}(d) + m_i^R(d). \quad (8)$$

Таким образом, для каждого пикселя из всех допустимых значений диспаратета становится возможным выбрать то, при котором значение энергии, получающееся в (8), равно минимальному значению функции энергии  $E_{min}$ . Значения диспаратета в совокупности с номерами пикселей, которым они соответствуют, и формируют кратчайший путь по графу.

#### **Литература**

1. Wang, L. *Global stereo matching leveraged by sparse ground control points* / L. Wang, R. Yang // *CVPR*. – 2011. – P. 3033–3040.
2. Scharstein, D. *A taxonomy and evaluation of dense two-frame stereo correspondence algorithms* / D. Scharstein, R. Szeliski // *IJCV*. – 2002. – Vol. 47, no. 1–3. – P. 7–42.
3. Hirschmüller, H. *Evaluation of Cost Functions for Stereo Matching* / H. Hirschmüller, D. Scharstein // *IEEE Conference on Computer Vision and Pattern Recognition*. – 2007. – Vol. 1. – P. 1–8.
4. Birchfield, S. *Depth discontinuities by pixel to-pixel stereo* / S. Birchfield, C. Tomasi // *IJCV*. – 1999. – Vol. 35, no. 3. – P. 269–293.
5. Ohta, Y. *Stereo by intra- and inter-scanline search* / Y. Ohta and T. Kanade // *TPAMI*. – 1985. – Vol. 7, no. 2. – P. 139–154.
6. Cox, I. *A maximum likelihood stereo algorithm* / I. Cox, S. Hingorani, S. Rao, and B. Maggs // *Computer Vision, Graphics and Image Processing*. – 1996. – Vol. 63, no. 3. – P. 542–567.
7. *Оптимизация энергии в задачах компьютерного зрения и алгоритмы на графах, лекция 1* [Электронный ресурс]. – <http://www.lektorium.tv/lecture/?id=12997>

**Поступила в редакцию 27 ноября 2012 г.**