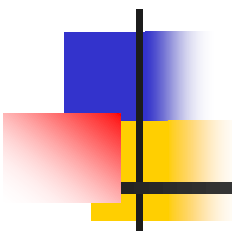


Некоторые методы сопоставления стерео



Алексей Федоров

Video Group
CS MSU Graphics & Media Lab

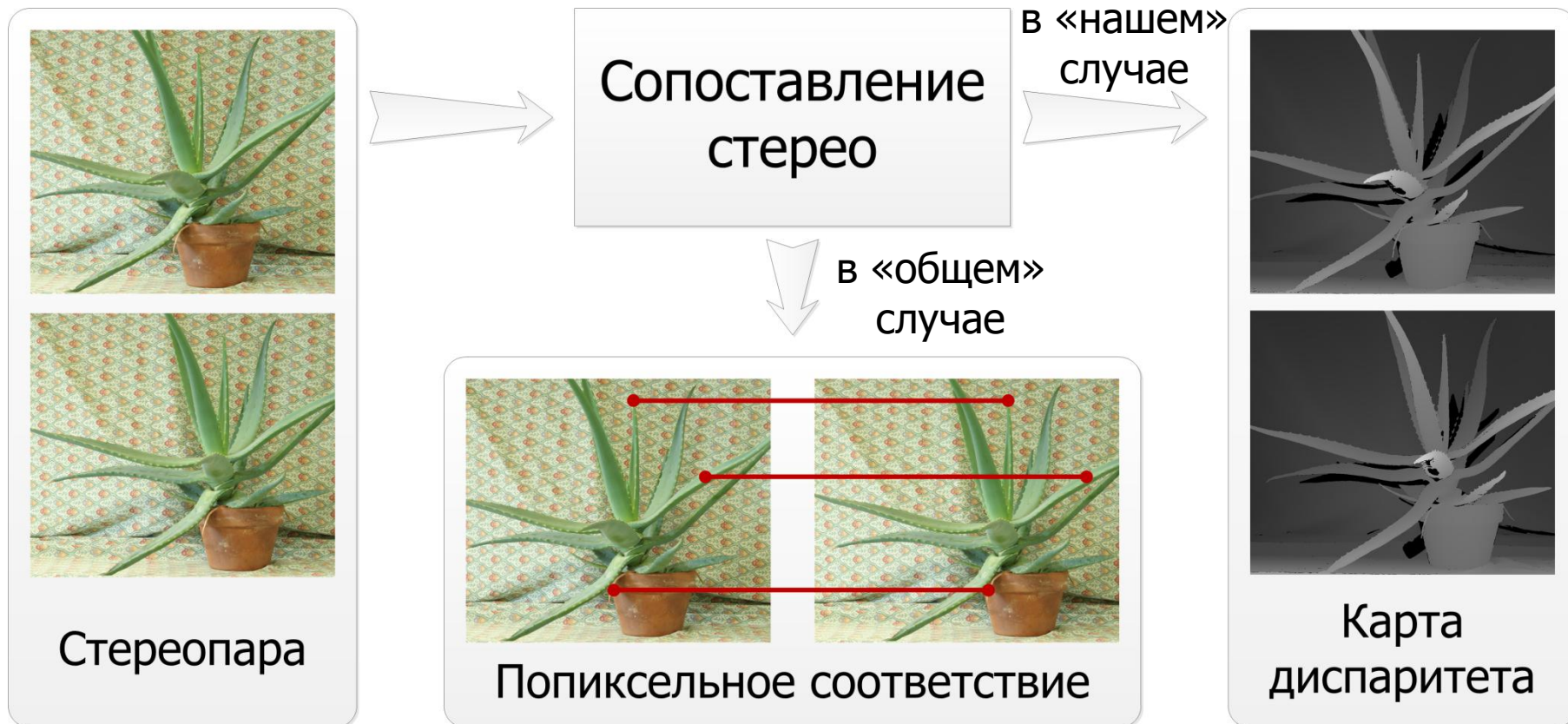


Содержание

- **Введение**
- Первичная обработка
 - Нелинейная ректификация стерео
- Методы сопоставления стерео
 - Bidirectional Stereo Matching with Occlusion Handling
 - Building an Accurate Stereo Matching System
- Заключение

Введение

Задача сопоставления стерео

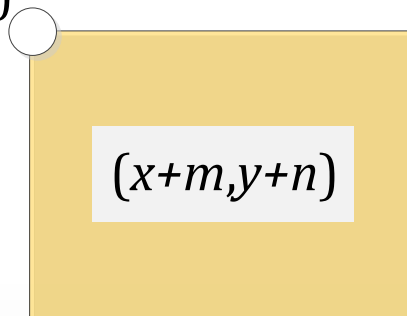


Сопоставление блоков

Метрики для блоков (1)

Левый ракурс (I_l)

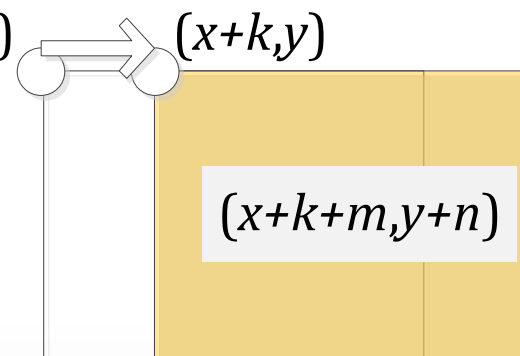
(x, y)



Блок для
сравнения

Правый ракурс (I_r)

(x, y)



Блок для
сравнения

k – диспаритет

Стандартный подход

Метрики для блоков (2)

- $SAD(I_r(x, y), I_l(x + k, y)) =$

$$\sum_{m, n \in W} |I_r(x + m, y + n) - I_l(x + m + k, y + n)|$$

- $SSD(I_r(x, y), I_l(x + k, y)) =$

$$\sum_{m, n \in W} (I_r(x + m, y + n) - I_l(x + m + k, y + n))^2$$

- $SM_{Fuzzy_similarity}(I_r(x, y), I_l(x + k, y)) =$

$$\frac{1}{m \times n} \sum_{m, n \in W} \left(1 - \frac{|I_r(x + m, y + n) - I_l(x + m + k, y + n)|}{\alpha} \right), \quad \alpha = 16$$

Сопоставление стерео

Частые проблемы

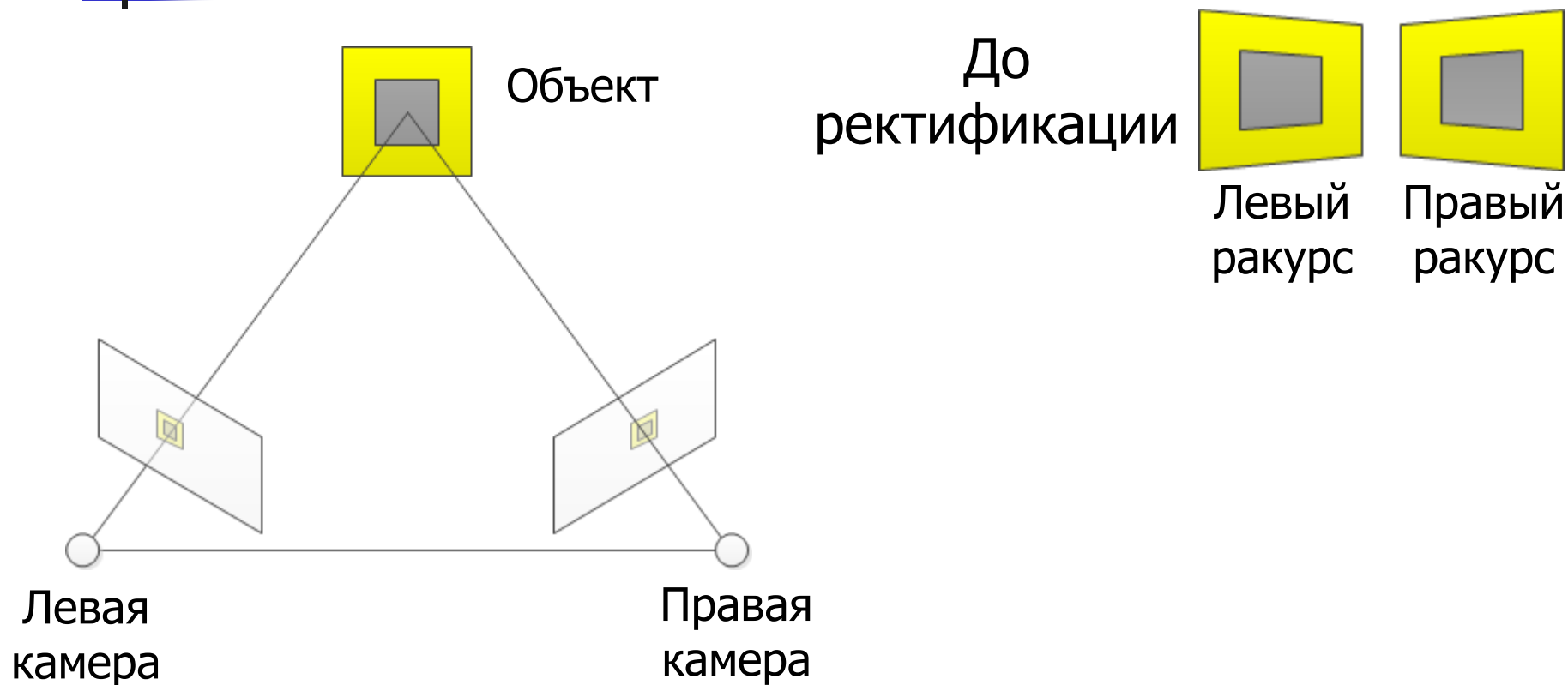
Возможны искажения изображений
на разных ракурсах

- по яркости
- по контрасту
- по фокусу
- по геометрии

Метрики, устойчивые
к расхождениям, рассмотрел
Нападовский Вячеслав

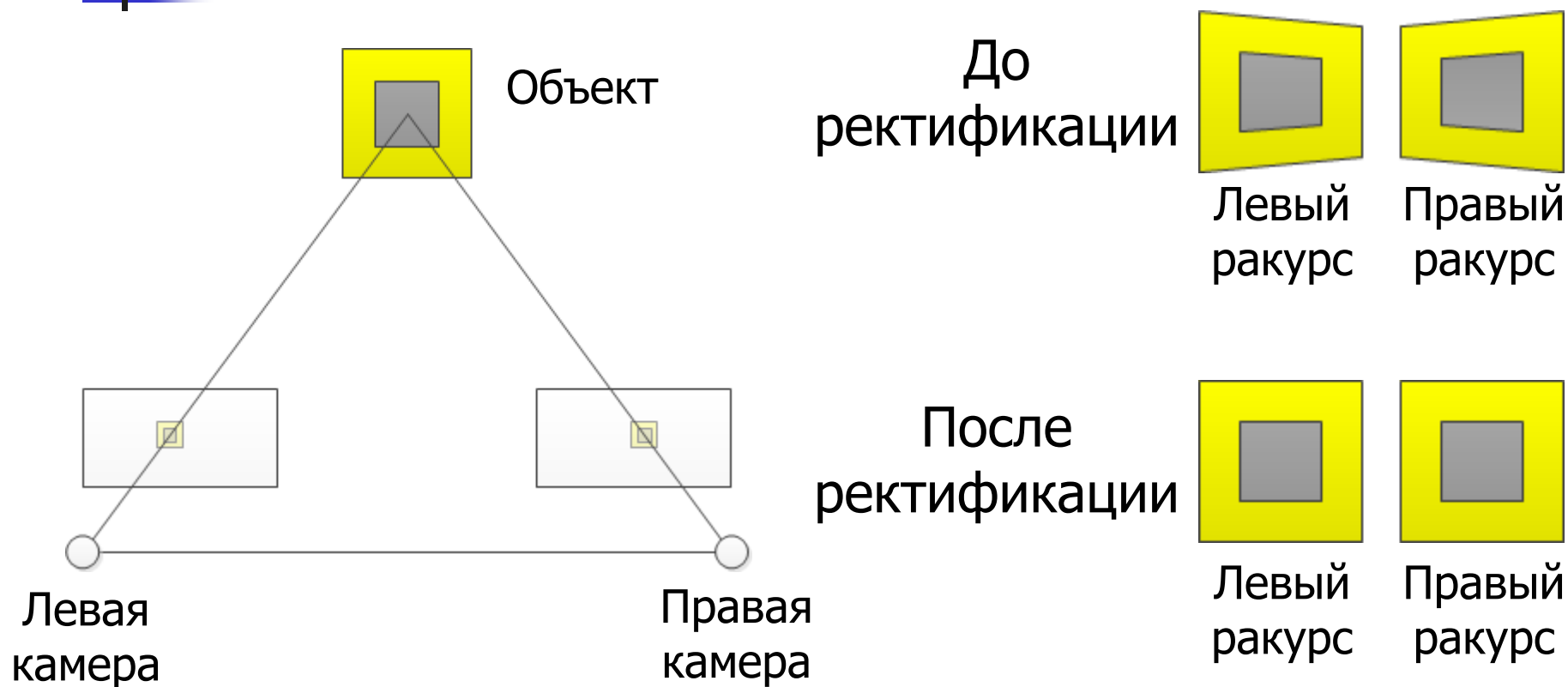
Пример

Ректификация изображений (1/2)



Пример

Ректификация изображений (2/2)



Содержание

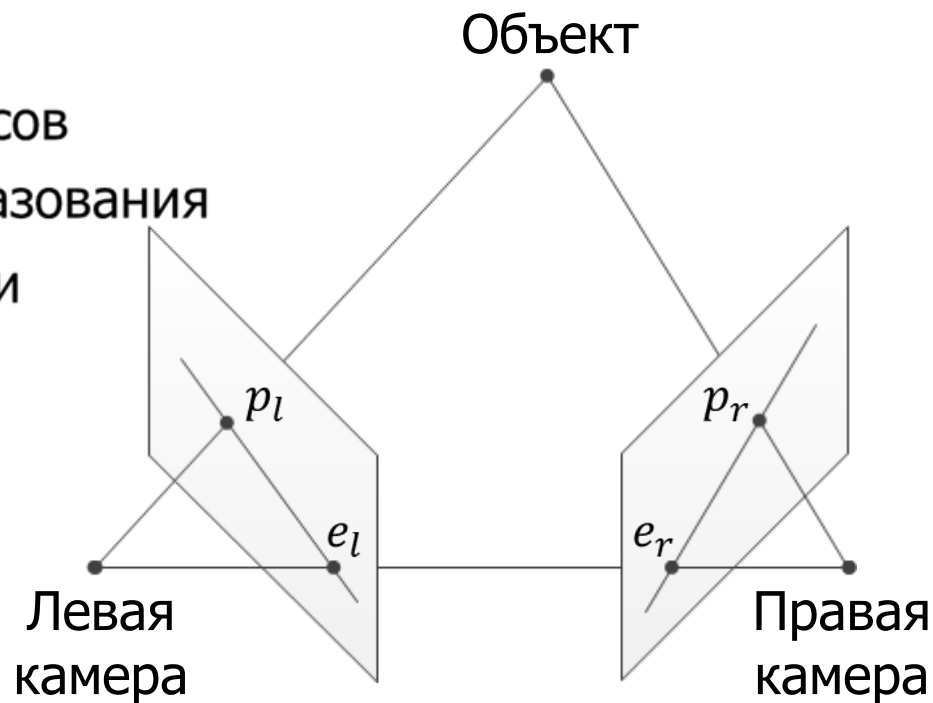
- Введение
- Первичная обработка
 - **Нелинейная ректификация стерео**
- Методы сопоставления стерео
 - Bidirectional Stereo Matching with Occlusion Handling
 - Building an Accurate Stereo Matching System
- Заключение

Ректификация стерео

Постановка задачи

$$I_{l(r)'} = H_{l(r)} I_{l(r)}$$

- I_l, I_r – исходные ракурсы
- H_l, H_r – преобразование ракурсов
- I_l', I_r' – ракурсы после преобразования
- p_l, p_r – соответствующие точки на разных ракурсах
- e_l, e_r – эпиполярные точки



Поиск преобразования

Минимизируемый функционал

$$\sum_i \left(\underbrace{A_l R_l G_l T_l p_{li}}_{H_l} - \underbrace{R_r G_r T_r p_{ri}}_{H_r} \right)^2$$

- $T_{l(r)}$ – компонента смещения
- $G_{l(r)}$ – эпиполярная компонента
- $R_{l(r)}$ – компонента вращения
- A_l – компонента вертикального соответствия
- i – номер сопоставленной точки

Поиск преобразования

Компонента смещения

$$\sum_i \left(\underbrace{A_l R_l G_l T_l}_{H_l} p_{li} - \underbrace{R_r G_r T_r}_{H_r} p_{ri} \right)^2$$

Выбираем по одной точке на каждом ракурсе, относительно которых будем строить преобразование

$$T_{l(r)} = \begin{pmatrix} 1 & 0 & -p_{cl(r)x} \\ 0 & 1 & -p_{cl(r)y} \\ 0 & 0 & 1 \end{pmatrix}$$

p_{cl} , p_{cr} – выбранные точки (обычно центральные)

Поиск преобразования

Эпиполярная компонента

$$\sum_i (A_l R_l \underbrace{G_l T_l}_{H_l} p_{li} - R_r \underbrace{G_r T_r}_{H_r} p_{ri})^2$$

Эпиполяры уводим на бесконечность.

Теперь прямые на обоих ракурсах параллельны.

$$G_{l(r)} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -\frac{e_{l(r)x}}{e_{l(r)x}^2 + e_{l(r)y}^2} & -\frac{e_{l(r)y}}{e_{l(r)x}^2 + e_{l(r)y}^2} & 1 \end{pmatrix}$$

Поиск преобразования

Компонента вращения

$$\sum_i (A_l \underbrace{R_l G_l T_l}_{H_l} p_{li} - \underbrace{R_r G_r T_r}_{H_r} p_{ri})^2$$

$$R_{l(r)} = \begin{pmatrix} \cos(\theta_{l(r)}) & -\sin(\theta_{l(r)}) & 0 \\ \sin(\theta_{l(r)}) & \cos(\theta_{l(r)}) & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

θ_l – выбирается вручную

$$\theta_r = \theta_l - \arctan(e_{ry}/e_{rx}) + \arctan(e_{ly}/e_{lx})$$

Поиск преобразования

Компонента вертикального соответствия

$$\sum_i \left(\underbrace{A_l R_l G_l T_l p_{li}}_{H_l} - \underbrace{R_r G_r T_r p_{ri}}_{H_r} \right)^2$$

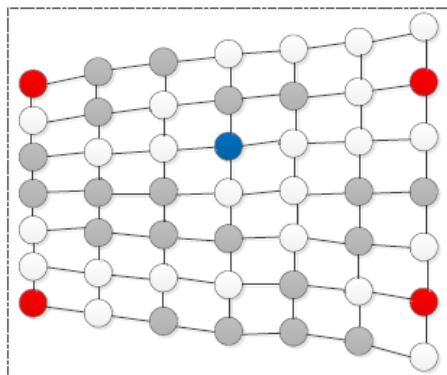


$$A_l = \begin{pmatrix} 1 & 0 & 0 \\ a & b & c \\ 0 & 0 & 1 \end{pmatrix}$$

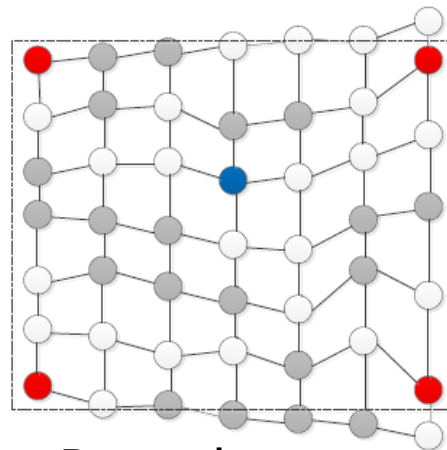
a, b, c – параметры

Задача сведена к поиску линейных параметров
методом наименьших квадратов

И что же такого в этой статье?



Ректификация



Ректификация

- Ожидаемый край
- Выбранная точка
- Заметная точка
- Менее заметная точка

без учета заметности с учетом заметности

Степень заметности формируется из следующей информации:

- Резкость края
- Скорость движения
- Результат Face Detection

Нелинейная ректификация

Учет заметности

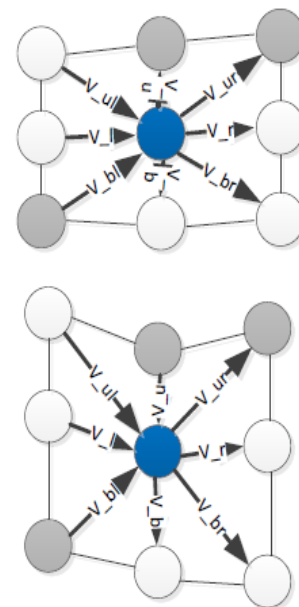


Человеческое восприятие чувствительно к искажению относительно соседних пикселей. Введем вектор относительной позиции:

$$\check{H}(p_0 - p_i) = H(p_0 - p_1),$$

$$i = 1, 2, \dots, 8$$

p_i – соседи p_0



● Выбранная точка

● Заметная точка

○ Менее заметная точка

Результаты

Сравнение с конкурентами (1/3)



Исходная стереопара

Результаты

Сравнение с конкурентами (2/3)



Предложенный метод

Результаты

Сравнение с конкурентами (3/3)



Результат конкурента



Выводы

Достоинство:

- Учет степени заметности

Недостатки:

- Высокая требовательность к ресурсам (Real-time на GPU)
- Возможность искажения прямых на Т-образных перекрестках



Содержание

- Введение
- Первичная обработка
 - Нелинейная ректификация стерео
- Методы сопоставления стерео
 - **Bidirectional Stereo Matching with Occlusion Handling**
 - Building an Accurate Stereo Matching System
- Заключение

Bidirectional Stereo Matching



Особенности метода

- Используется двунаправленное сопоставление стерео
- Использование Graph-Cut для поиска диспаратета
- Учитываются области открытия (закрытия)

Поиск диспаратета

Graph Cut

V – множество вершин, E – множество ребер

$G(V, E)$ – граф

$C \subset E$ – разрез графа. Подмножество ребер, которое приводит граф $G'(V, E/C)$ к индуцированному виду

$|C|$ – сумма весов ребер, входящих в C

Задача: найти разрез с наименьшей стоимостью среди всех возможных разрезов

Альфа-расширение

ALPHA EXPANTION[3]

1. Start with an arbitrary labeling f'
2. Repeat
 - For each label $a \in L$
 - 2.1 Find $f^a = \arg \min C(f)$ where f is an α -expansion of f'
 - 2.2 if $C(f^a) < C(f')$, set $f' := f^a$
3. Until converged

f – диспаратет

Минимизируемая функция стоимости

$$C(f) = C_{data}(f) + C_{smoothness}(f) + C_{visibility}(f) + C_{occluded}(f)$$

- C_{data} – компонента интенсивности
- $C_{smoothness}$ – компонента сглаженности
- $C_{visibility}$ – компонента видимости
- $C_{occluded}$ – компонента областей открытия (закрытия)

ФУНКЦИЯ СТОИМОСТИ

Компонента интенсивности

$$C(f) = \underline{C_{data}(f)} + C_{smoothness}(f) + C_{visibility}(f) + C_{occluded}(f)$$

$$C_{data}(f) = \sum_{\{(p,l)\}, \{(q,l)\} \in I} D_{p,q,l}(f_p, f_q)$$

(p, l) – пиксель p левого ракурса, имеющий глубину l

$$D_{p,q,l}(f_p, f_q) = D(p, q) \cdot T[f_p = f_q = l]$$

(q, l) – пиксель q правого ракурса, имеющий глубину l

$$D(p, q) = \frac{1}{3} \sum_{Channel} \frac{\sum_N S_p S_q}{\sqrt{\sum_N S_p^2} \sqrt{\sum_N S_q^2}}$$

T – функция-индикатор

$W_{p(q)(channel)}$ – значение цвета пикселя $p(q)$

$\overline{W_{p(q)(channel)}}$ – среднее значение цвета в окрестности пикселя $p(q)$

$$S_p = W_{p(Channel)} - \overline{W_{p(Channel)}}, S_q = W_{q(Channel)} - \overline{W_{q(Channel)}}$$

Результаты (1)

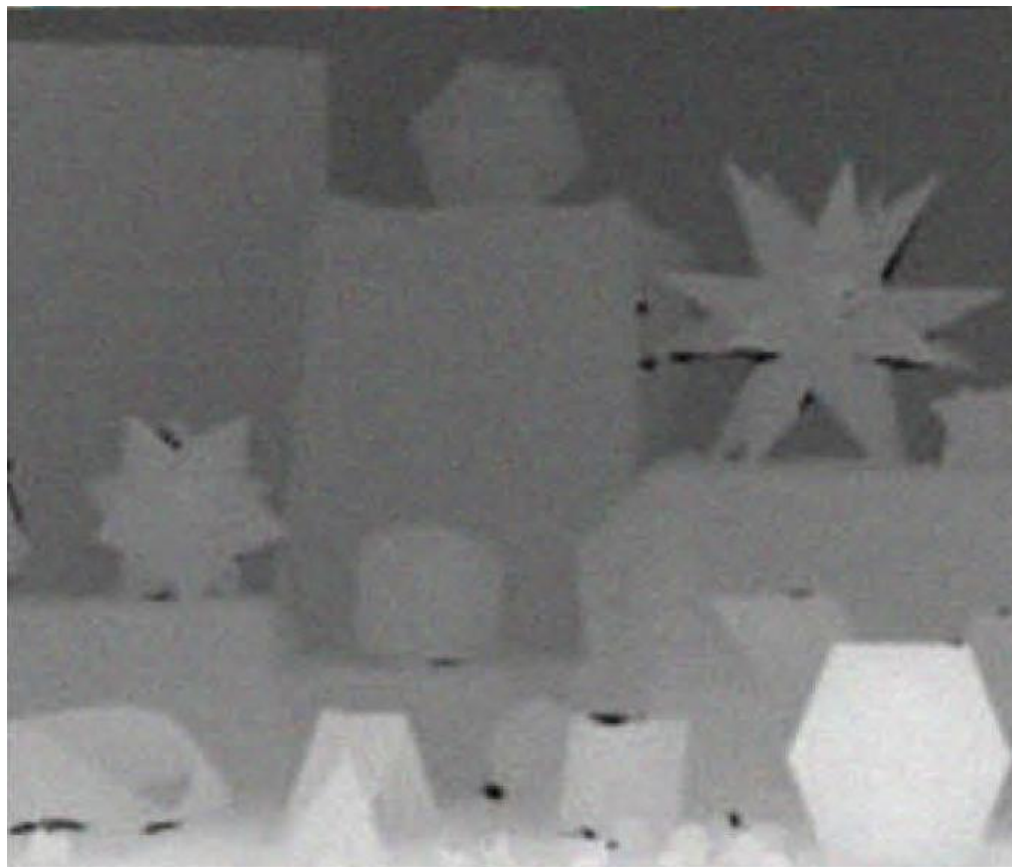
Изображение Моебиуса (1/3)



Исходное изображение

Результаты (1)

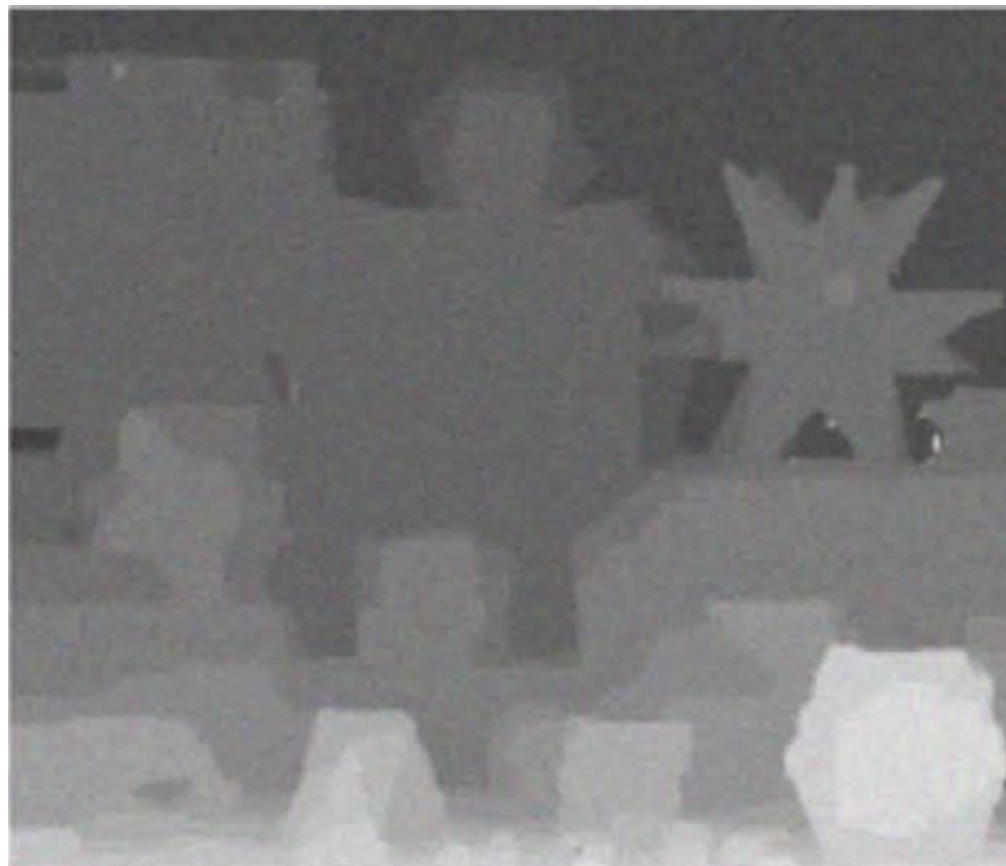
Изображение Моебиуса (2/3)



Ground truth

Результаты (1)

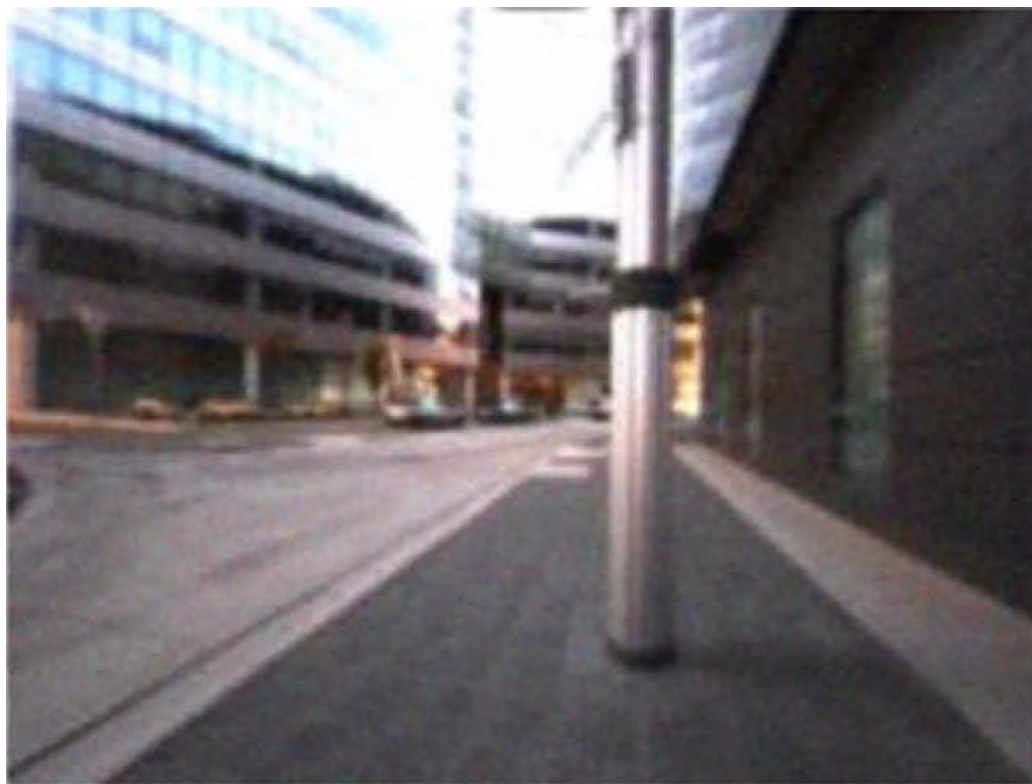
Изображение Моебиуса (3/3)



Предложенный метод

Результаты (2)

Изображение Street (1/2)



Исходное изображение

Результаты (2)

Изображение Street (2/2)



Предложенный метод



Выводы

Достоинства:

- Двухнаправленное сопоставление
- Учет областей открытия

Недостаток:

- Высокая требовательность к ресурсам



Содержание

- Введение
- Первичная обработка
 - Нелинейная ректификация стерео
- Методы сопоставления стерео
 - Bidirectional Stereo Matching with Occlusion Handling
 - **Building an Accurate Stereo Matching System**
- Заключение

Accurate Stereo Matching

Особенности метода

- Используется метрика, основанная на расстоянии Хэмминга
- Используется адаптивная форма сопоставляемых блоков
- Учитывается возможность параллельных вычислений

Метрика отличия блоков

AD-Census

$$C(\mathbf{p}, d) = \rho(C_{census}(\mathbf{p}, d), \lambda_{census}) + \rho(C_{AD}(\mathbf{p}, d), \lambda_{AD})$$

$$\rho(c, \lambda) = 1 - \exp\left(-\frac{c}{\lambda}\right)$$

$$C_{AD}(\mathbf{p}, d) = \frac{1}{3} \sum_{i=R,G,B} |I_i^{Left}(\mathbf{p}) - I_i^{Right}(\mathbf{p}d)|$$

Absolute Differences

C_{census} – метрика, равная расстоянию Хемминга между блоками (используются блоки 9×7)

$\mathbf{p}d = \mathbf{p}(x, y)d = \mathbf{p}(x + d, y)$, \mathbf{p} – пиксель, d – глубина

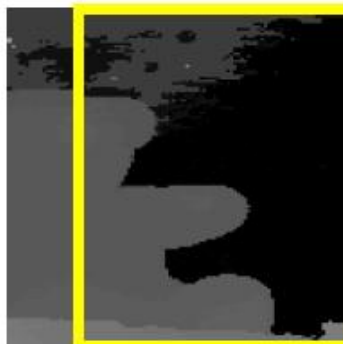
$\lambda_{census}, \lambda_{AD}$ – КОНСТАНТЫ

Сравнение метрик

repetitive
structures



textureless
regions



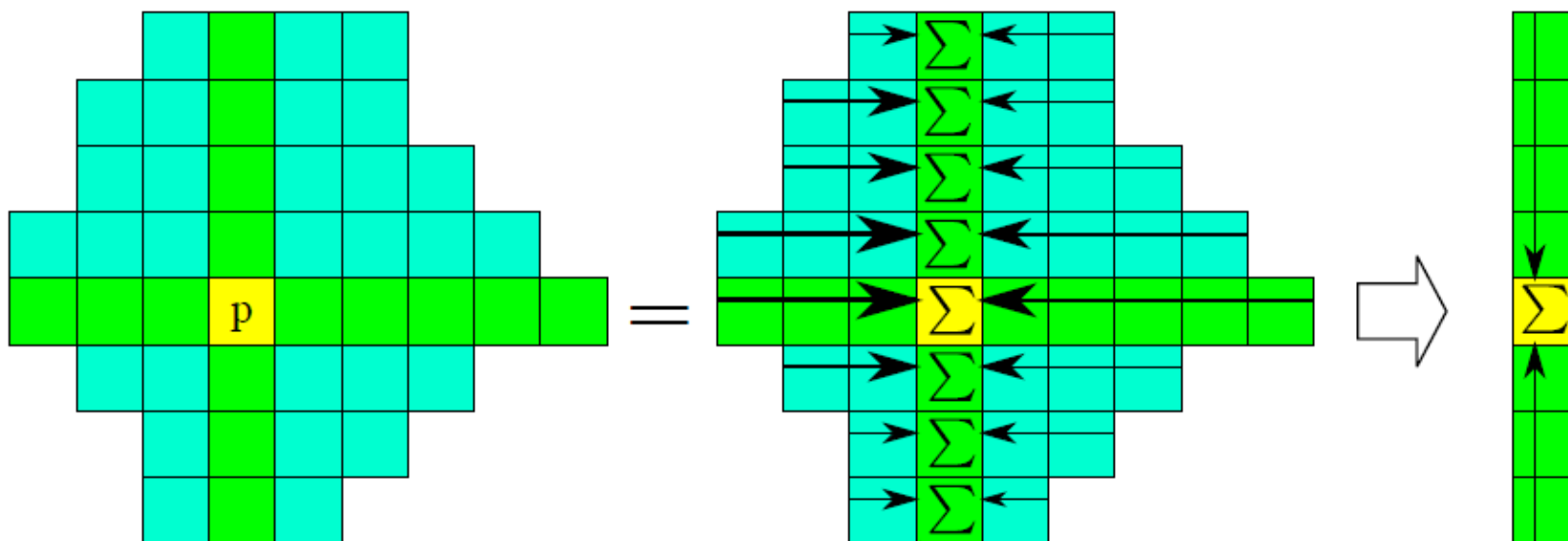
AD

Census

AD-Census

Крестообразные области

Агрегирование стоимости



Пример

Результаты

Изображение Cones (1/3)



Исходное изображение

Результаты

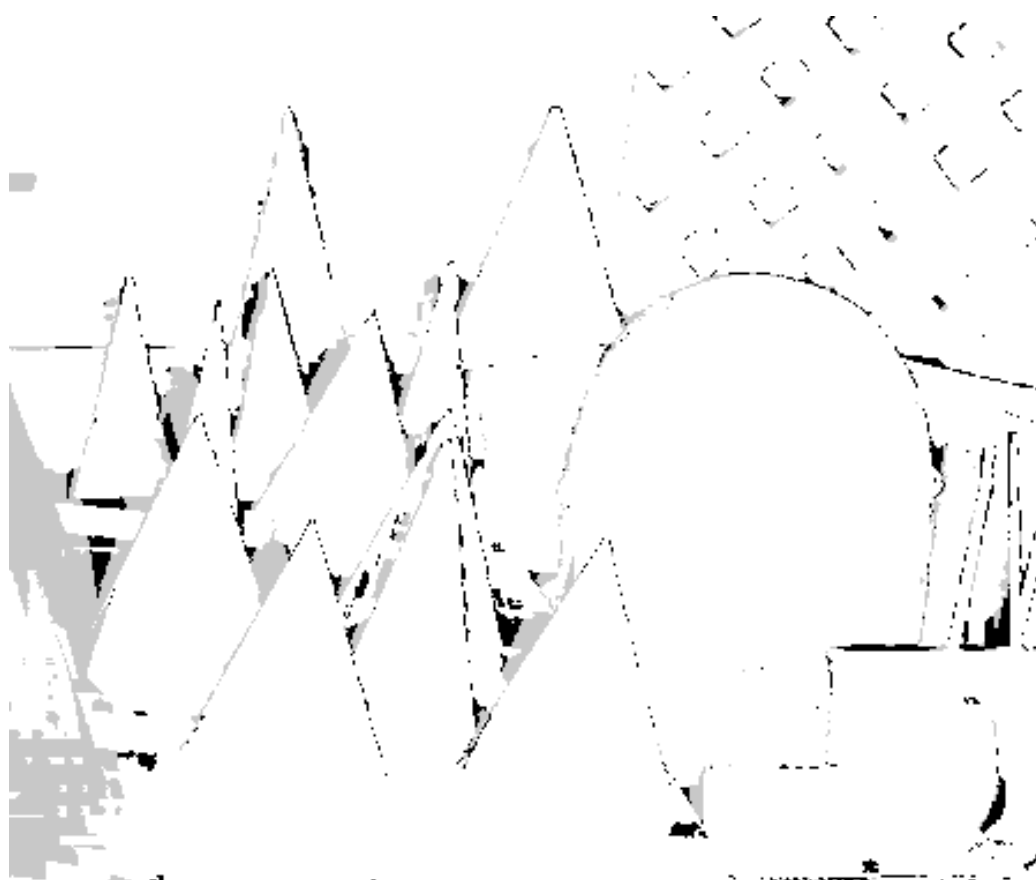
Изображение Cones (2/3)



Предложенный метод

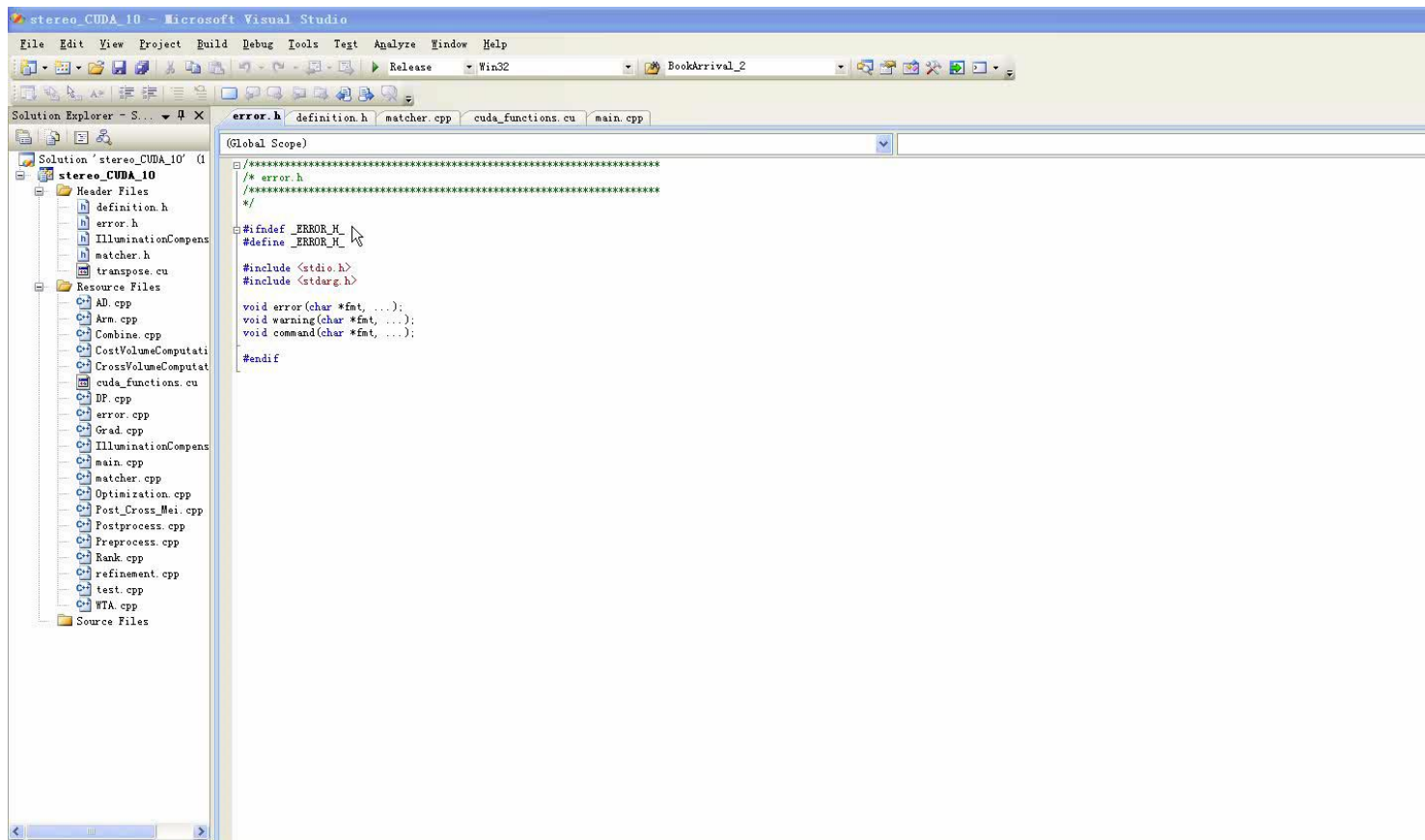
Результаты

Изображение Cones (3/3)



Разница с Ground truth

Результаты Мультик =)



```
error.h
definition.h
matcher.cpp
cuda_functions.cu
main.cpp

(Global Scope)
/* error.h
*/
#ifdef _ERROR_H
#define _ERROR_H

#include <stdio.h>
#include <stdarg.h>

void error(char *fmt, ...);
void warning(char *fmt, ...);
void command(char *fmt, ...);

#endif
```



Выводы

Достоинства:

- Использование метрики AD-Census
- Возможность распараллеливания (реализовать на GPU)

Недостатки:

- Чувствителен к зашумленным данным
- Чувствителен к искажениям по цвету



Содержание

- Введение
- Первичная обработка
 - Нелинейная ректификация стерео
- Методы сопоставления стерео
 - Bidirectional Stereo Matching with Occlusion Handling
 - Building an Accurate Stereo Matching System
- **Заключение**



Заключение

В презентации обозначены следующие идеи, которые можно использовать в дальнейшем:

- Учитывать степень заметности при ректификации
- Использовать двунаправленное сопоставление
- Попробовать метрику AD-Census

Наши результаты 1

Изображение Cones (1/3)



Исходное изображение

Наши результаты 1

Изображение Cones (2/3)



Наш метод

Наши результаты 1

Изображение Cones (3/3)



Метод из статьи Xun Sun (2011)

Наши результаты 2

Изображение Tsukuba (1/3)



Исходное изображение

Наши результаты 2

Изображение Tsukuba (2/3)



Наш метод

Наши результаты 2

Изображение Tsukuba (3/3)



Метод из статьи Xun Sun (2011)

Литература

1. Yu Zhao, Hang Shao, and Qionghai Dai, "Gravity-oriented nonlinear stereo rectification for stereoscopic 3D," in *3DTV Conference: The True Vision – Capture, Transmission and Display of 3D Video (3DTV-CON)*, 2011, pp. 1-4.
2. Shaobo Zhang, Sheng Liu, Yuhang Mao, and Xin Wang, "Global Optimization for Bidirectional Stereo Matching with Occlusion Handling," *International Conference on Measurement, Information and Control (MIC)*, vol. 2, 2012, pp. 553-557.
3. Xun Sun, Mingcai Zhou, Shaohui Jiao, Haitao Wang, and Xiaopeng Zhang, "On Building an Accurate Stereo Matching System on Graphics Hardware," *IEEE International Conference on Computer Vision Workshops (ICCV Workshops)*, 2011, pp. 467-474.
4. Mikel Galar, Miguel Pagola, Edurne Barrenechea, Carlos Lopez-Molina, and Humberto Bustince, "Aggregation of Color Information in Stereo Matching Problem: A Comparison Study," *Lecture Notes in Computer Science*, vol. 6098, 2010, pp. 369–378.

Лаборатория компьютерной графики и мультимедиа



Видеогруппа — это:

- Выпускники в аспирантурах Англии, Франции, Швейцарии (в России в МГУ и ИПМ им. Келдыша)
- Выпускниками защищены 5 диссертаций
- Наиболее популярные в мире сравнения видеокодеков
- Более 3 миллионов скачанных фильтров обработки видео