



Школа Анализа Данных
Яндекса

Курс «Анализ изображений и видео, ч.2»

Лекции №11
«Многовидовая реконструкция»

Антон Конушин

Заведующий лабораторией компьютерной графики и мультимедиа
ВМК МГУ

28 апреля 2017 года

Благодарности



В презентации использованы слайды курса «3D Shape Reconstruction from Photographs: a Multi-View Stereo Approach» от Carlos Hernandez, George Vogiatzis and Yasutaka Furukawa

План лекции



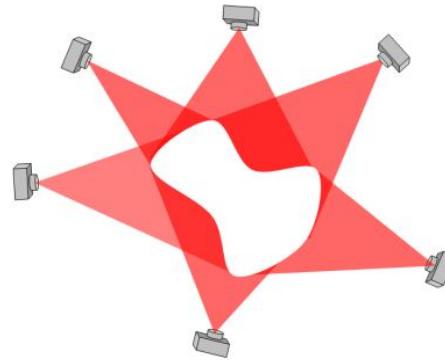
- Введение
- Многовидовое пассивное стерео
 - Объединение сеток
 - Разрастание регионов
- Приложение: реконструкция головы человека

План лекции



- Введение
- Многовидовое пассивное стерео
 - Объединение сеток
 - Разрастание регионов
- Приложение: реконструкция головы человека

Многовидовое стерео - формально



Вход:

- N фотографий (в один момент времени) объекта с известной калибровкой

Выход:

- 3D модель объекта – геометрическая модель и текстура

Многовидовое стерео



- Сравнения и оценки:
 - **A Comparison and Evaluation of Multi-View Stereo Reconstruction Algorithms**,
S. Seitz et al., CVPR 2006, vol. 1, pages 519-526.
 - <http://vision.middlebury.edu/mview/>
- Недавно много новых алгоритмов появилось
- Высокая точность
- Почти все работают с небольшим количеством изображений (~100) основное исключение [Pollefeys08]
- Работают очень долго (off-line)

Sort By	Temple		Temple		Temple		Dino		Dino		Dino	
	Full		Ring		Sparse		Full		Ring		Sparse	
	312 views	47 views	16 views	363 views	48 views	16 views	Acc	Comp	Acc	Comp	Acc	Comp
Furukawa 3	0.49	99.6	0.47	99.6	0.63	99.3	0.33	99.8	0.28	99.8	0.37	99.2
Campbell	0.41	99.9	0.48	99.4	0.53	98.6						
Hernandez	0.36	99.7	0.52	99.5	0.75	95.3	0.49	99.6	0.45	97.9	0.6	98.5
Furukawa 2	0.54	99.3	0.55	99.1	0.62	99.2	0.32	99.9	0.33	99.6	0.42	99.2
Zaharescu			0.55	99.2	0.78	95.8			0.42	98.6	0.45	99.2
Zach2	0.51	98.8	0.56	99.0			0.55	98.7	0.51	99.1		
SurfEvolution			0.56	98.9	0.78	96.8			0.56	97.7	0.66	97.6
Bradley			0.57	98.1	0.48	93.7			0.39	97.6	0.38	94.7
Zach			0.58	99.0					0.67	98.0		
Furukawa	0.65	98.7	0.58	98.5	0.82	94.3	0.52	99.2	0.42	98.8	0.58	96.9
Pons			0.6	99.5	0.9	95.4			0.55	99.0	0.71	97.7
Goesele	0.42	98.0	0.61	86.2	0.87	56.6	0.56	80.0	0.46	57.8	0.56	26.0
Zaharescu2			0.62	98.5					0.5	98.5		
Vogiatzis2	0.5	98.4	0.64	99.2	0.69	96.9						
Sormann			0.69	97.2					0.81	95.2		
ICCV_1500	0.65	85.8	0.7	78.9	0.59	74.9	0.91	73.8	0.71	76.6	0.66	74.9
Kolev2			0.72	97.8	1.04	91.8			0.43	99.4	0.53	98.3
Surfel Cut			0.73	97.5					0.69	98.7		
Merrell Stability			0.76	85.2					0.73	73.1		
Vogiatzis	1.07	90.7	0.76	96.2	2.77	79.4	0.42	99.0	0.49	96.7	1.18	90.8
Kolev			0.79	96.0					0.53	96.9		
Sinha			0.79	94.9					0.69	97.2		
Merrell Confidence			0.83	88.0					0.84	83.1		
Auclair			0.86	96.2	1.03	92.5			0.62	96.7	0.74	96.8
Strecha			0.86	97.6	1.05	94.1			1.21	92.4	1.41	91.5
Gargallo			0.88	84.3	1.05	81.9			0.6	92.9	0.76	90.7
Tran			1.12	92.3	1.53	85.4			1.12	92.0	1.26	89.3
Kolmogorov			1.86	90.4					2.81	86.0		
Continuous Probab			1.89	92.1					2.61	91.4		
3DIM_155	0.55	99.7					0.76	99.0				
Delaunoy					0.73	95.9				0.89	93.9	
Goesele 2007	0.42	98.2					0.46	96.7				
Habbecke	0.66	98.0					0.43	99.7				
Hornung	0.58	98.7					0.79	95.1				
JancosekCVWW									0.79	95.9		
Ladikos										0.89	95.0	
Liu					0.96	89.6				0.59	98.3	
Liu2						0.65	96.9				0.51	98.7
Massively Parallel	0.79	92.2								1.01	90.7	
Starck							1.27	87.7				

План лекции



- Введение
- Многовидовое пассивное стерео
 - Объединение сеток
 - Разрастание регионов
- Приложение: реконструкция головы человека

Фото-согласование



Основной принцип
многовидового
стерео

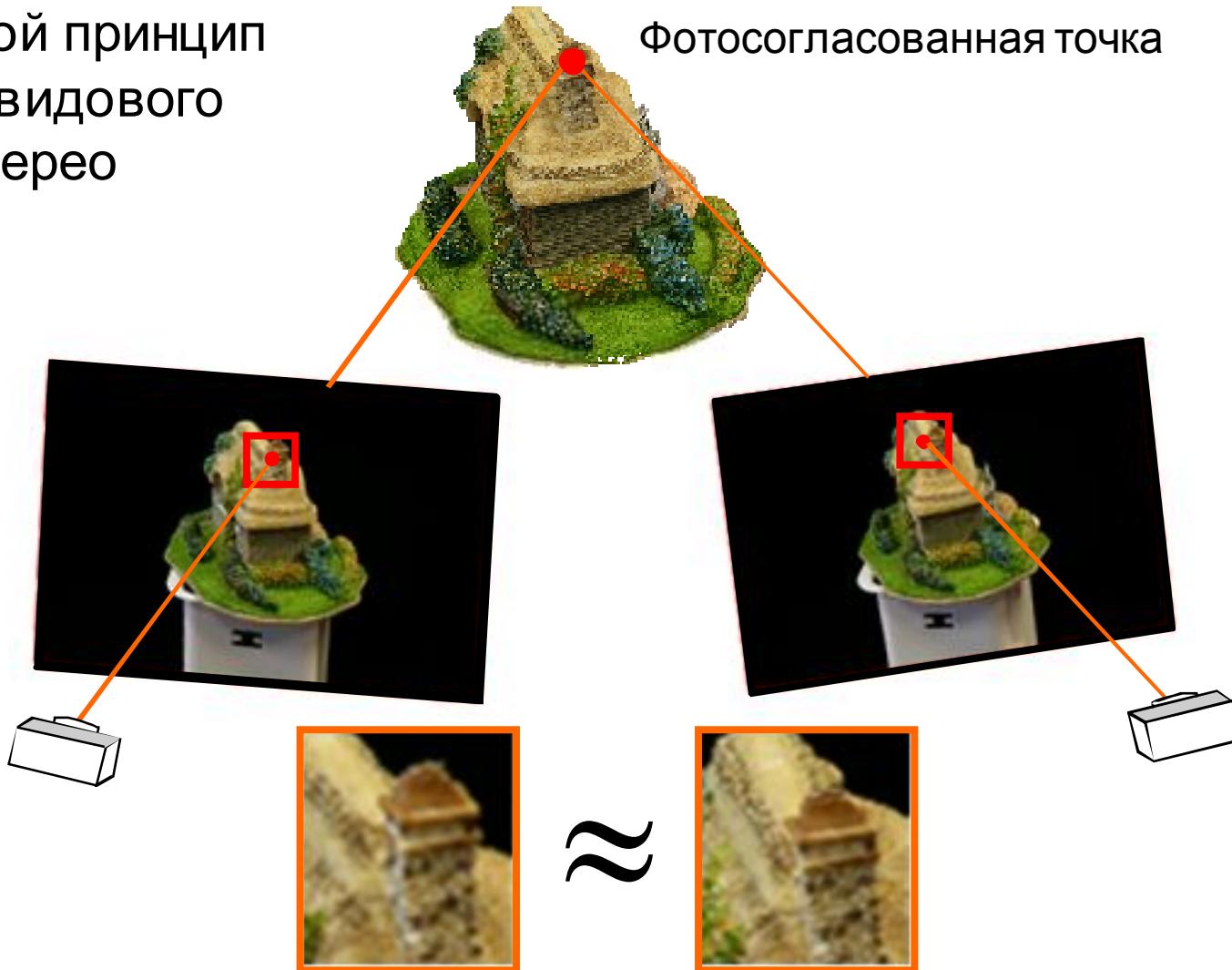
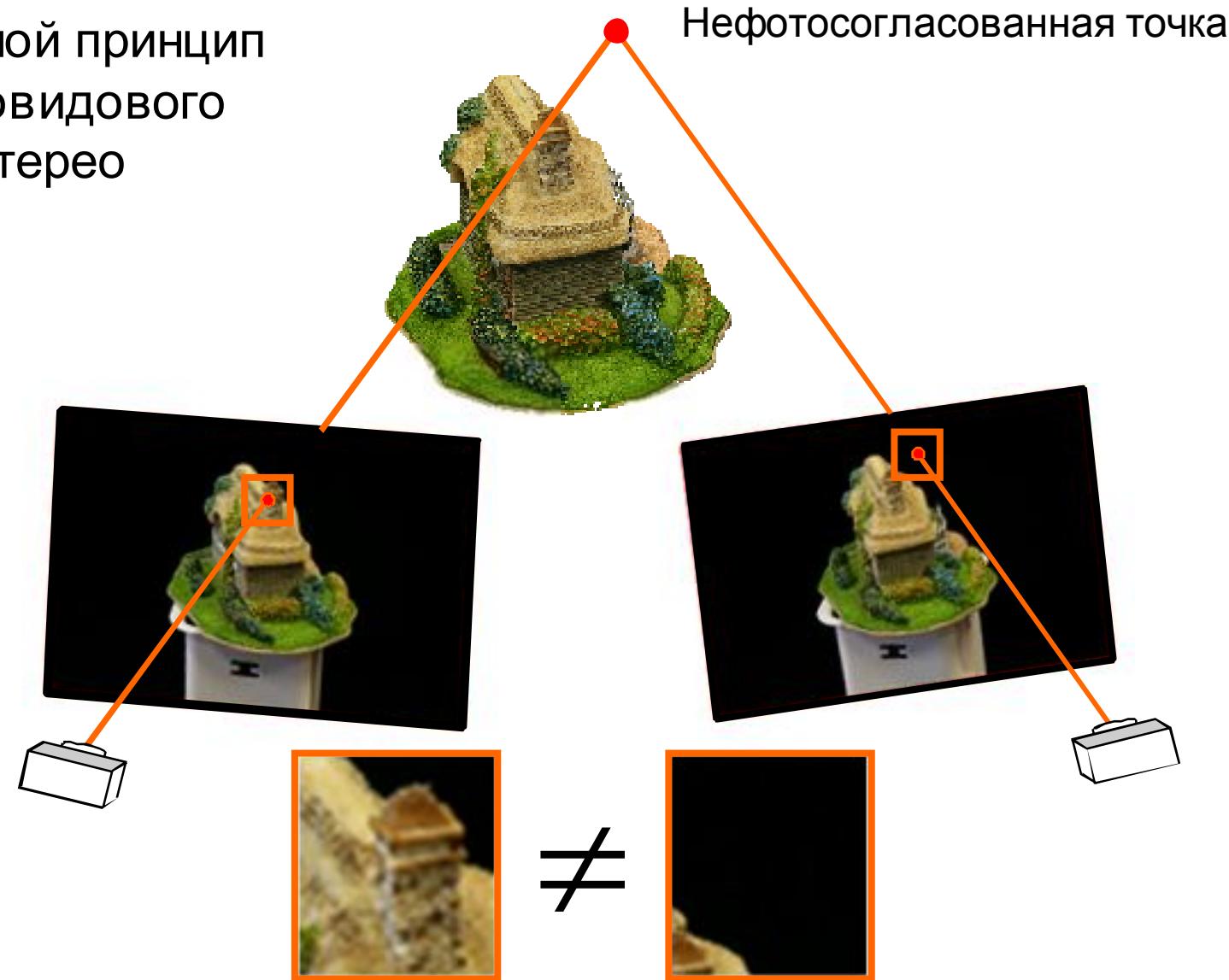


Фото-согласование



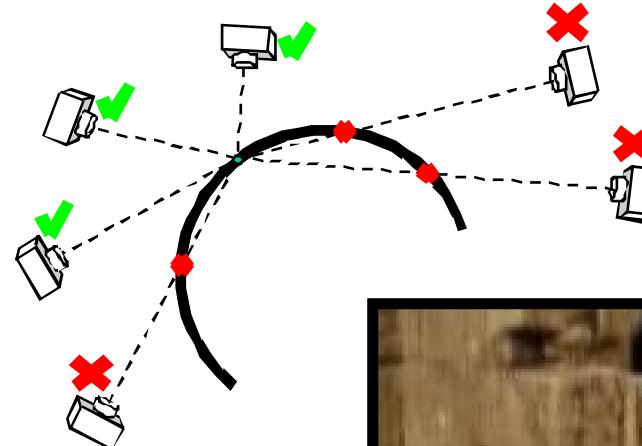
Основной принцип
многовидового
стерео



Проблемы фото-согласования



- Видимость камеры
- Ошибки метрики
 - Повтор текстуры
 - Отсутствие текстуры
 - Блики



Разные методы по
разному подходят к
решению указанных
проблем

План лекции

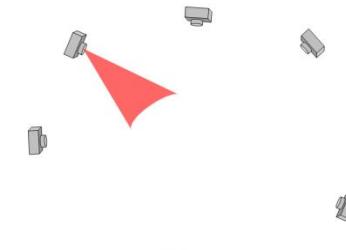


- Введение
- Многовидовое пассивное стерео
 - *Объединение сеток*
 - Разрастание регионов
- Приложение: реконструкция головы человека

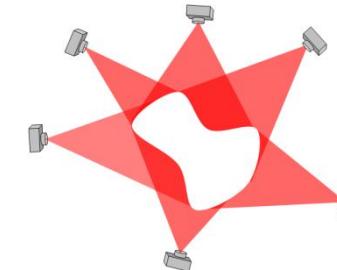
Схема объединения карт глубины



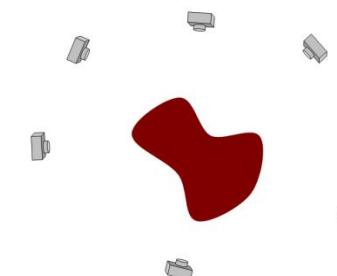
1. Вычисляем карту глубины



2. Объединяем карты глубины в объеме



3. Извлекаем 3д поверхность

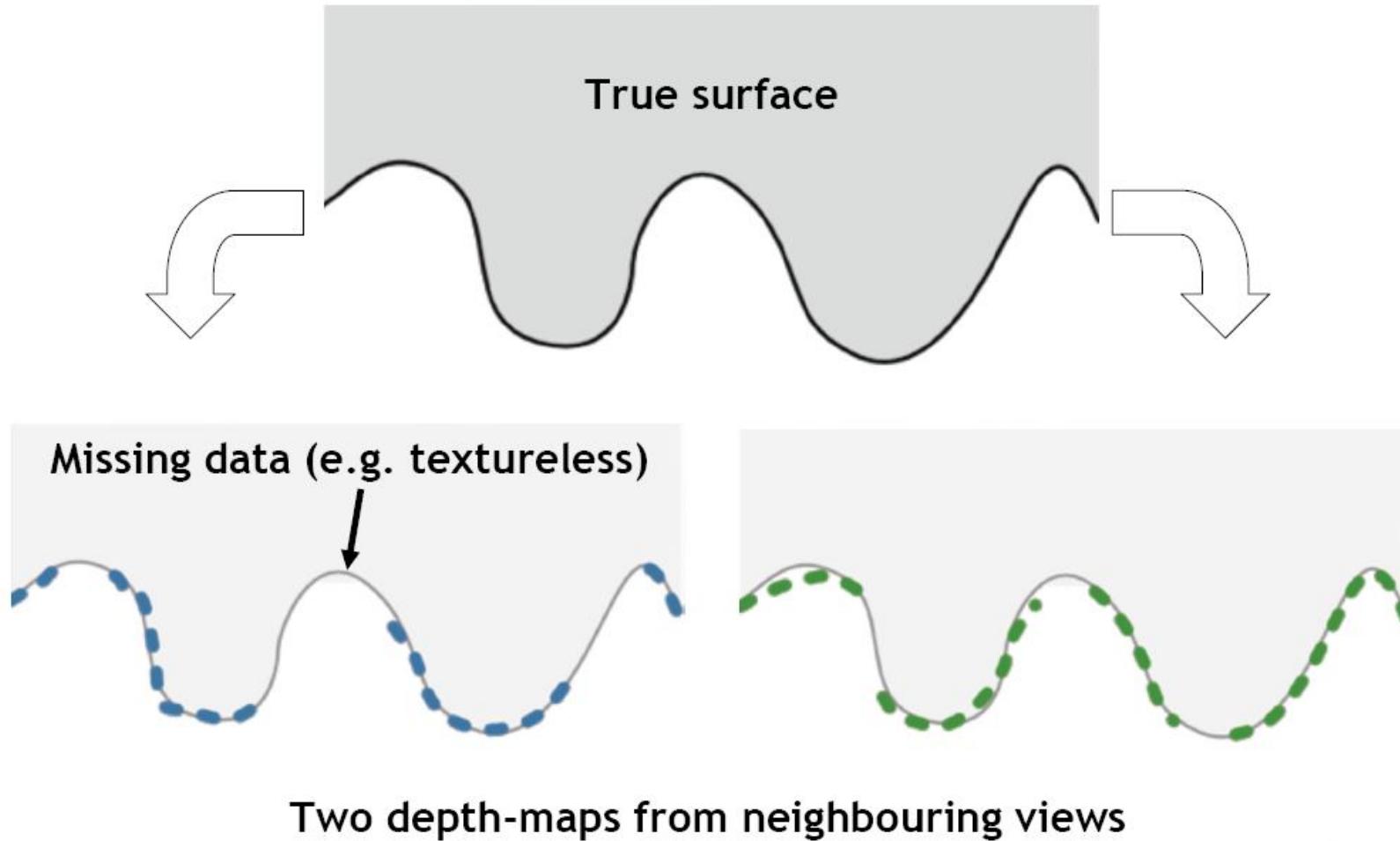


Проблема регуляризации

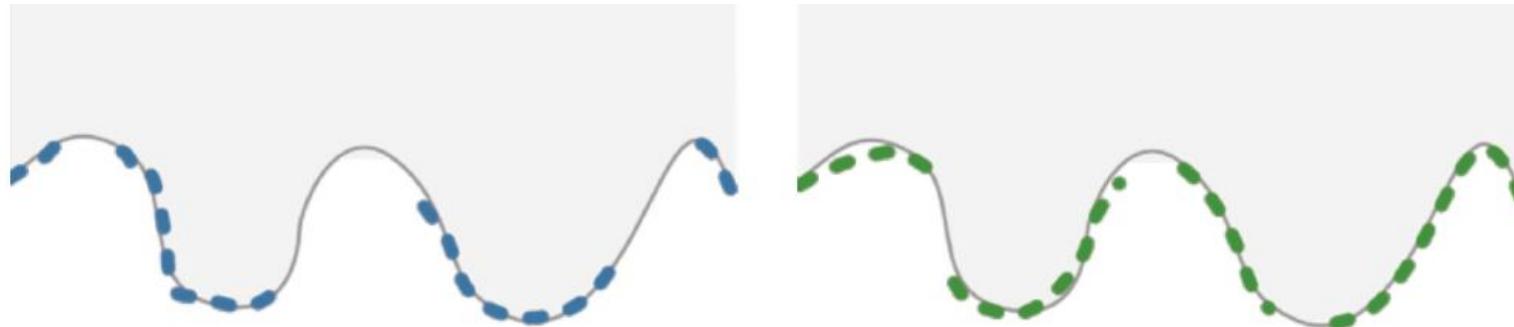


- Какие методы бинокулярного стерео будем использовать?
- Лучшие сейчас – на основе функционалов энергии и на основе пересегментации
- В чём их основная идея?
 - Накладываем на диспаритет дополнительные ограничения
 - Интерполируем значения в сложных областях
- Посмотрим, к чему это приводит при многовидовом стерео

Ранняя регуляризация



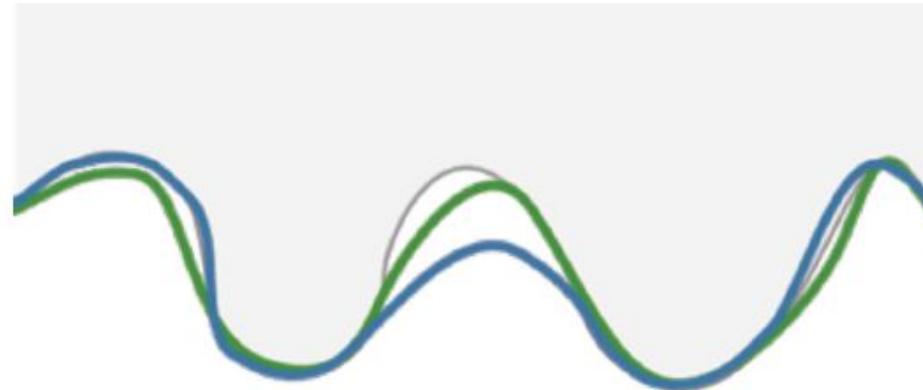
Ранняя регуляризация



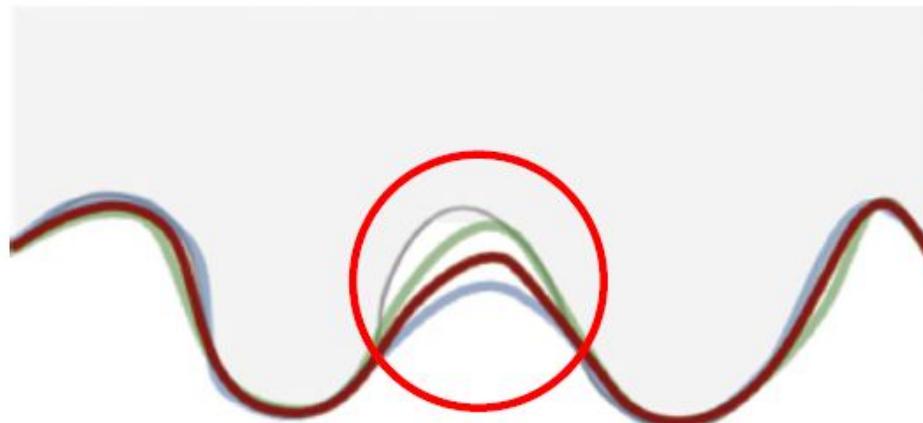
Регуляризация каждой карты глубины по
отдельности



Объединение оценок



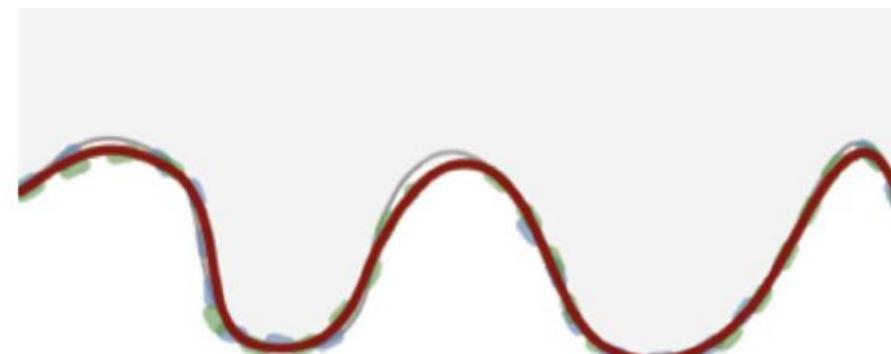
Объединяем две оценки карты глубины



Ранняя регуляризация



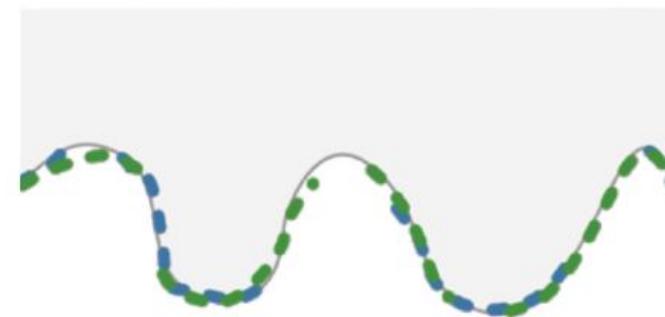
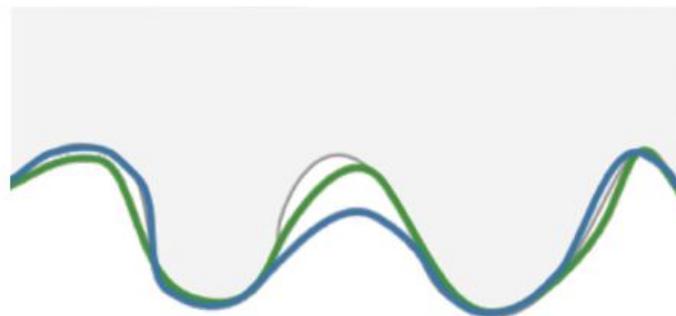
Объединяем две карты глубины, а затем
регуляризуем



Ранняя регуляризация



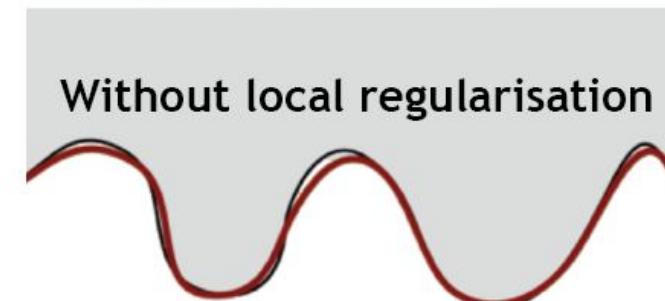
Before combining



Final result



With local regularisation



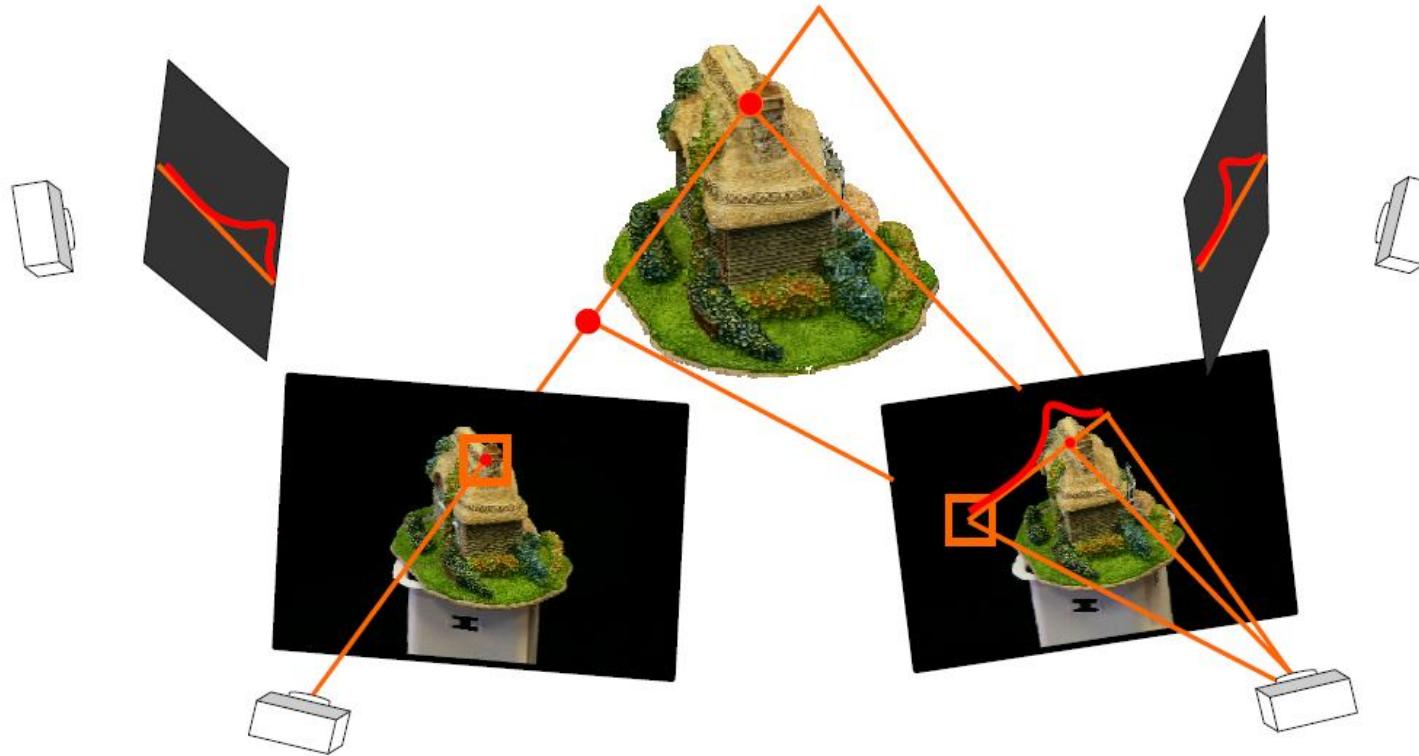
Without local regularisation

Общий подход



- Проблема регуляризации:
 - В сложных областях метод интерполирует значения за счёт использования регуляризации
 - Регуляризовать приходится из-за предположения, что дополнительных данных нет
 - В случае многовидового стерео дополнительная информация есть!
- Поэтому будем делать так:
 - Вычисляем диспаритет локальными методами, без регуляризации
 - Объединяем наблюдения между разными парами кадров
 - Затем накладываем дополнительные условия

Вычисление стерео

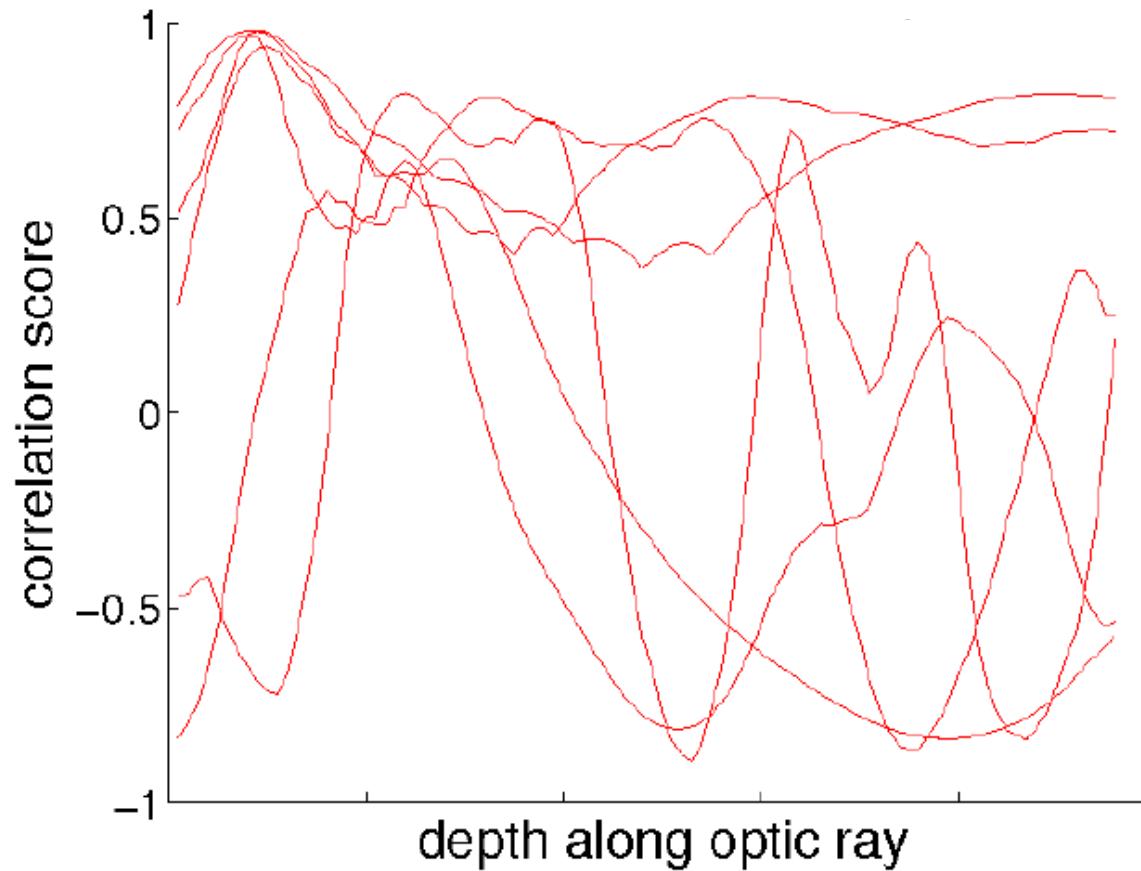


- Выберем опорное (базовое) изображение
- Для каждого пикселя на опорном изображении построим график меры фотосогласования в зависимости от глубины точки для всех дополнительных изображений

Объединение наблюдений

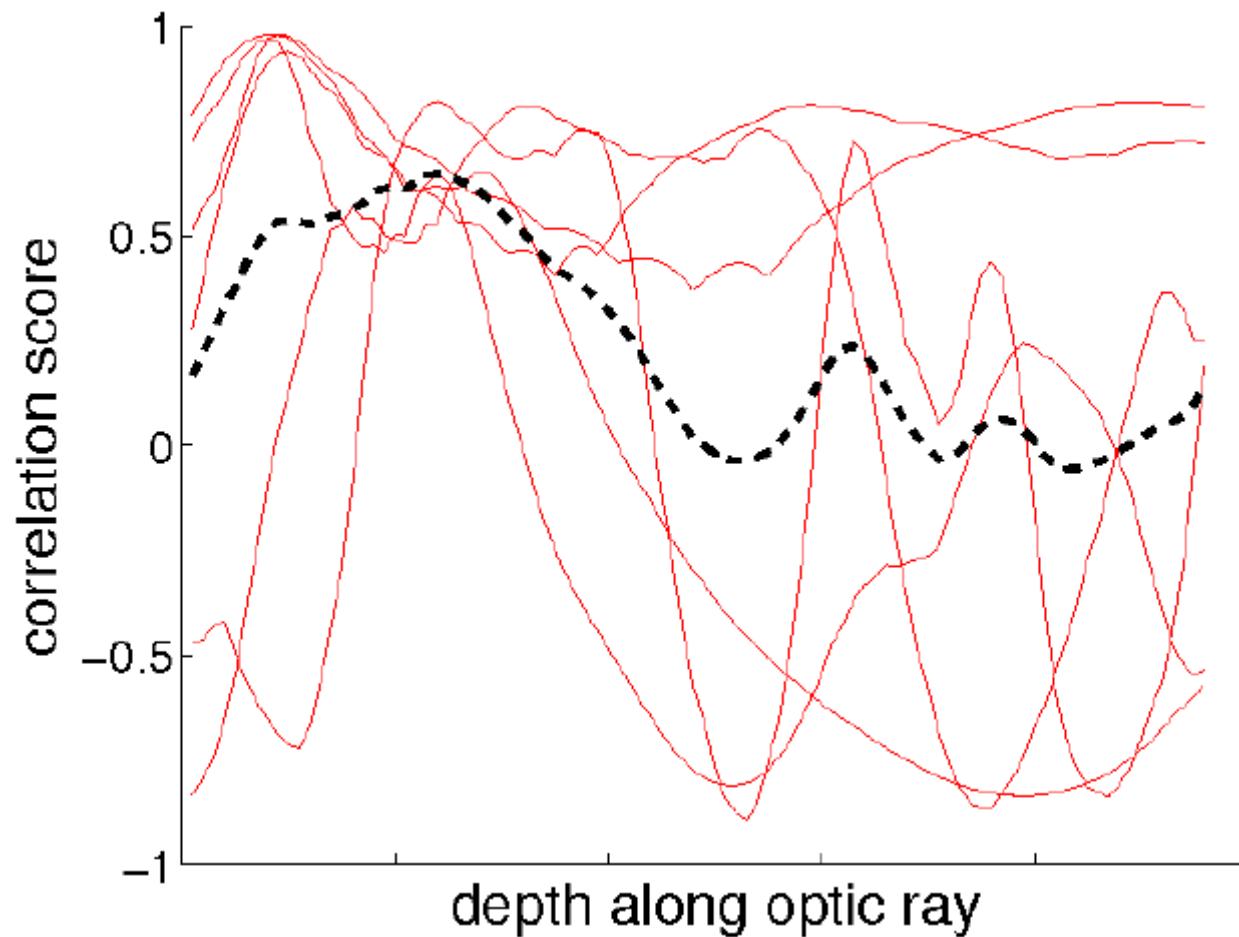


Где истинная глубина точки?

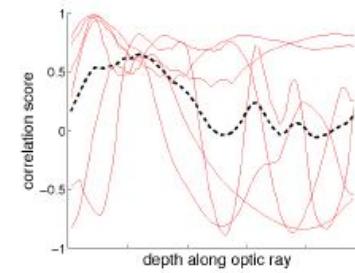
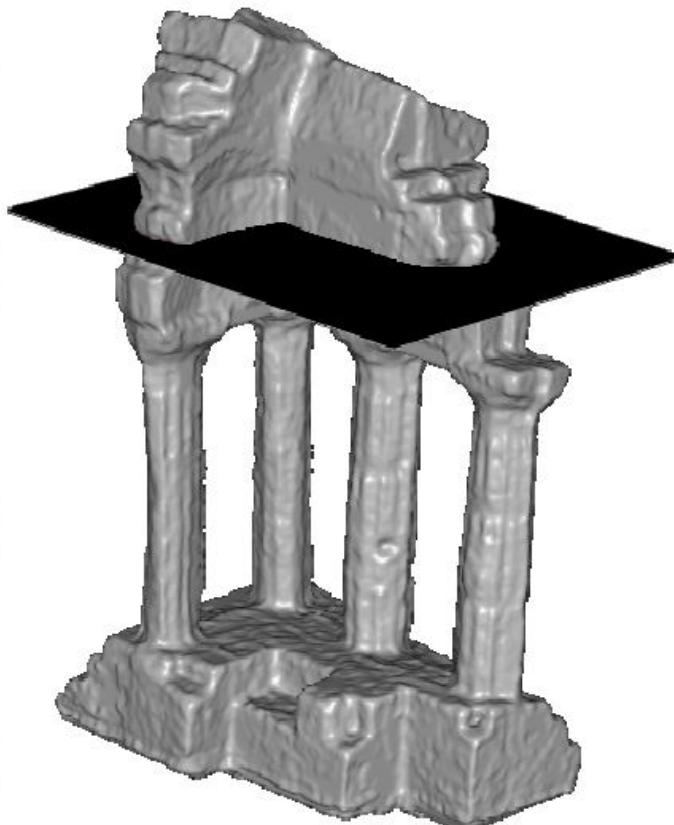


- Объединим кривые корреляции

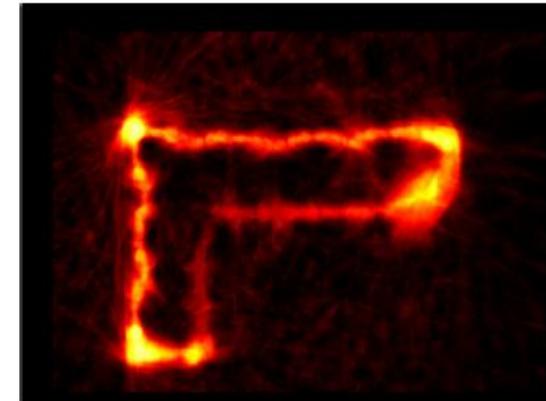
Простое усреднение



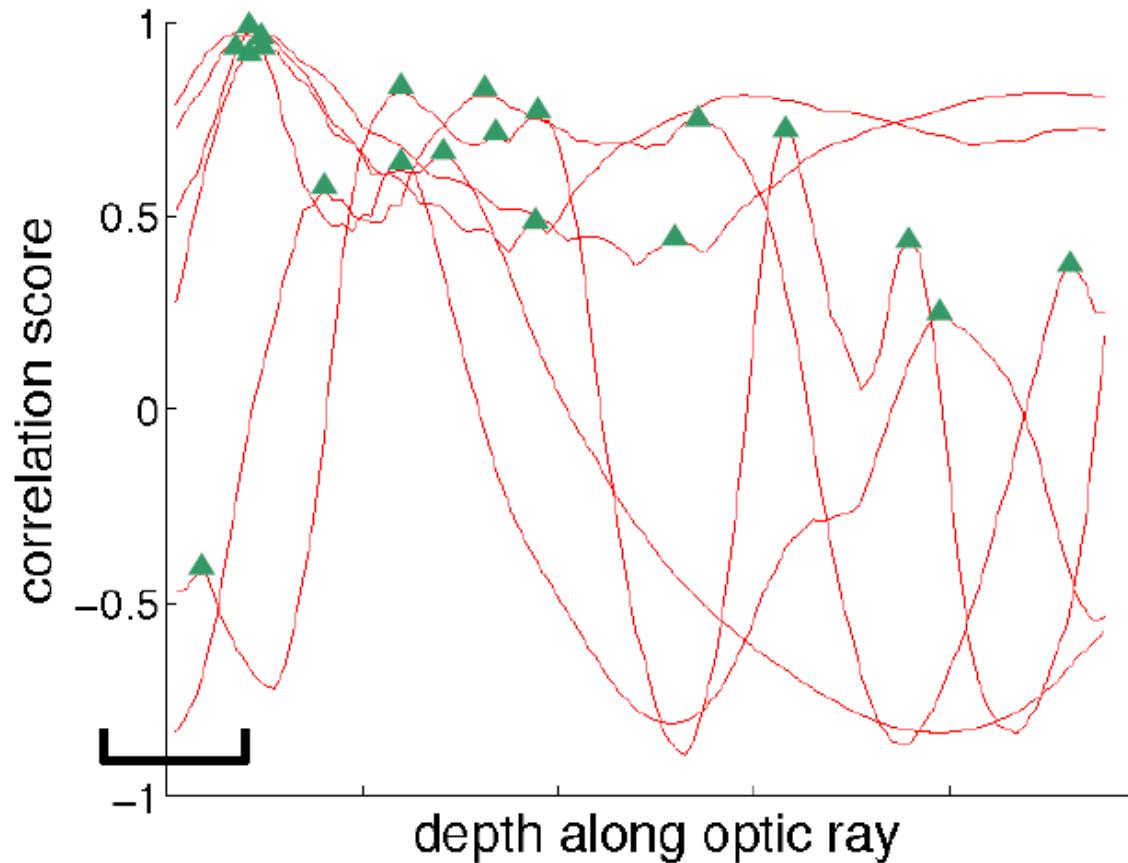
Простое усреднение



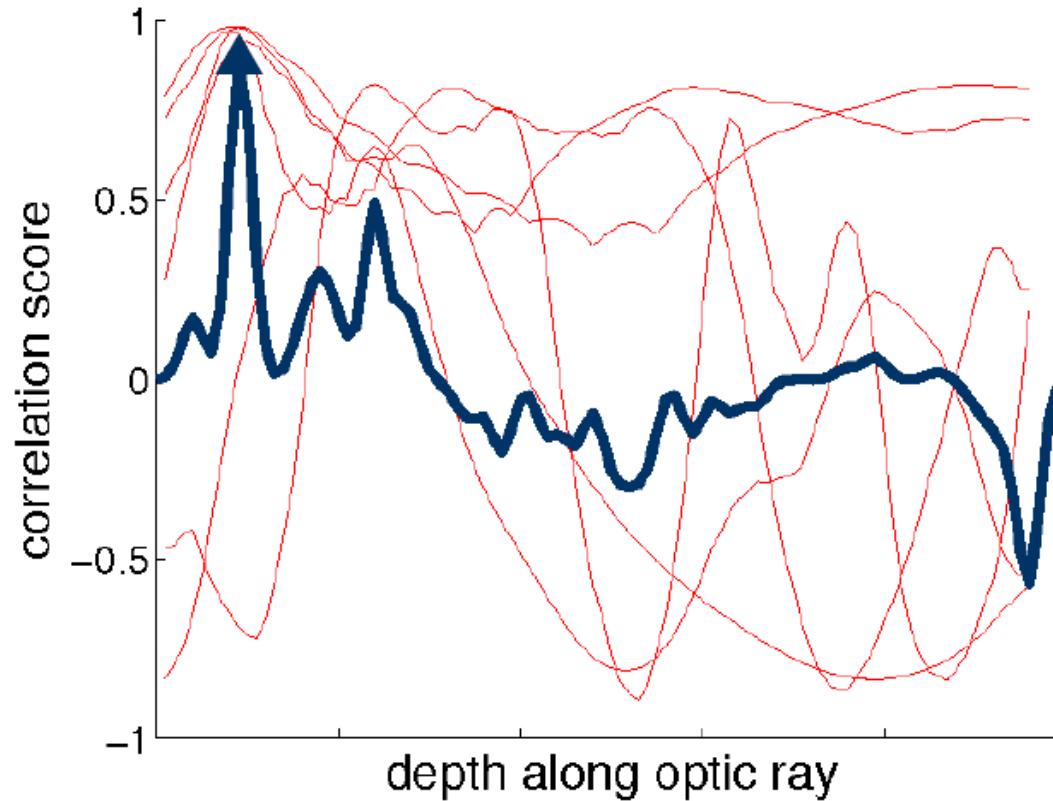
In 3d



При простом усреднении оценка глубины получается очень неточной

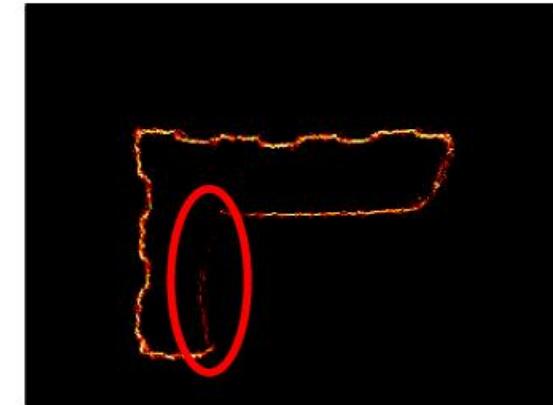
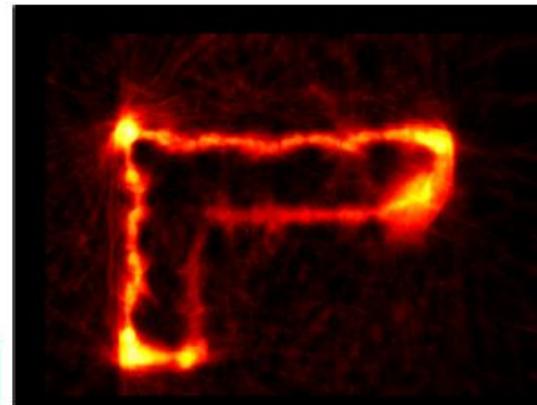
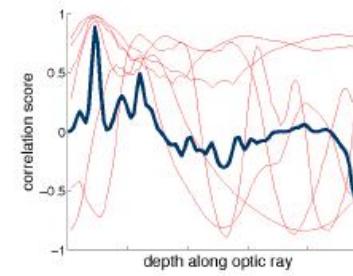
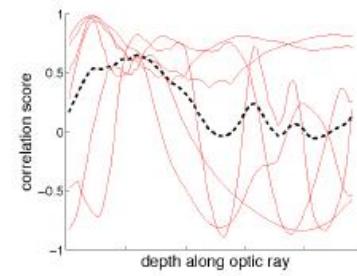
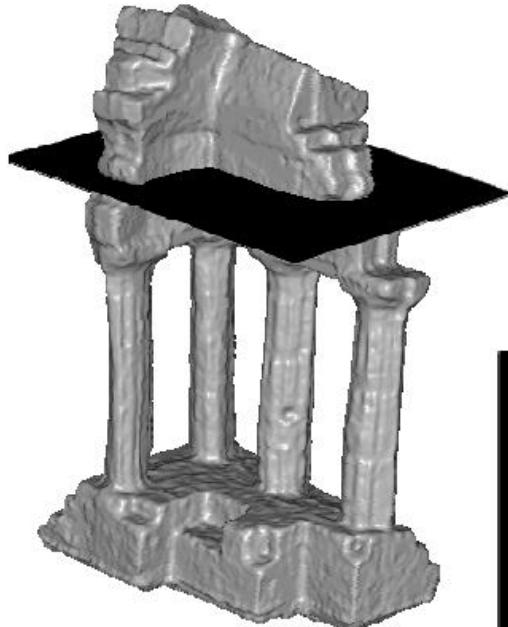


- Ключевое предположение
 - правильная глубина соответствует одному из локальных пиков одной из локальных кривых



- Окна Парзена – непараметрическая оценка количества локальных максимумов по окну

Максимальный отклик



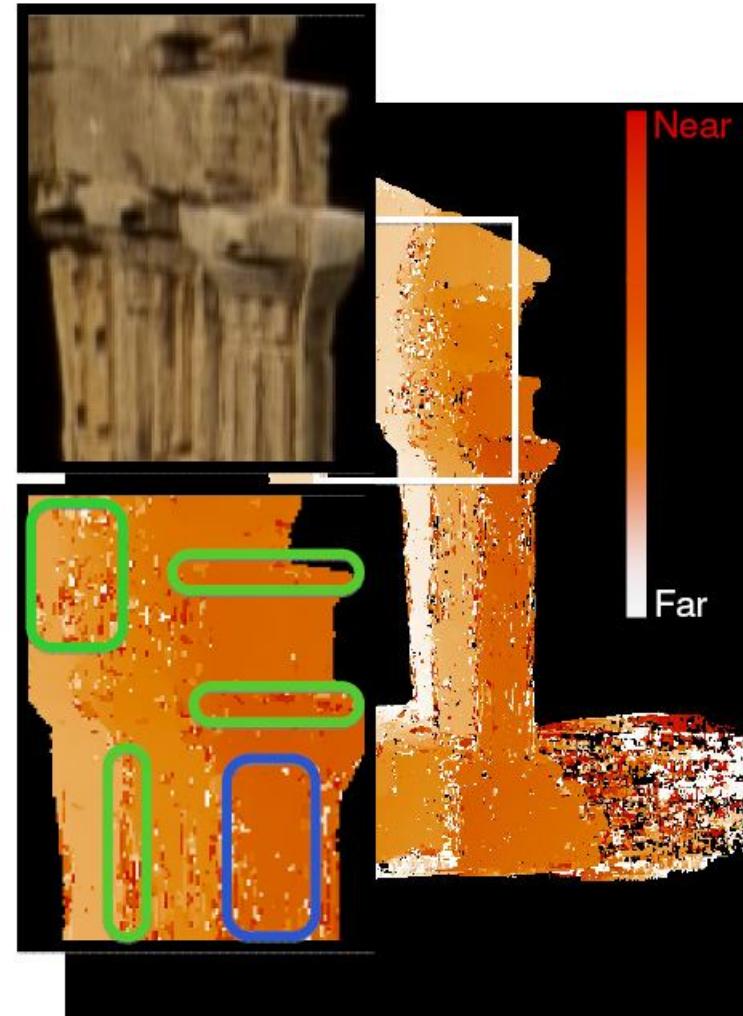
MATLAB

В тех областях, где совпадают много локальных максимумов получается хорошо. В других – не очень.

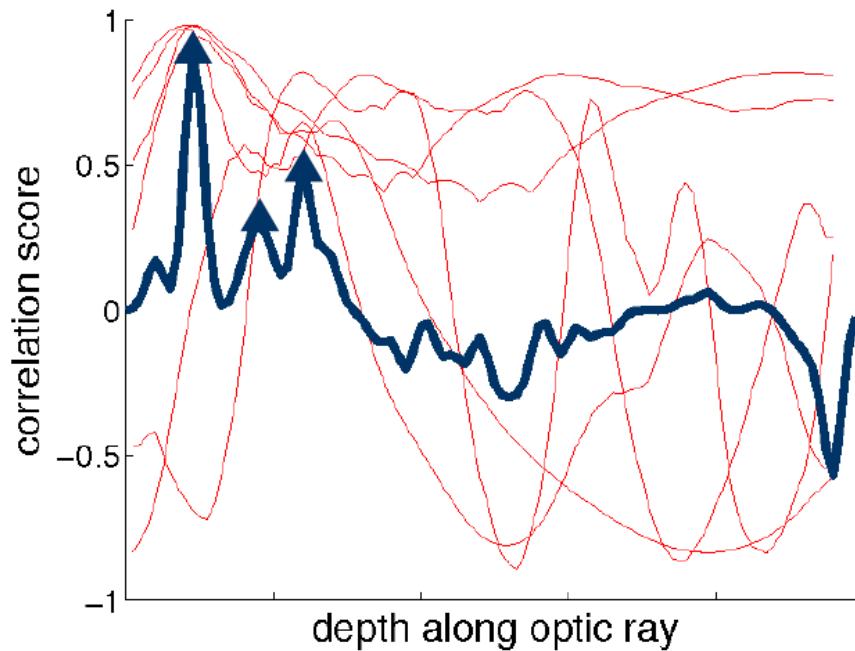


Ошибки

- Повторяющаяся текстура
- Ошибки сопоставления
 - Отсутствие текстуры
 - Шум в изображении и т.д.
- Типичные ошибки стерео!



Регуляризация



- Любой локальный пик может быть соответствием
- Используем избыточность данных и свойство непрерывности поверхности
 - У соседних пикселей должно быть похожая глубина!

Схема MRF



- Выберем k самых сильных пиков как гипотезы глубины для пикселя. Дополним U – «глубина неизвестна»

$$k_i \in \{1 \dots K, \mathcal{U}\}$$

- Зададим функционал энергии:

$$\sum_i \phi(k_i) + \sum_{(i,j)} \psi(k_i, k_j)$$

- Унарный потенциал пенализирует малые пики
- Парный потенциал пенализирует резкие изменения глубины между соседними пикселями
- Метка U нужна, чтобы добавить устойчивость к шуму
 - Небольшой штраф за U
 - Небольшой константный штраф за соседство с U

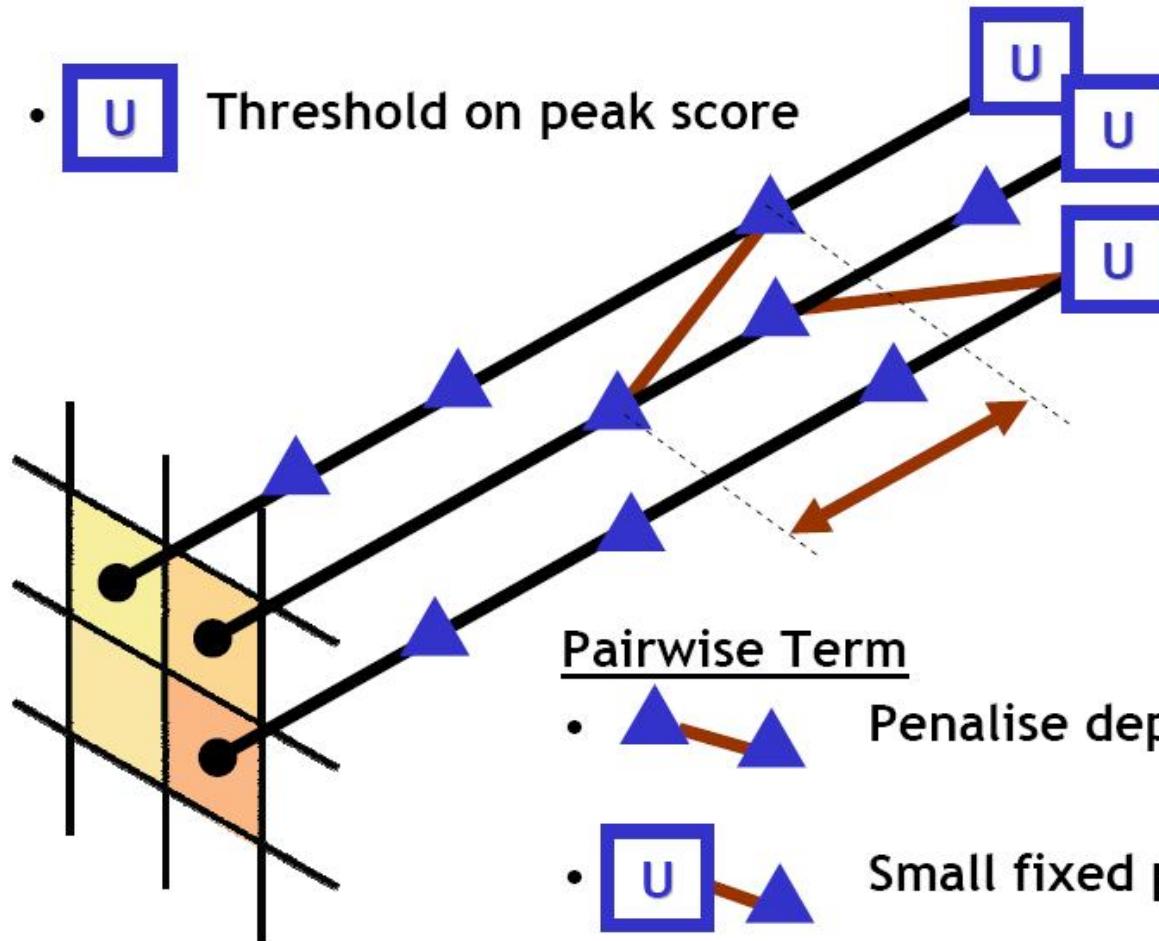
Схема



Unary Term

- Penalise peaks with low score

- Threshold on peak score



Pairwise Term

- Penalise depth disparity

- Small fixed penalty

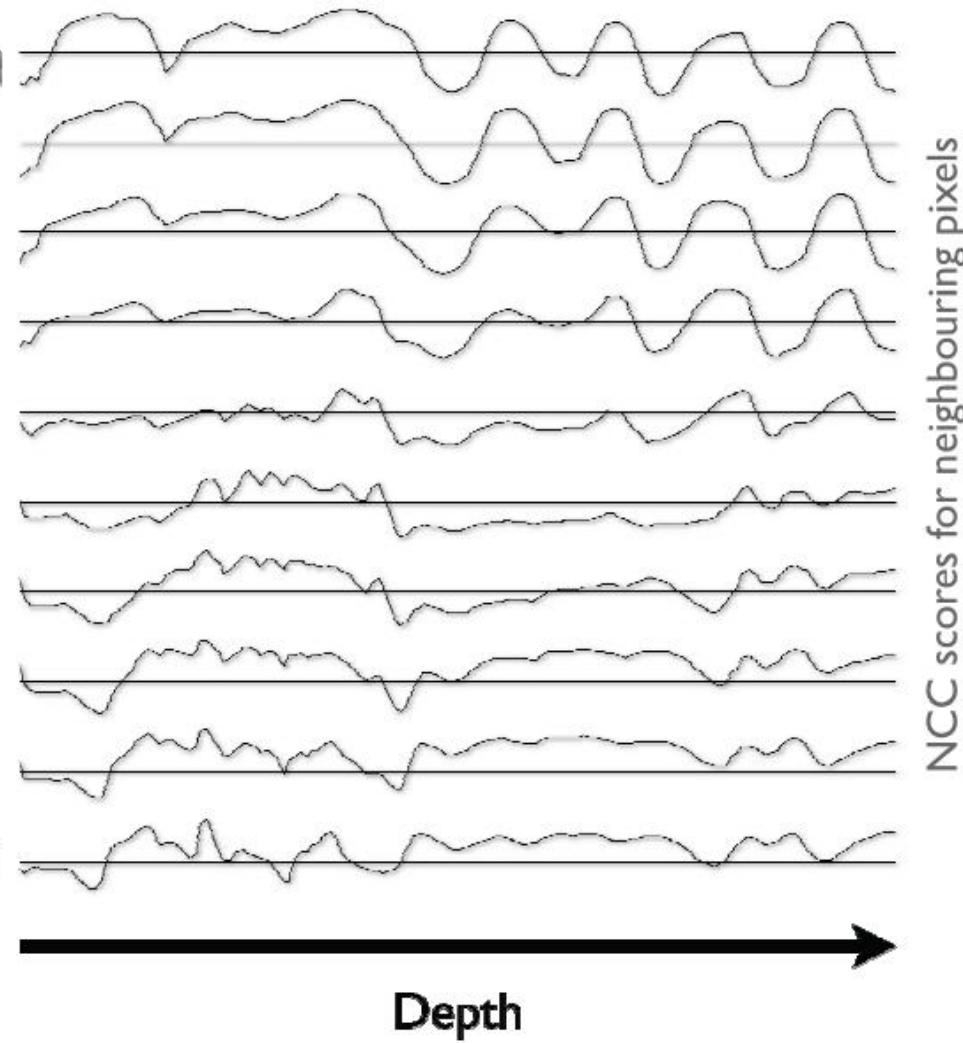
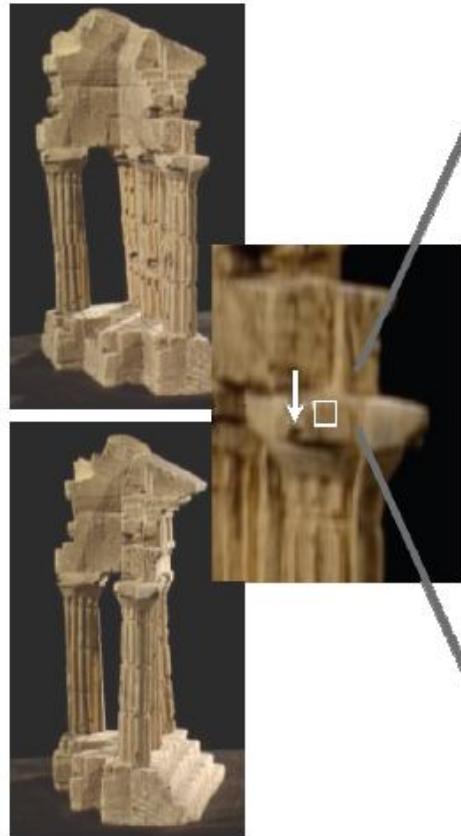
Astor

Ранняя регуляризация?



- Да, но с оговорками...
 - Мы фильтруем ненадежные данные
 - Мы не интерполируем промежуточные значения
 - Если для глубины пикселя нет хороших гипотез, мы записываем как «неизвестное значение»

Пример

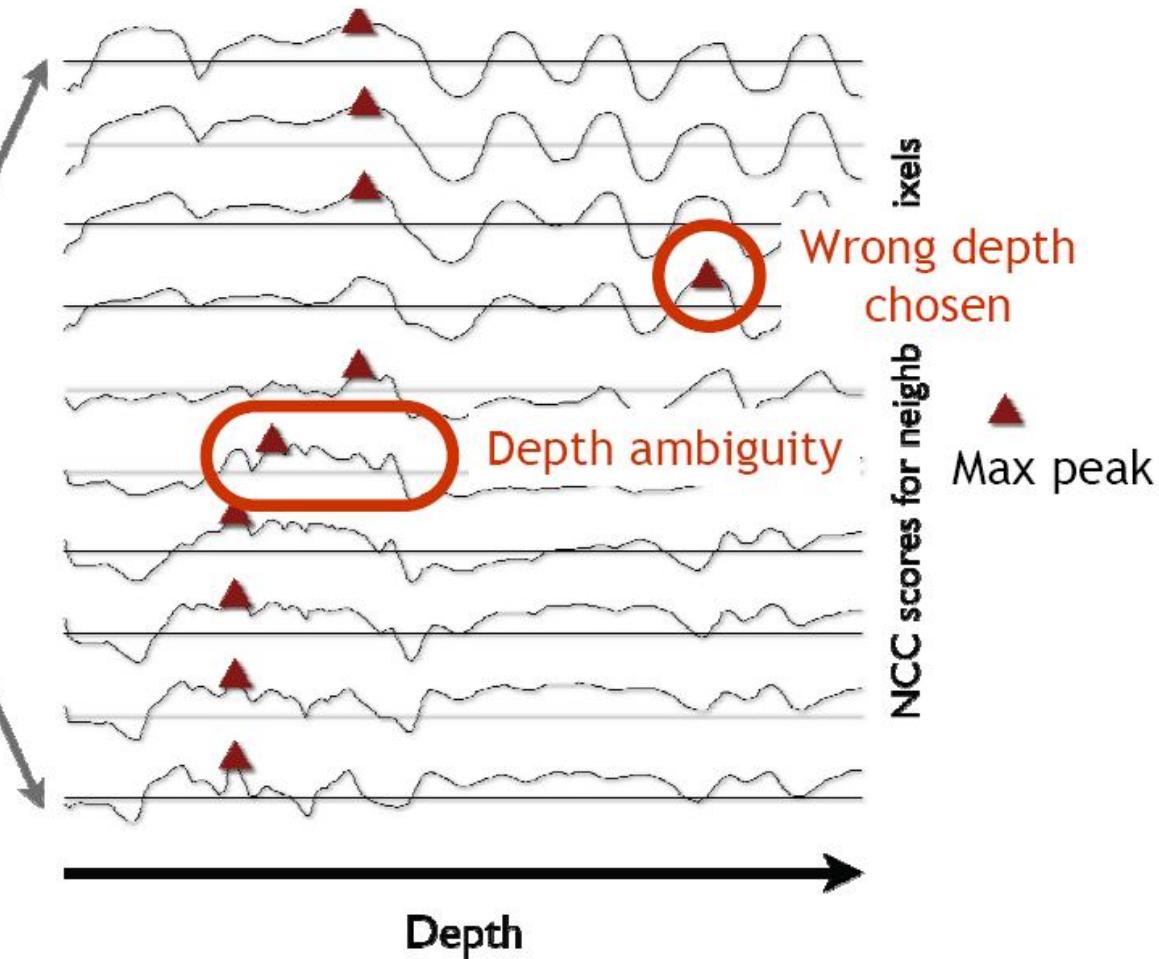
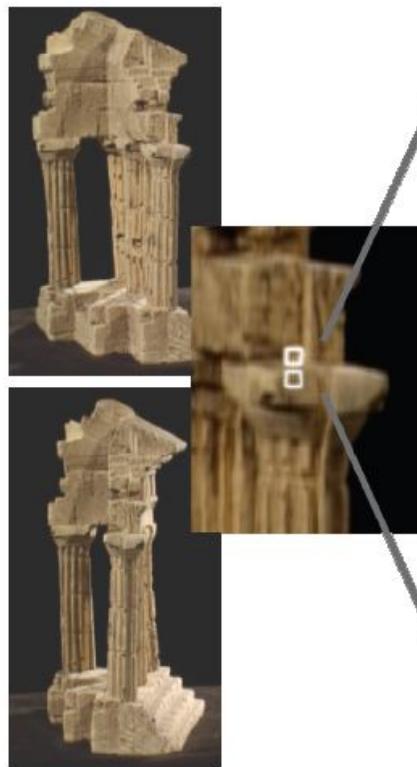


NCC scores for neighbouring pixels

Depth

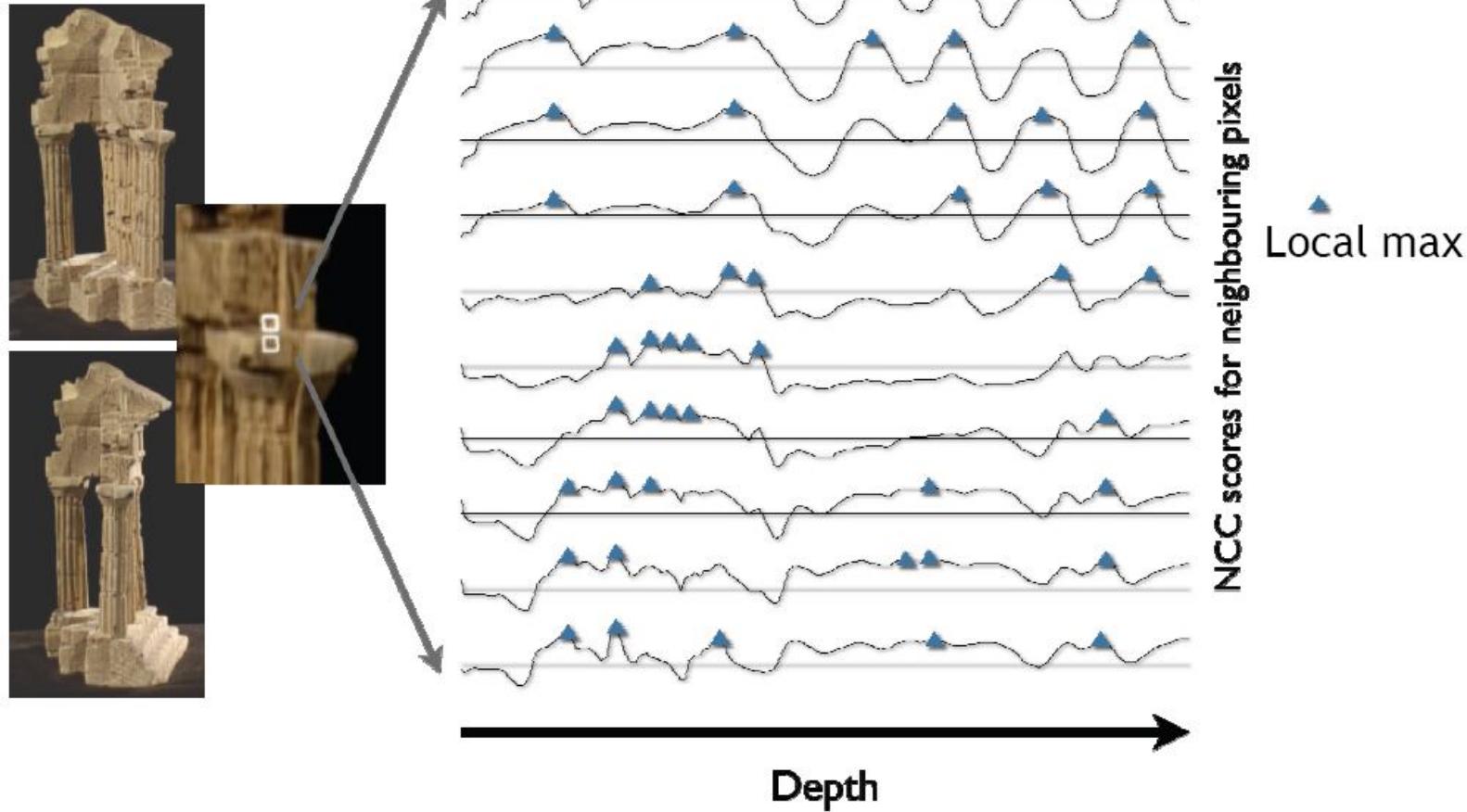
Рассмотрим несколько соседних пикселей

Пример



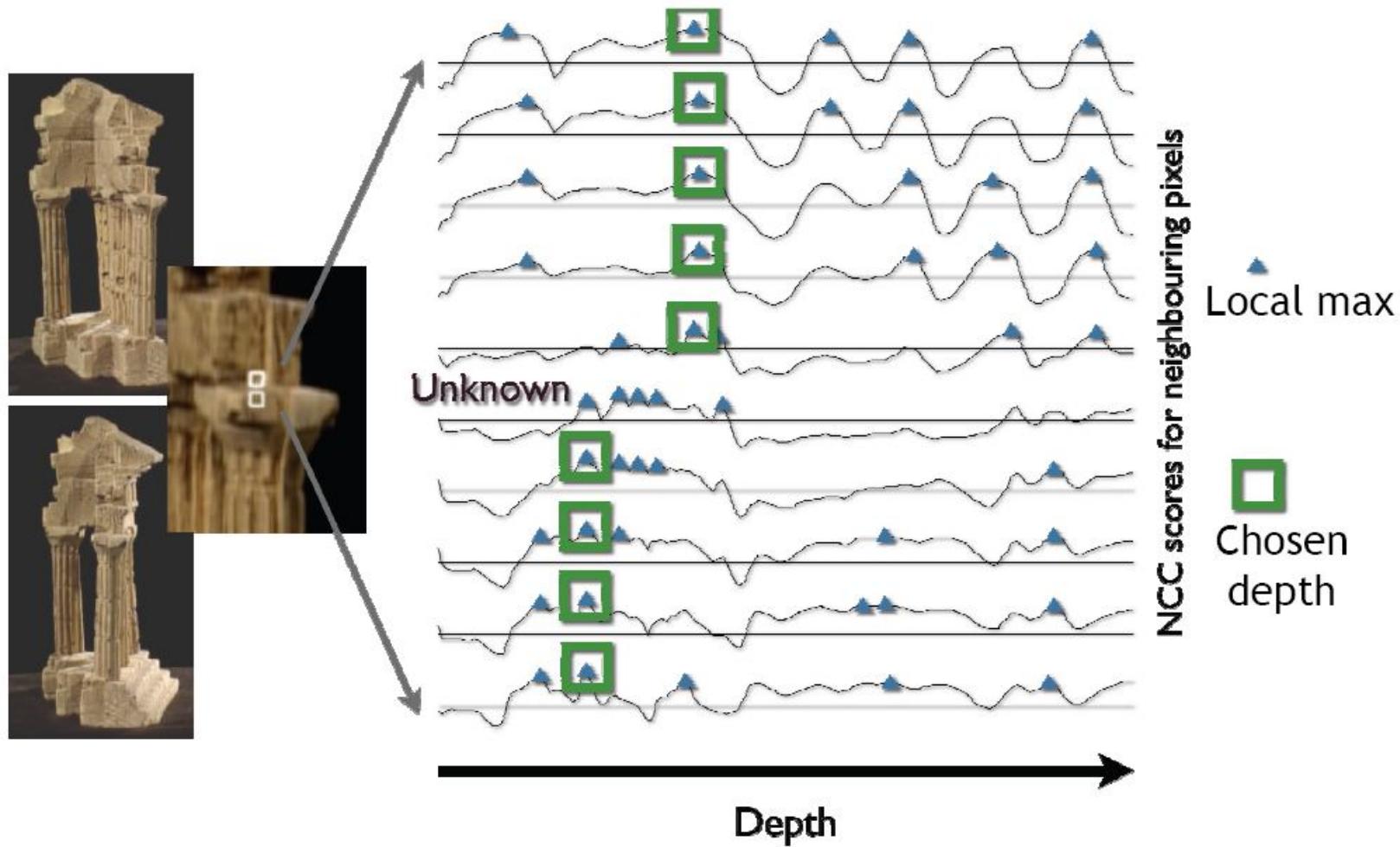
В некоторых пикселях ошибки или неоднозначное решение

Пример



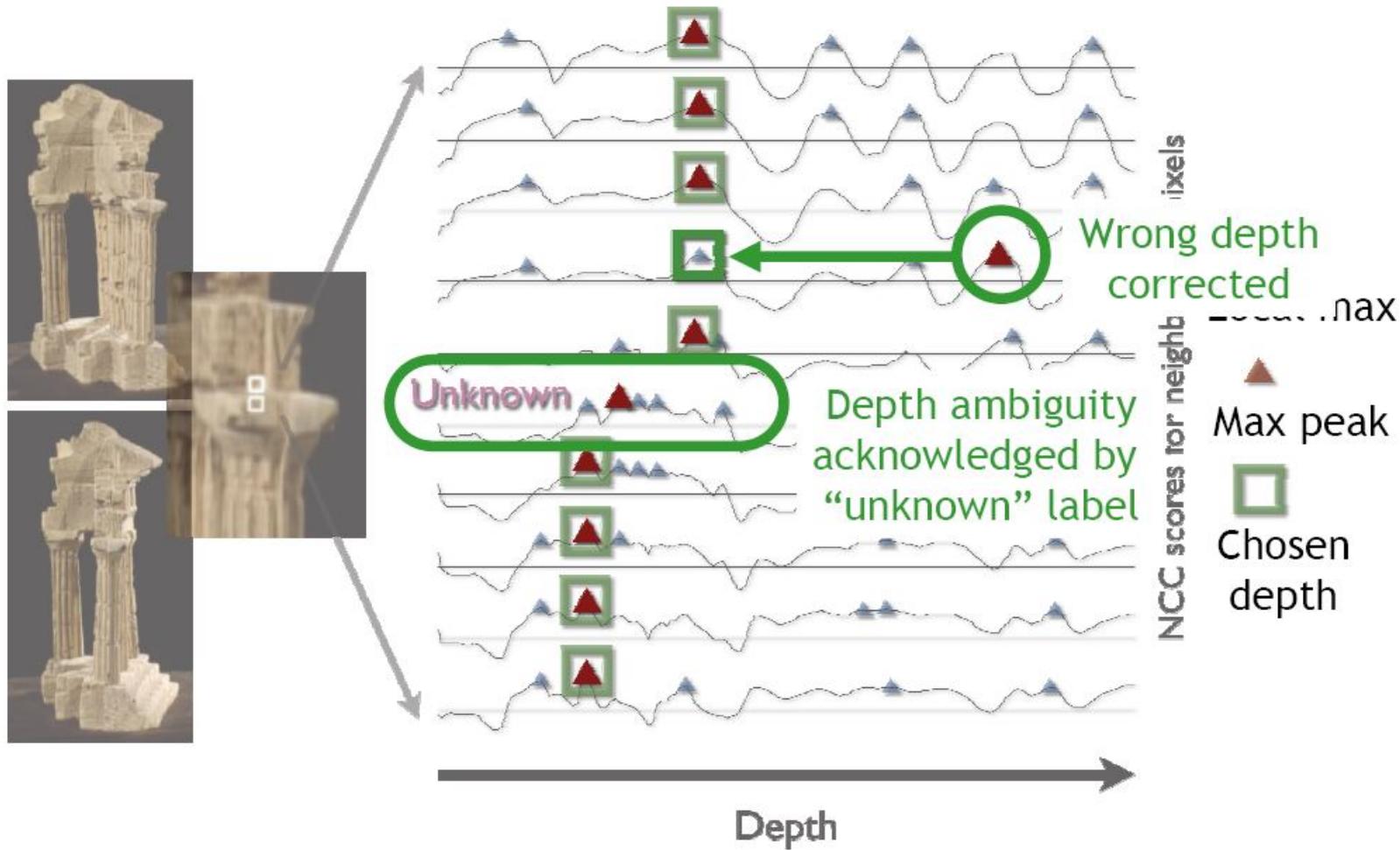
Найдём локальные максимумы для каждого пикселя

Пример



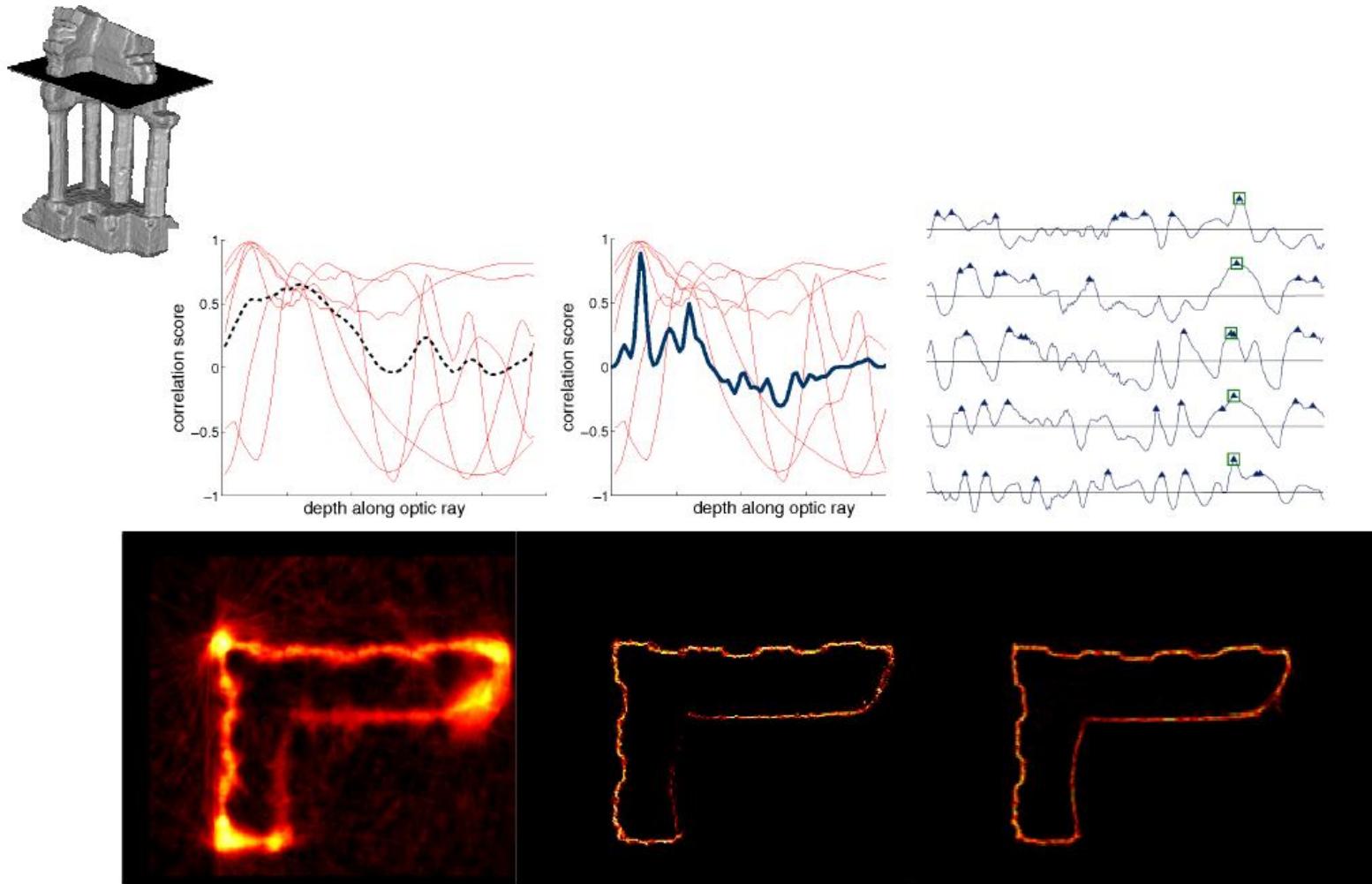
Оптимизируем качество с помощью MRF-схемы

Пример



Исправили ошибки и выделили область неопределенности

Результат



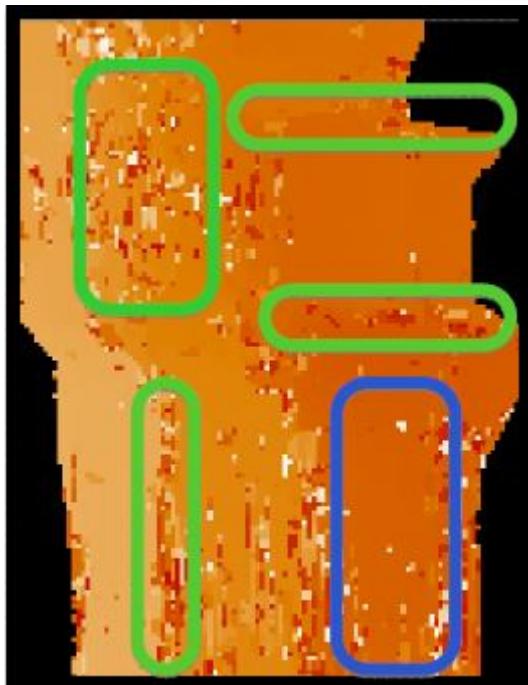
Точность и полнота реконструкции повысилась!

Пример



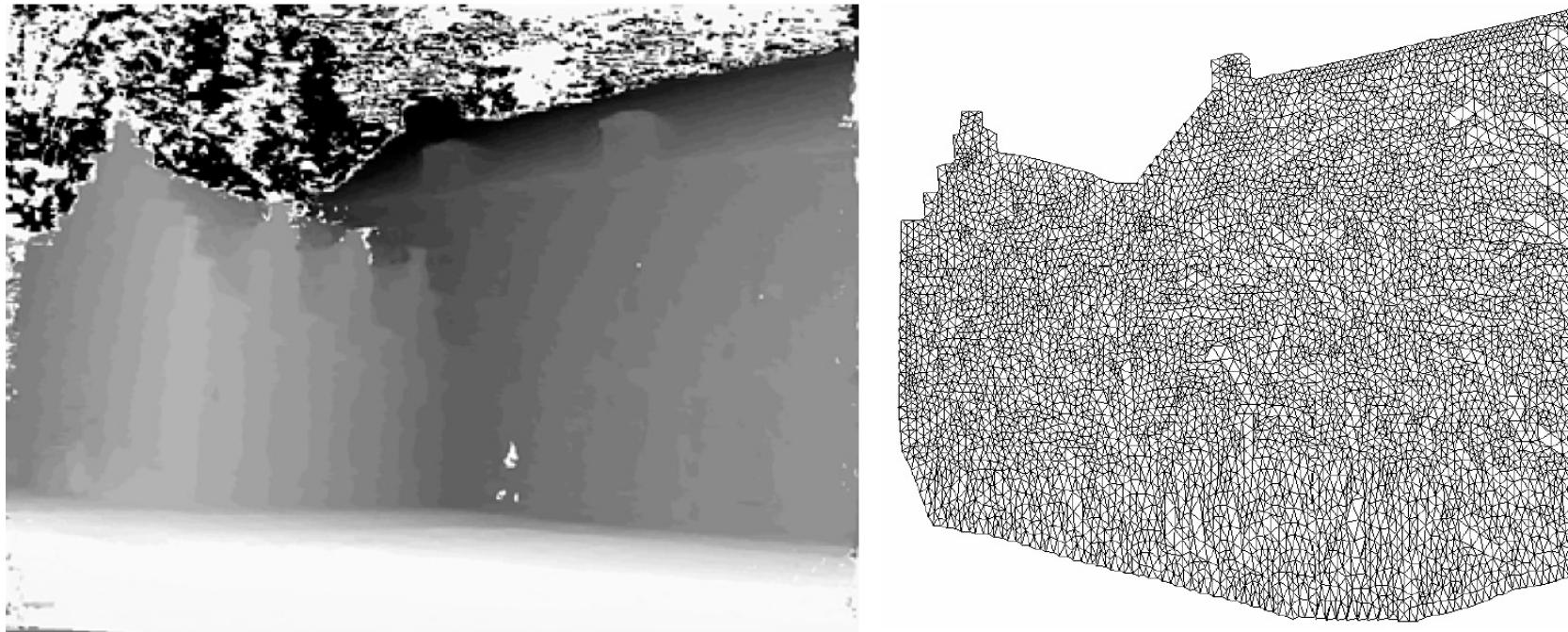
Посмотрим карту глубины опорного изображения

Пример



Выделились разрывы (области неопределенности) и
отфильтровались ошибки на ровных поверхностях

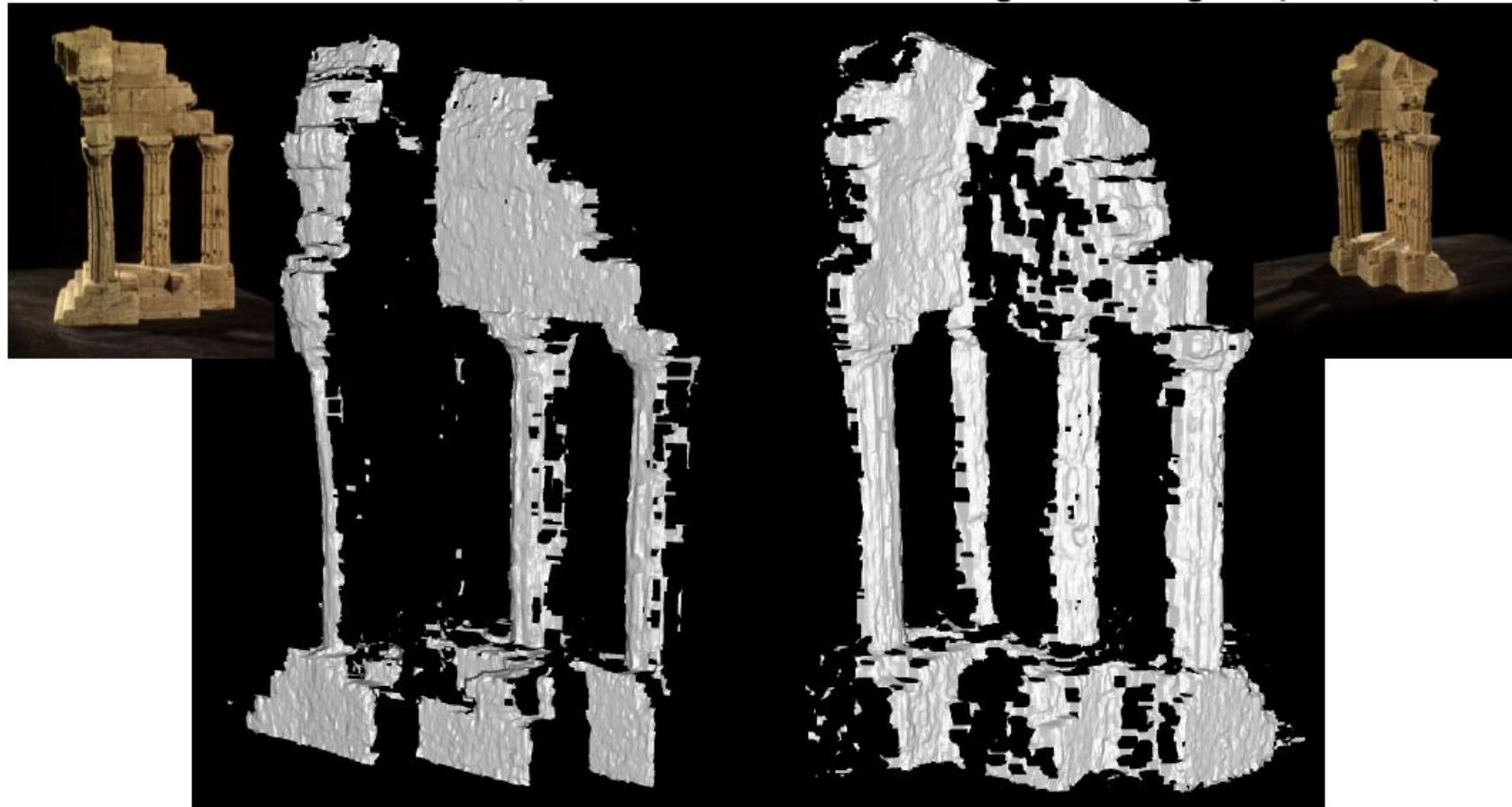
Триангуляция карты глубины



По карте глубины можно построить сеточную модель:

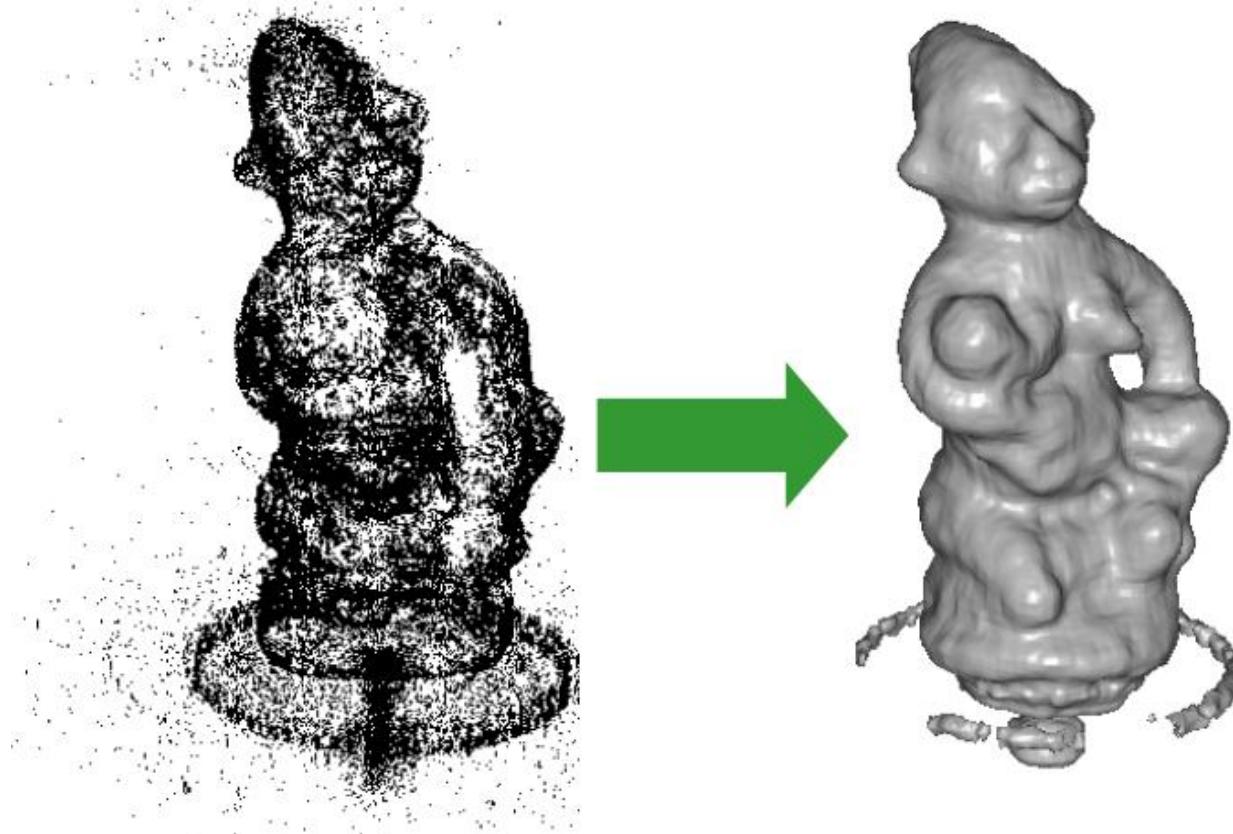
- Каждому пикслю изображения сопоставляется вершина
- Соседние вершины объединяются
- Глубина вершины берется из карты глубины
- Слишком узкие и наклоненные треугольники отбрасываются

Карты глубины



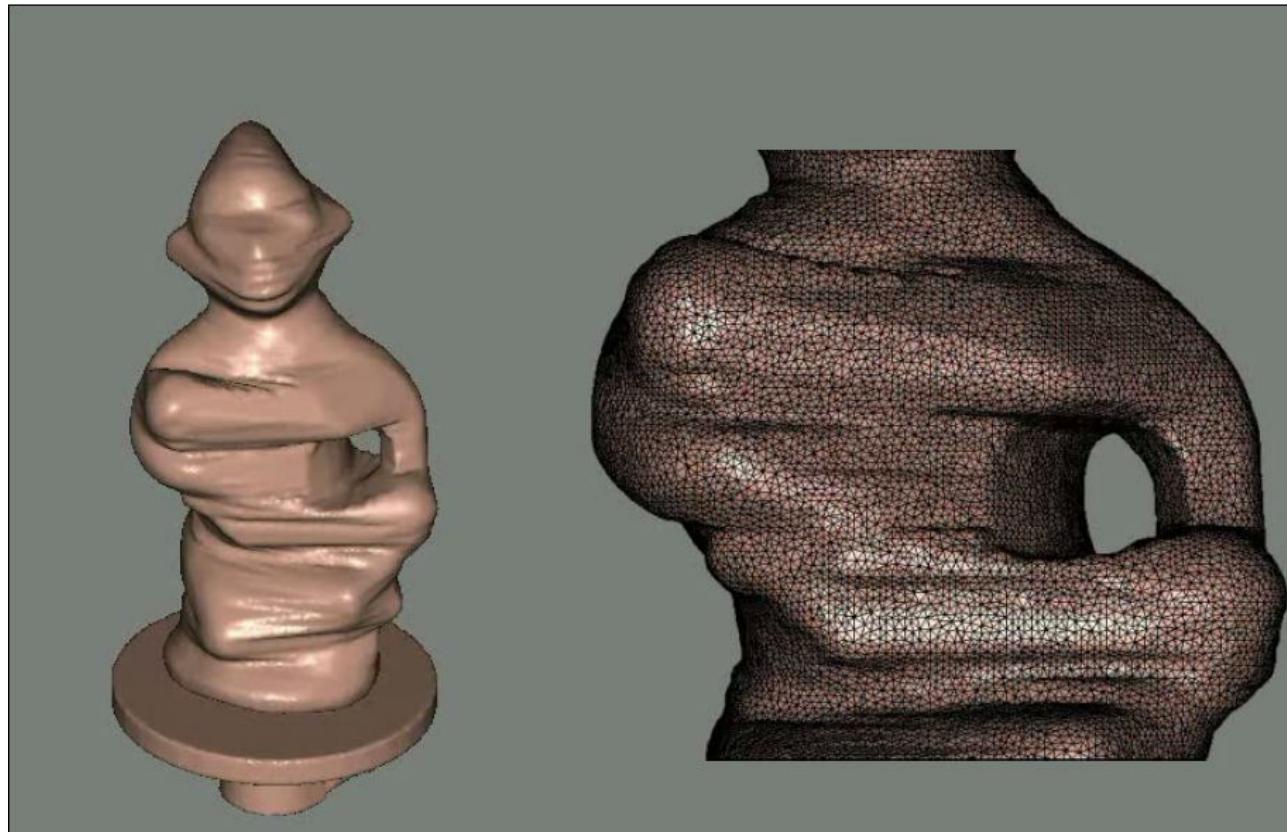
Пример построенных моделей по опорным кадрам

Извлечение поверхности



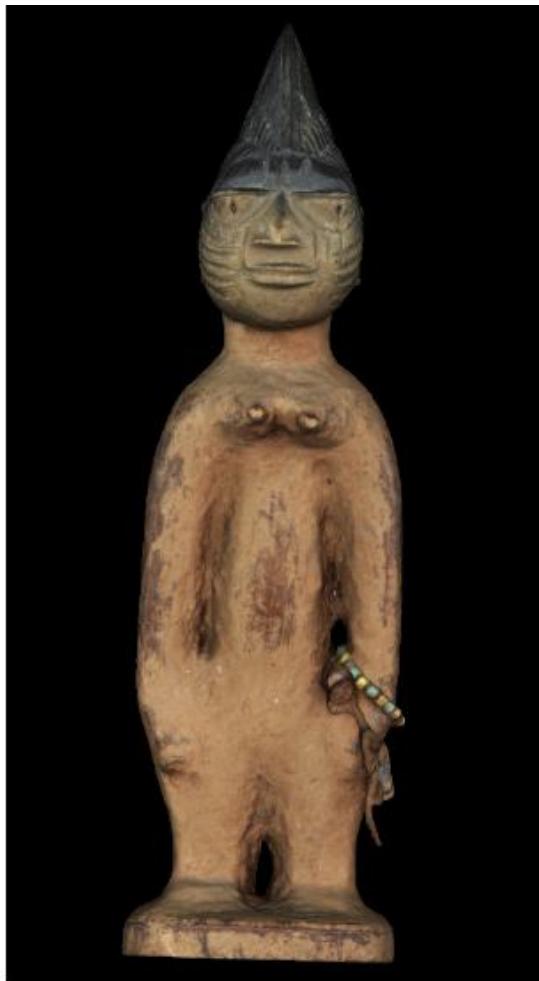
- Можно объединять отдельные сетки с каждого опорного изображения
- Можно построить сетку (триангулировать) по облаку точек (картам глубин) со всех опорных изображений

Деформируемые поверхности



- Строим сетку по облаку точек
- Оптимизируем сетку, добиваясь фото-согласованности поверхности
- Очень медленно, бывают локальные минимумы

Пример работы



Оптимизация сетки позволяет удалить ошибки из облака точек, и получить гладкие и точные поверхности

План лекции



- Введение
- Многовидовое пассивное стерео
 - Объединение сеток
 - **Разрастание регионов**
- Приложение: реконструкция головы человека

Методы на основе фрагментов



[Lhuillier and Quan,
PAMI 05]



[Furukawa and Ponce,
CVPR 07]



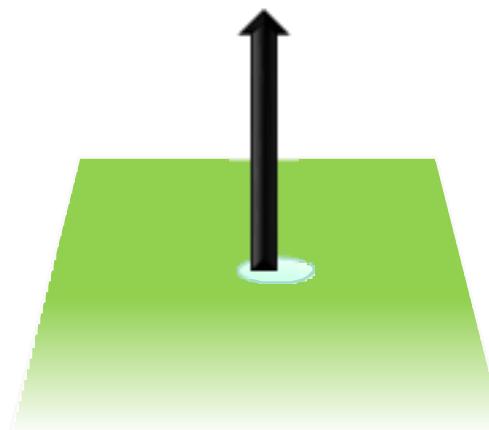
[Habbecke and Kobbelt,
CVPR 07]

Одни из самых точных современных алгоритмов основаны на фрагментах

Определение фрагмента



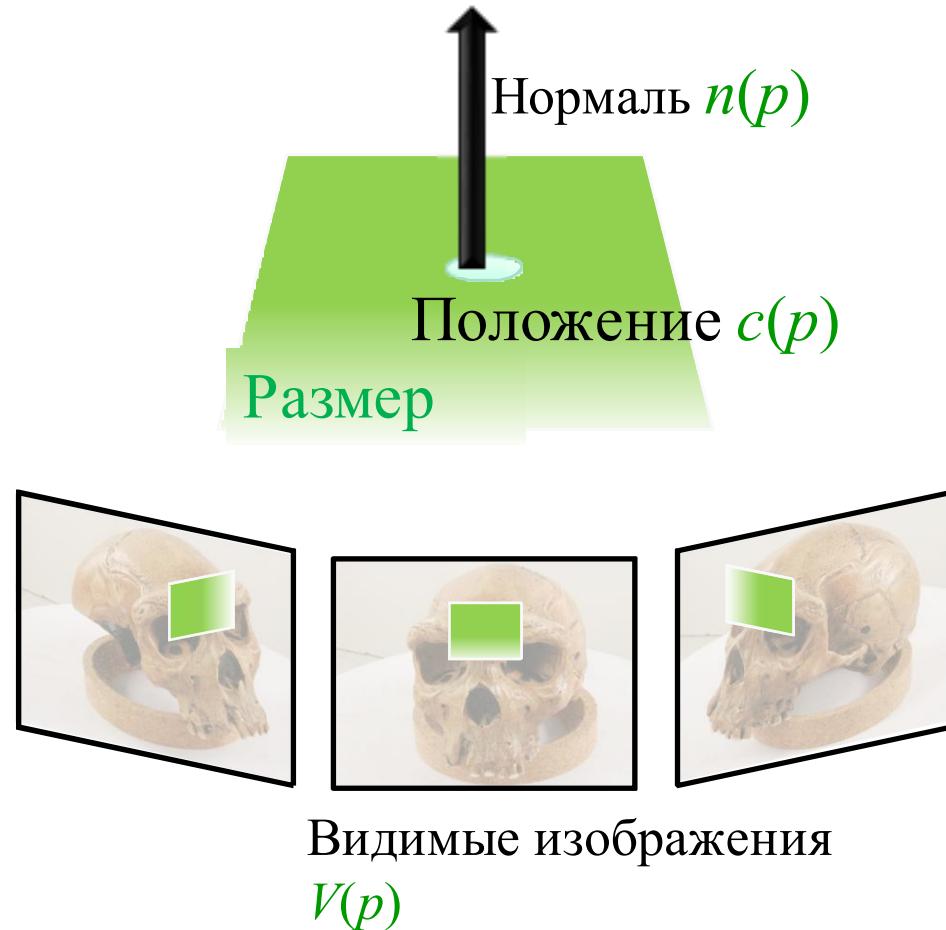
- «Фрагмент» (patch) – это модель небольшого участка поверхности объекта
- Поскольку участок небольшой, то можно описать его плоской поверхностью (квадратом)



Определение фрагмента



- Фрагмент p определен
 - Положением $c(p)$
 - Нормалью $n(p)$
 - Видимыми изображениями $V(p)$
- Размер выбирается так, чтобы p был примерно 9x9 пикселей в $V(p)$

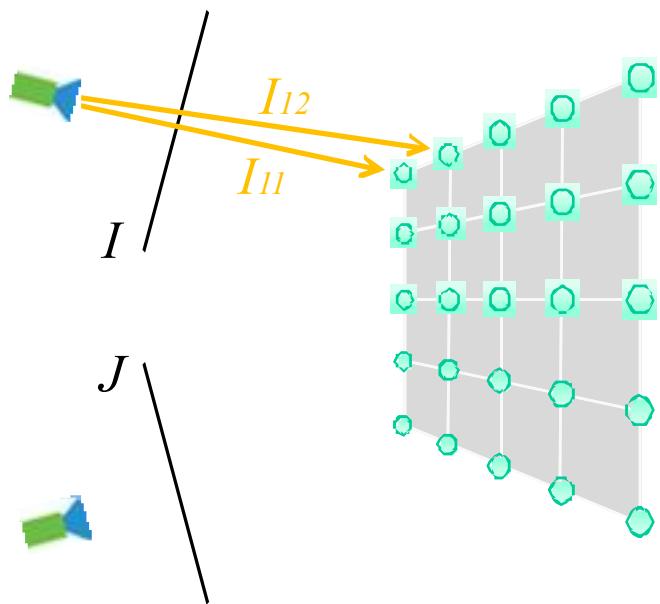


Фотосогласованность



- Вычислим фотосогласованность $N(I, J, p)$ фрагмента p между изображениями I и J

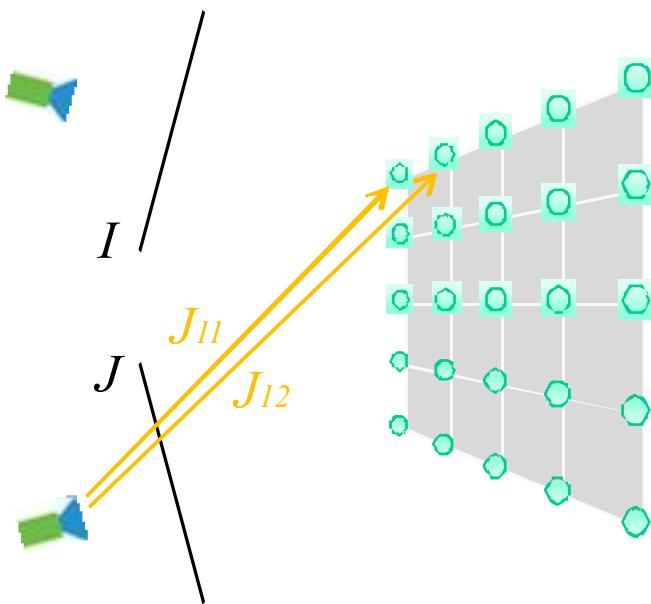
I_{xy} : цвет в изображении I



Фотосогласованность



- Вычислим фотосогласованность $N(I, J, p)$ фрагмента p между изображениями I и J



I_{xy} : цвет в изображении I

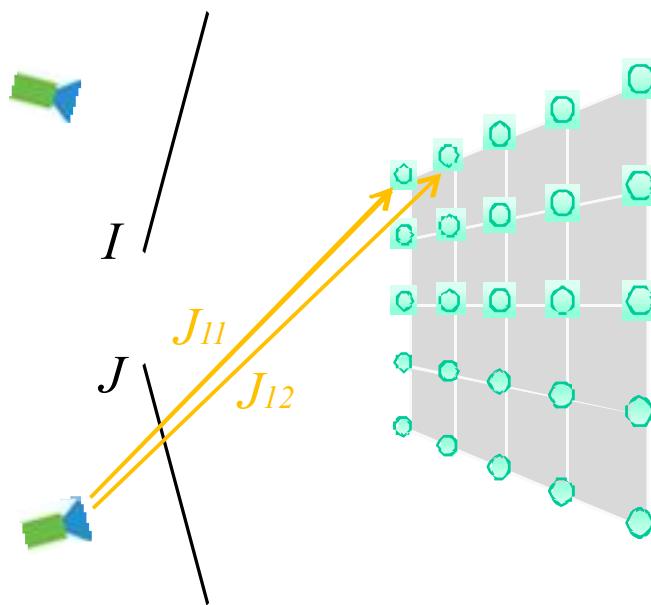
J_{xy} : цвет в изображении J

$$N(I, J, p) = \frac{\sum (I_{xy} - \bar{I}_{xy}) \cdot (J_{xy} - \bar{J}_{xy})}{\sqrt{(I_{xy} - \bar{I}_{xy})^2} \sqrt{(J_{xy} - \bar{J}_{xy})^2}}$$

Фотосогласованность



- Перейдём от случаев 2x изображений к случаю многих изображений:



$$N(I, J, p) = \frac{\sum (I_{xy} - \bar{I}_{xy}) \cdot (J_{xy} - \bar{J}_{xy})}{\sqrt{(I_{xy} - \bar{I}_{xy})^2} \sqrt{(J_{xy} - \bar{J}_{xy})^2}}$$

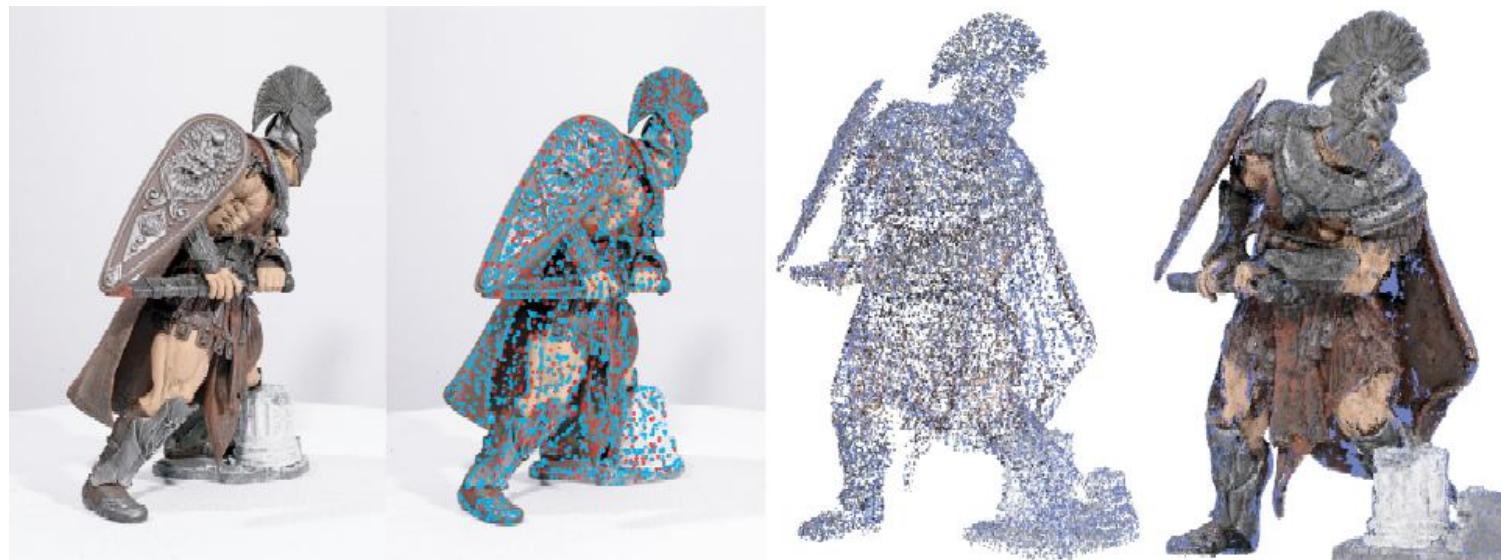


$$N(p) = \frac{\sum_{i=1}^n \sum_{j=i+1}^n N(I_i, I_j, p)}{(n+1)n/2}$$

Алгоритм на основе фрагментов



- #1. Поиск особых точек
- #2. Инициализация точек
- #3. Расширение фрагментов и фильтрация



Input image

#1

#2

#3

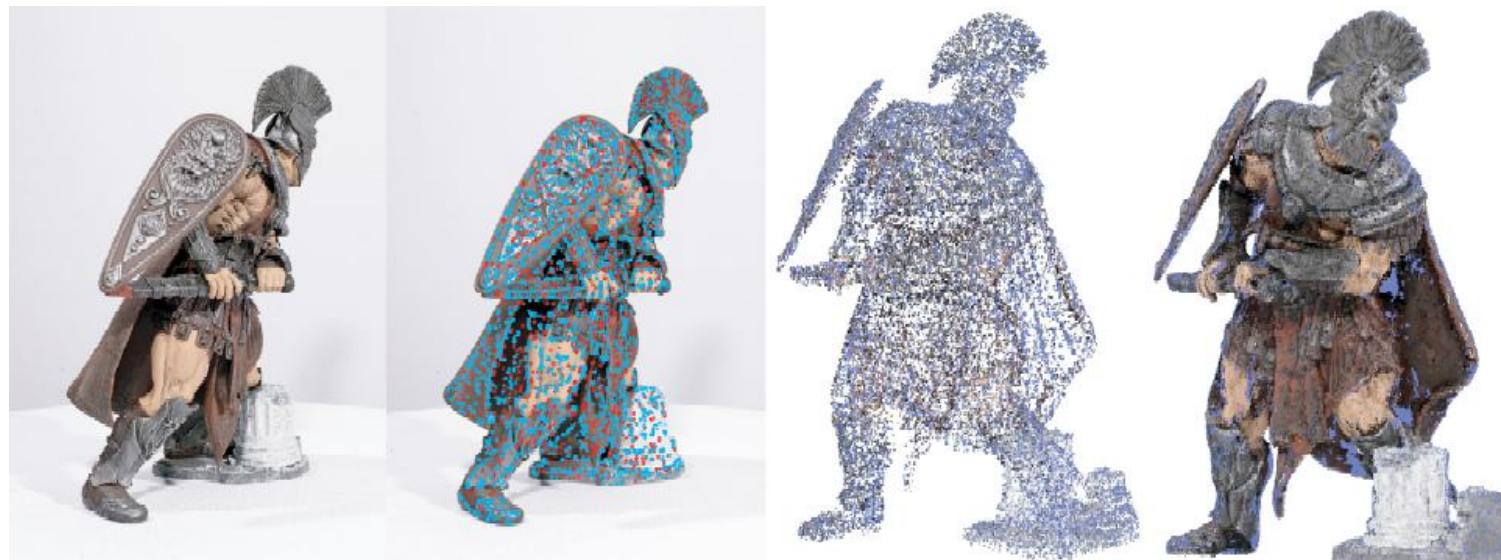
Алгоритм на основе фрагментов



#1. Поиск особых точек

#2. Инициализация фрагментов

#3. Расширение фрагментов и фильтрация



Input image

#1

#2

#3

Поиск особых точек



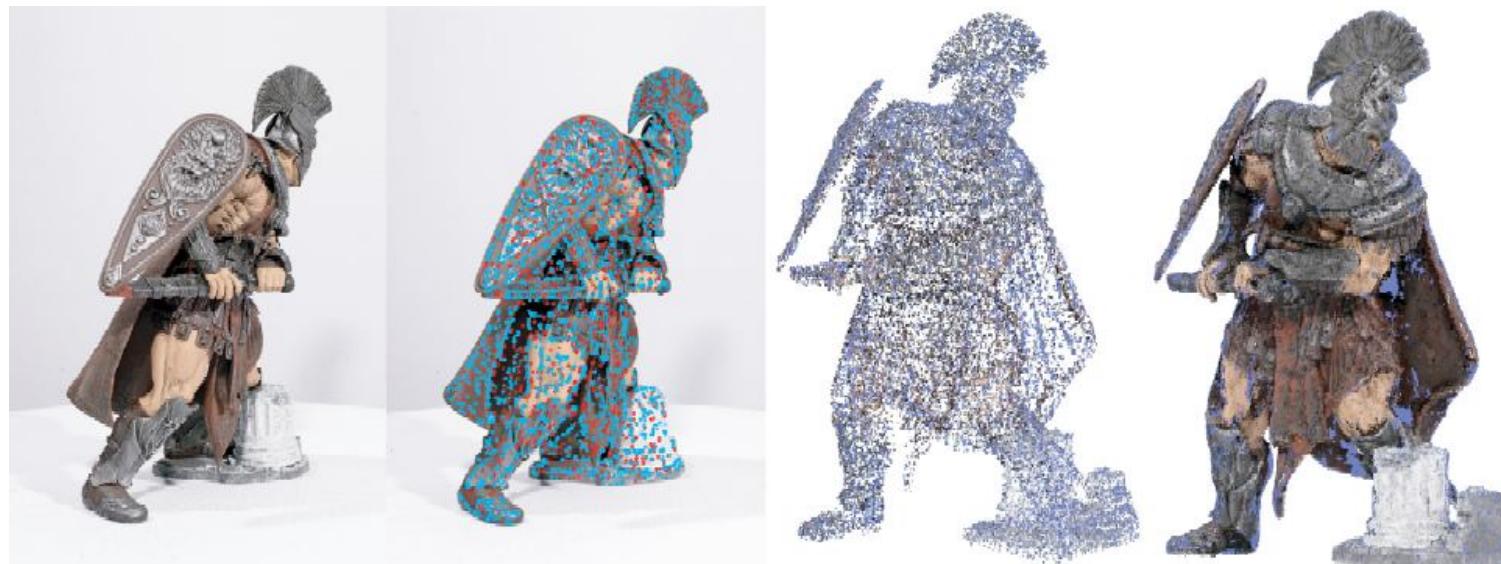
- Harris corner detector (corners)
- Difference of Gaussian (blobs)



Алгоритм на основе фрагментов



- #1. Поиск особых точек
- #2. Инициализация фрагментов
- #3. Расширение фрагментов и фильтрация



Input image

#1

#2

#3

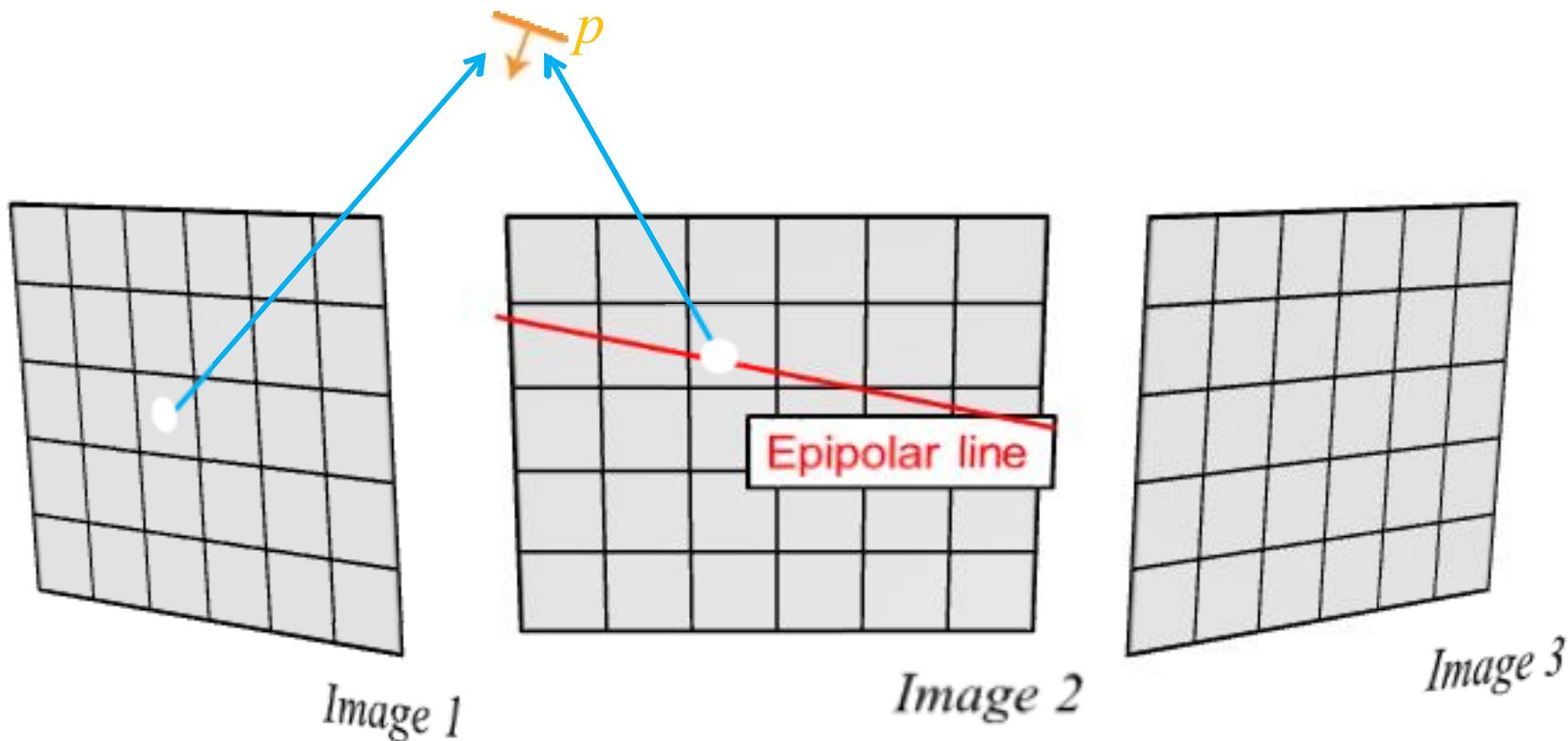
Инициализация фрагментов



$c(p)$: триангуляция

$n(p)$:

$V(p)$:



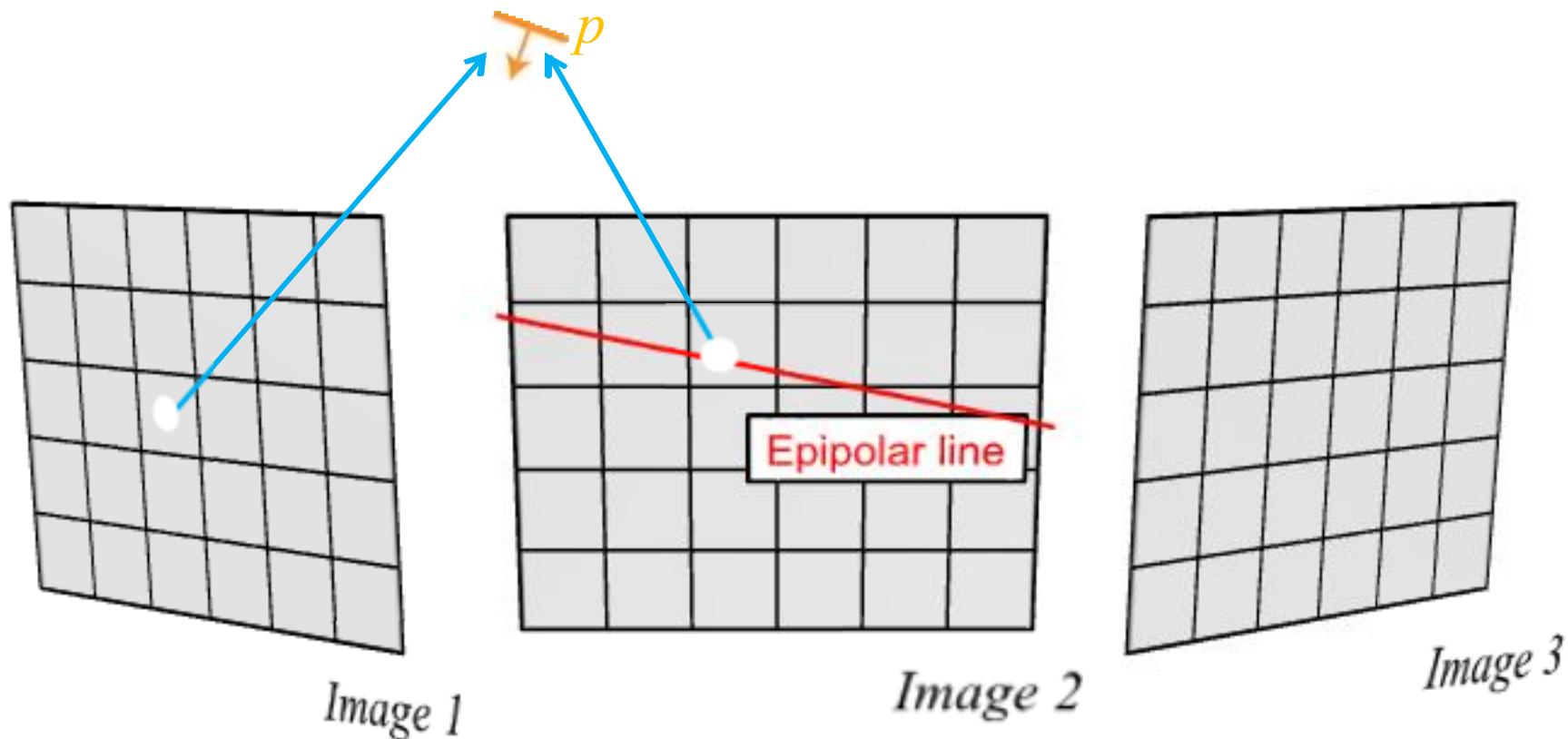
Инициализация фрагментов



$c(p)$: триангуляция

$n(p)$: параллельно к $Image1$

$V(p)$: $\{Image1, Image2\}$



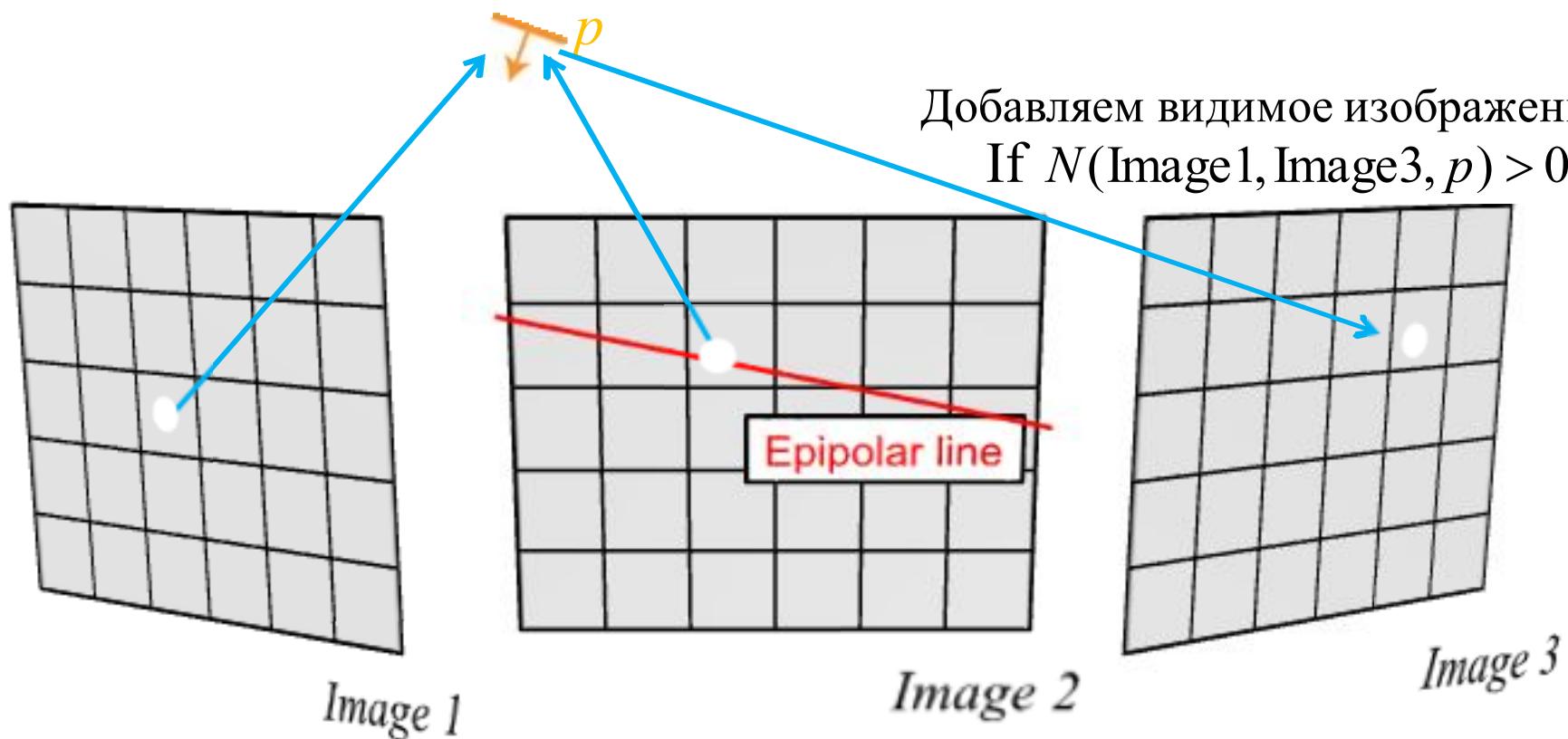
Инициализация фрагментов



$c(p)$: триангуляция

$n(p)$: параллельно к $Image1$

$V(p)$: $\{Image1, Image2, Image3\}$



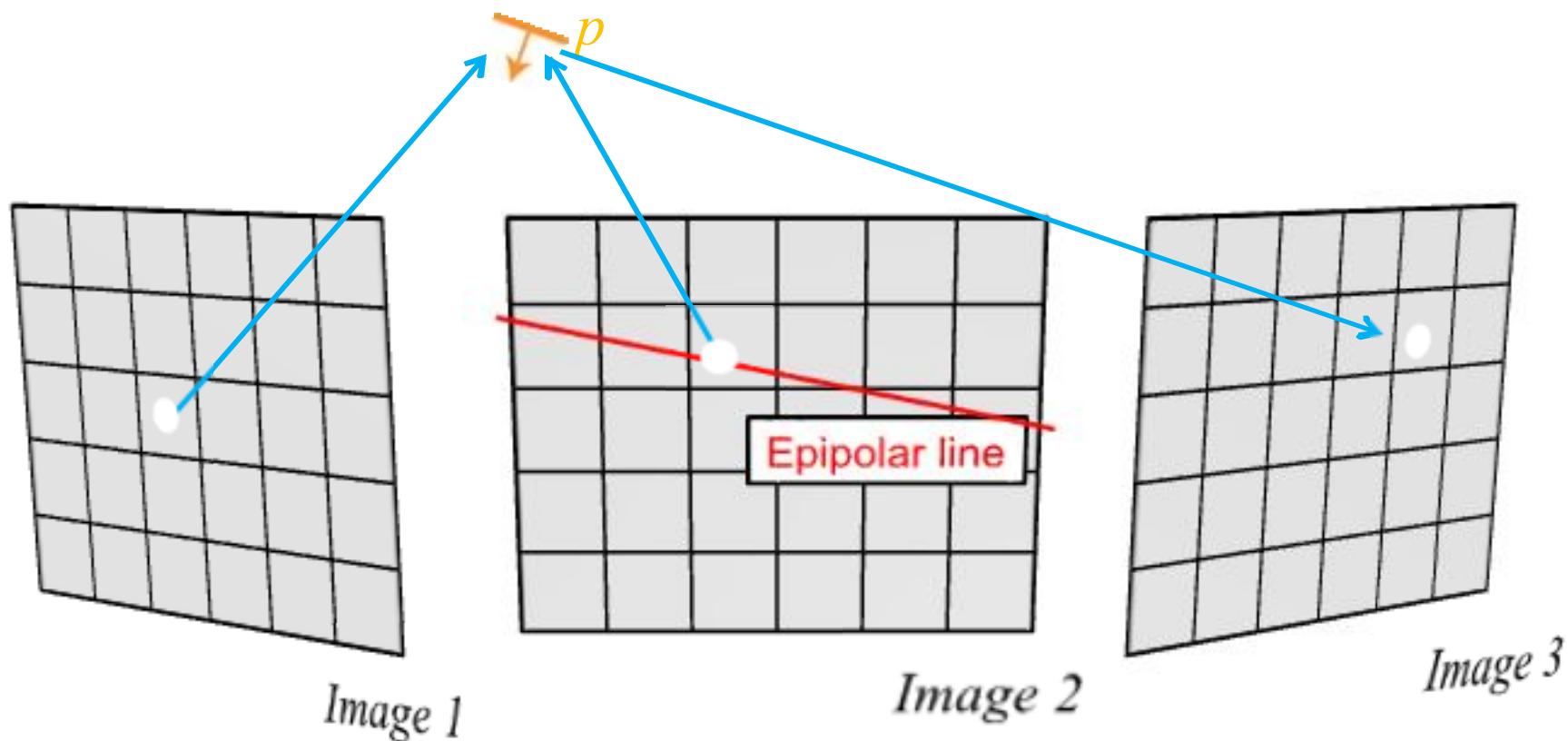
Инициализация фрагментов



$c(p)$: триангуляция

$n(p)$: параллельно к $Image1$

$V(p)$: $\{Image1, Image2, Image3\}$



Инициализация фрагментов

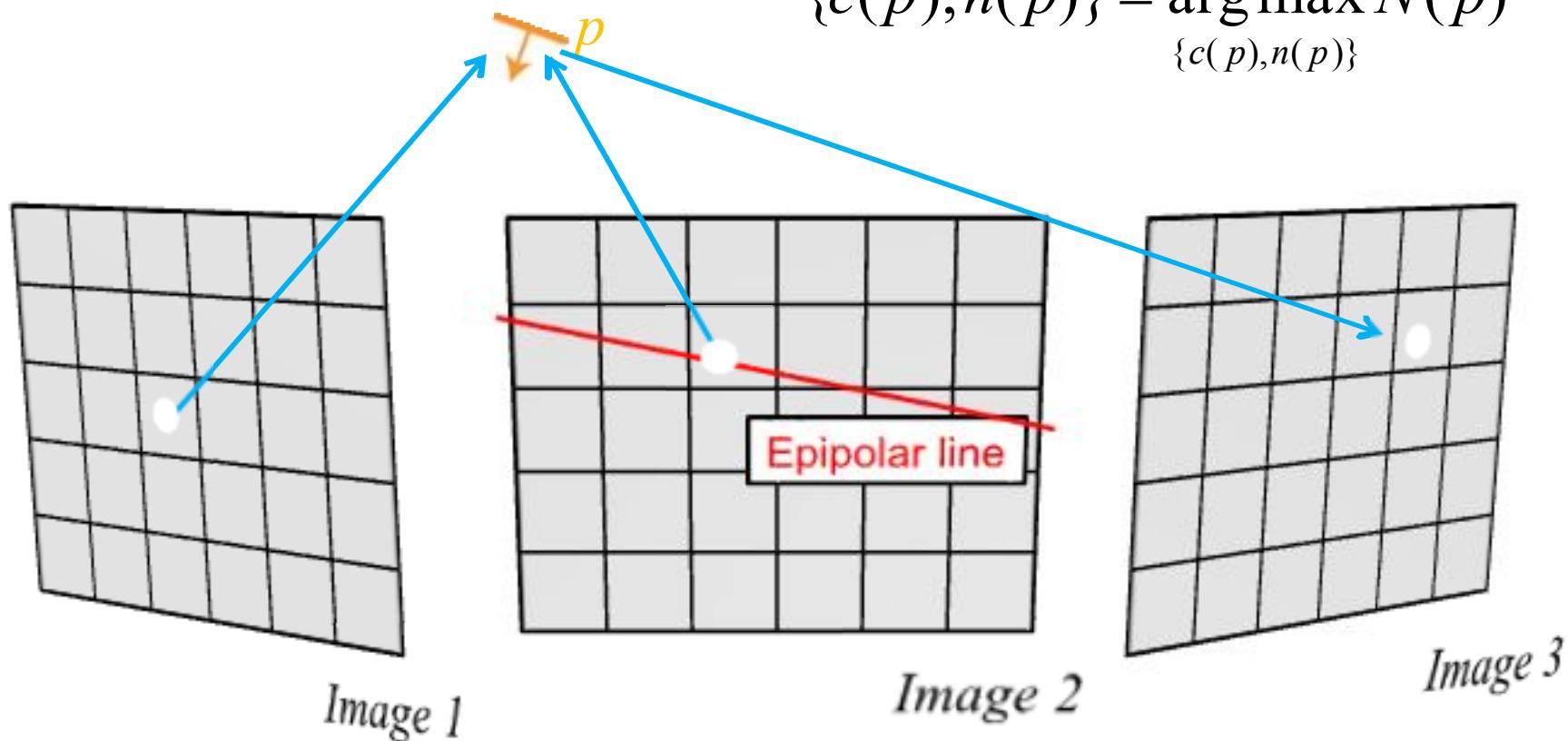


$c(p)$: уточним

$n(p)$: уточним

$V(p)$: $\{Image1, Image2, Image3\}$

$$\{c(p), n(p)\} = \arg \max_{\{c(p), n(p)\}} N(p)$$



Инициализация фрагментов

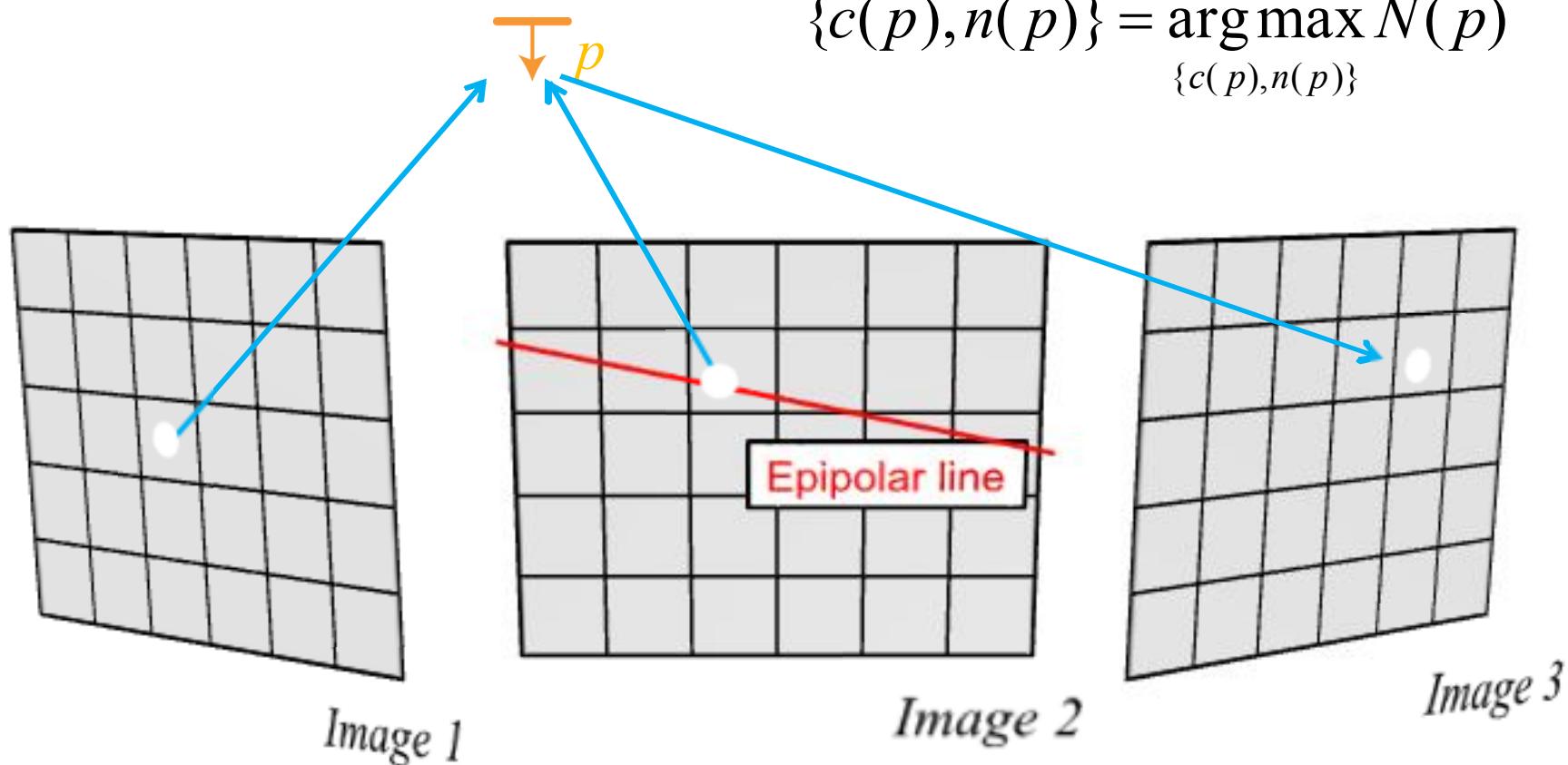


$c(p)$: уточним

$n(p)$: уточним

$V(p)$: $\{Image1, Image2, Image3\}$

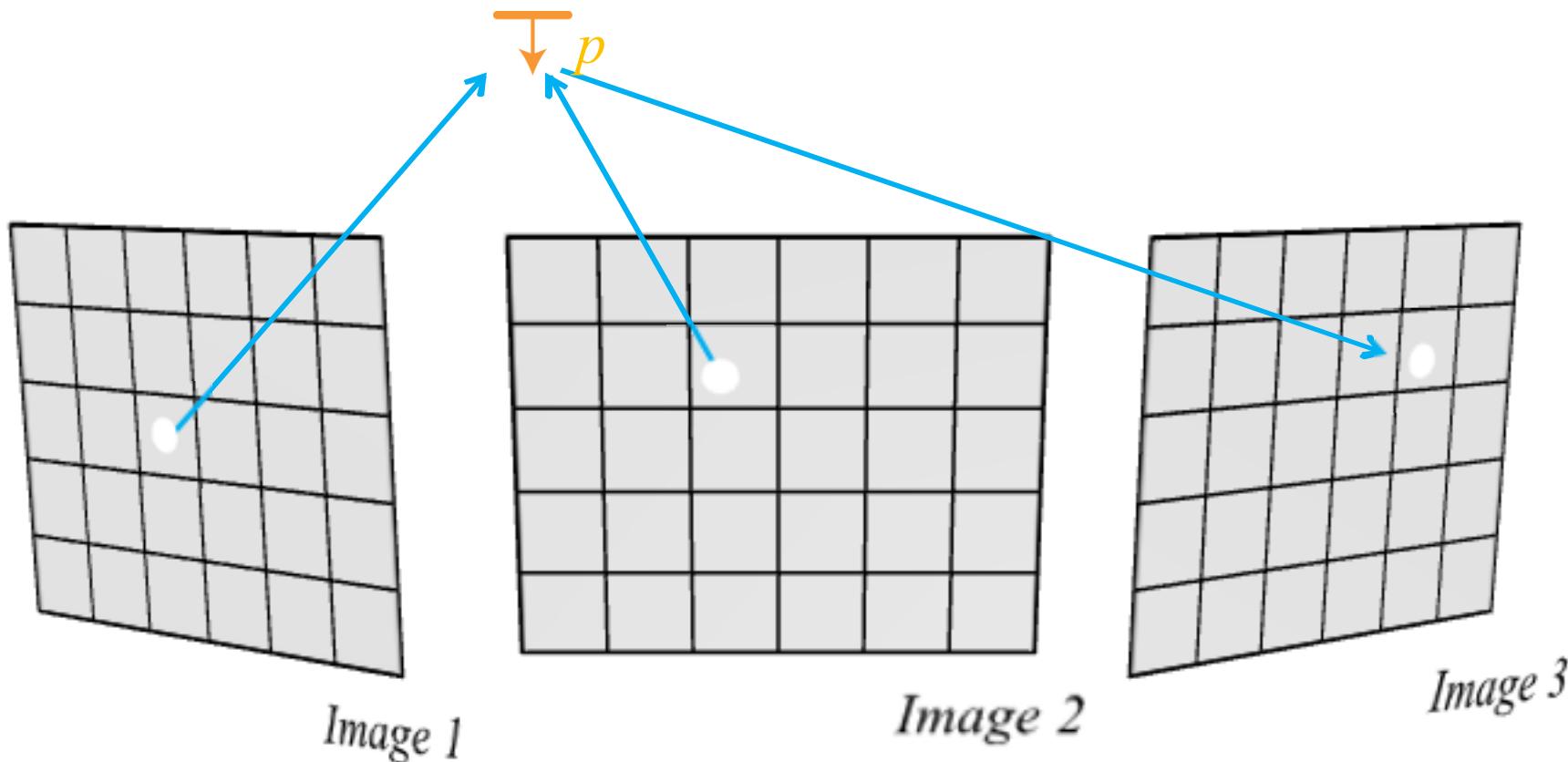
$$\{c(p), n(p)\} = \arg \max_{\{c(p), n(p)\}} N(p)$$



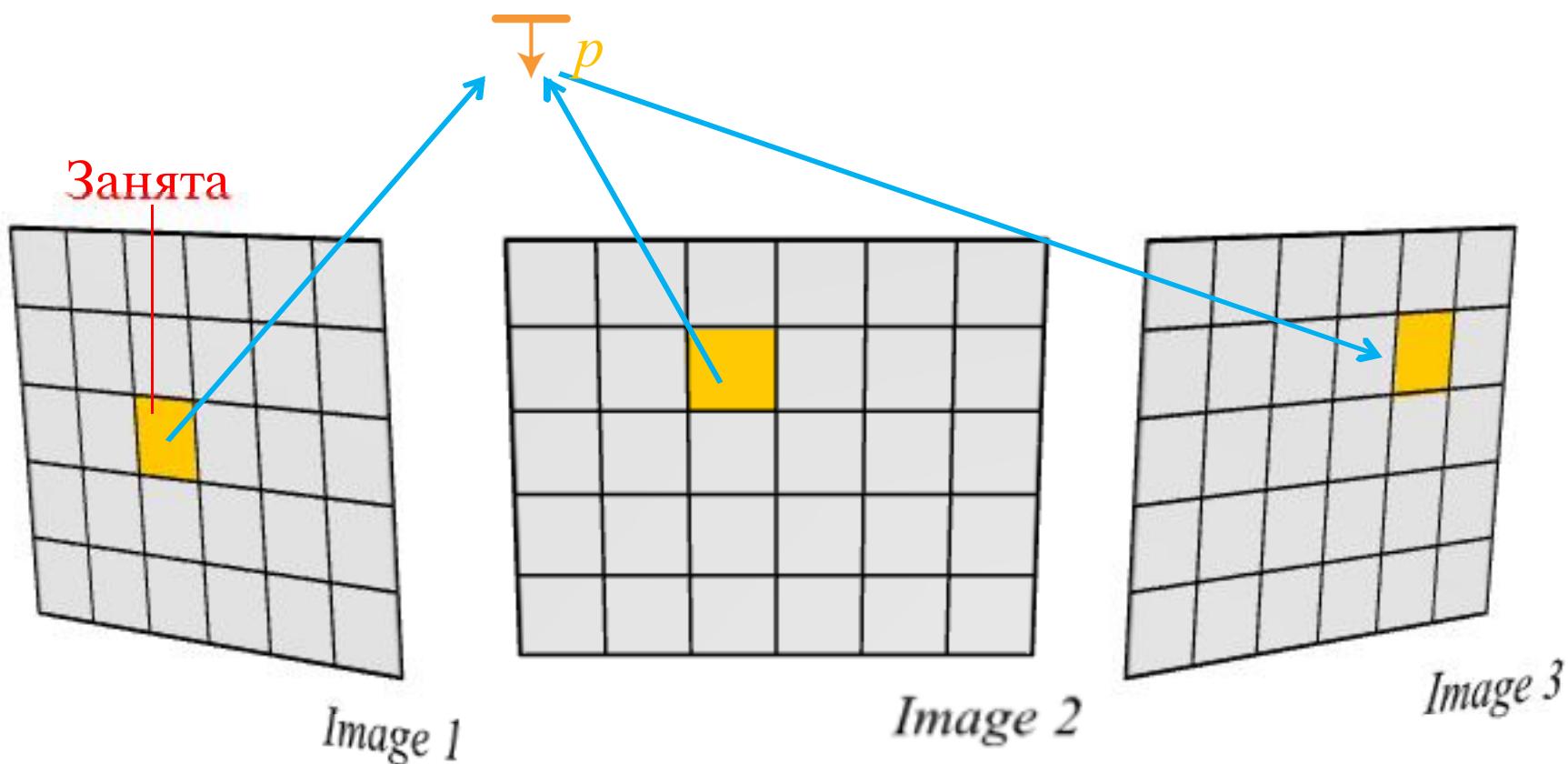
Инициализация фрагментов



Верификация
(обновим $V(p)$ и проверим $|V(p)| \geq 3$)



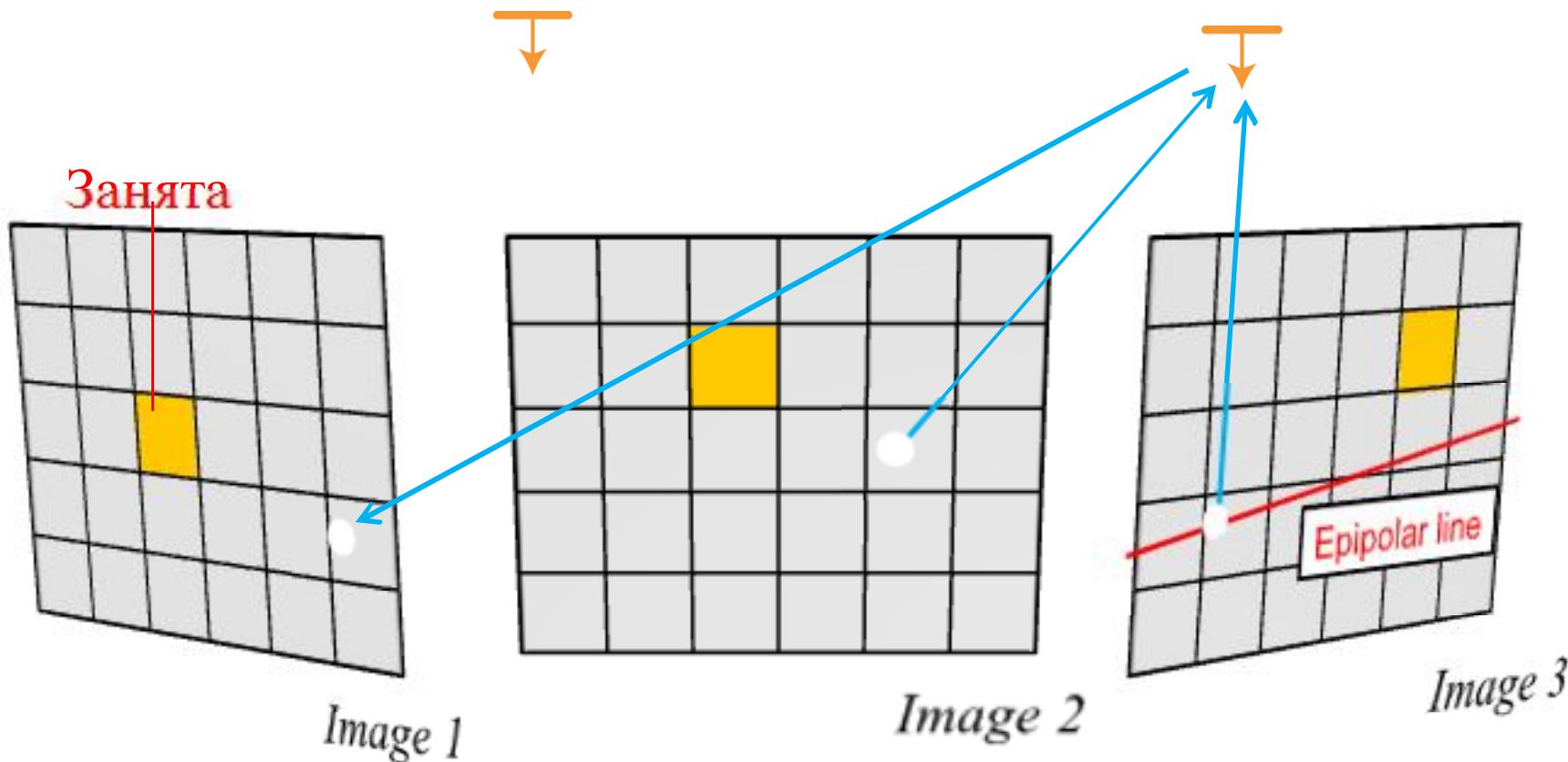
Инициализация фрагментов



Инициализация фрагментов



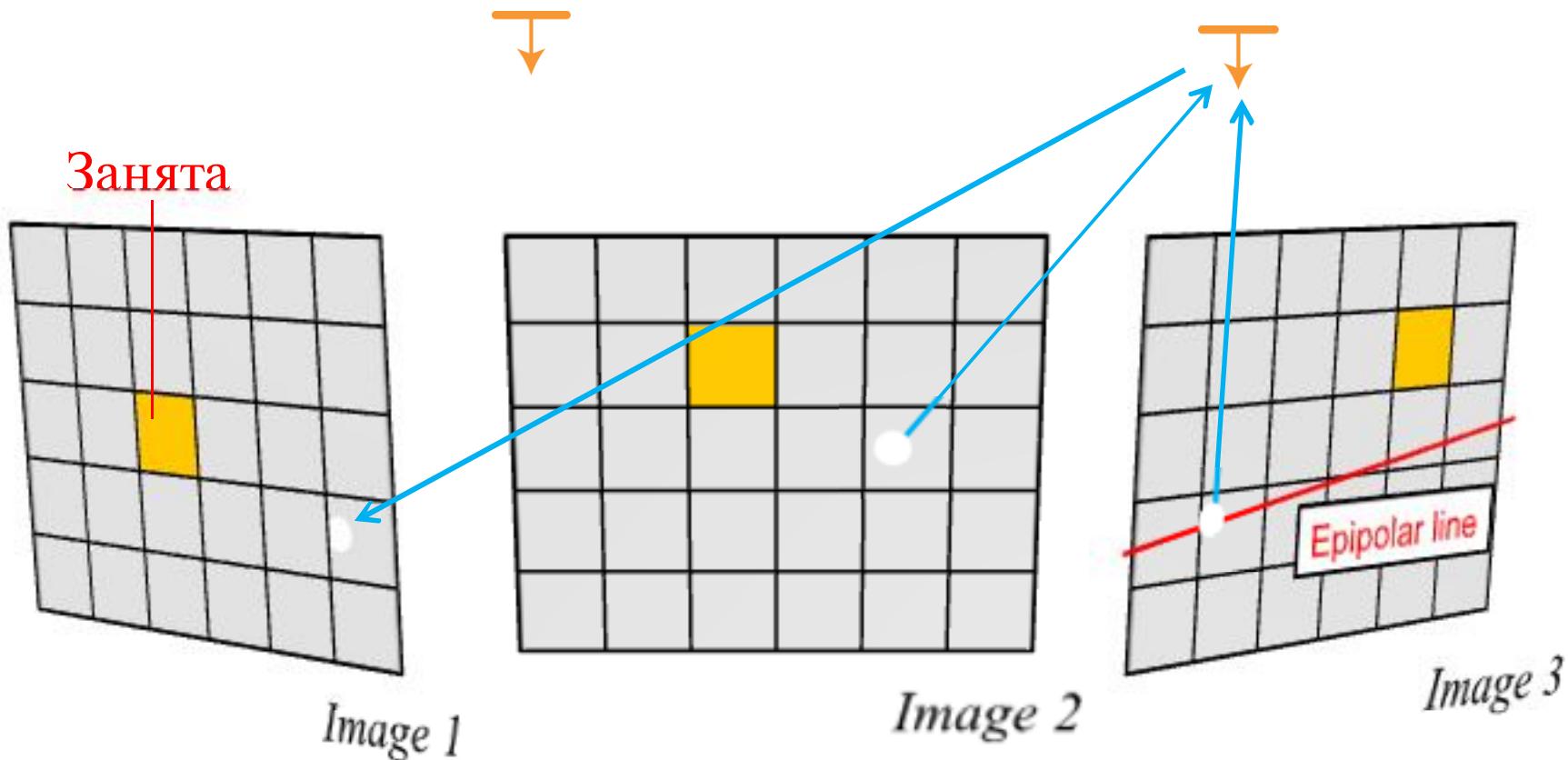
- Повторяем для всех найденных особых точек



Инициализация фрагментов



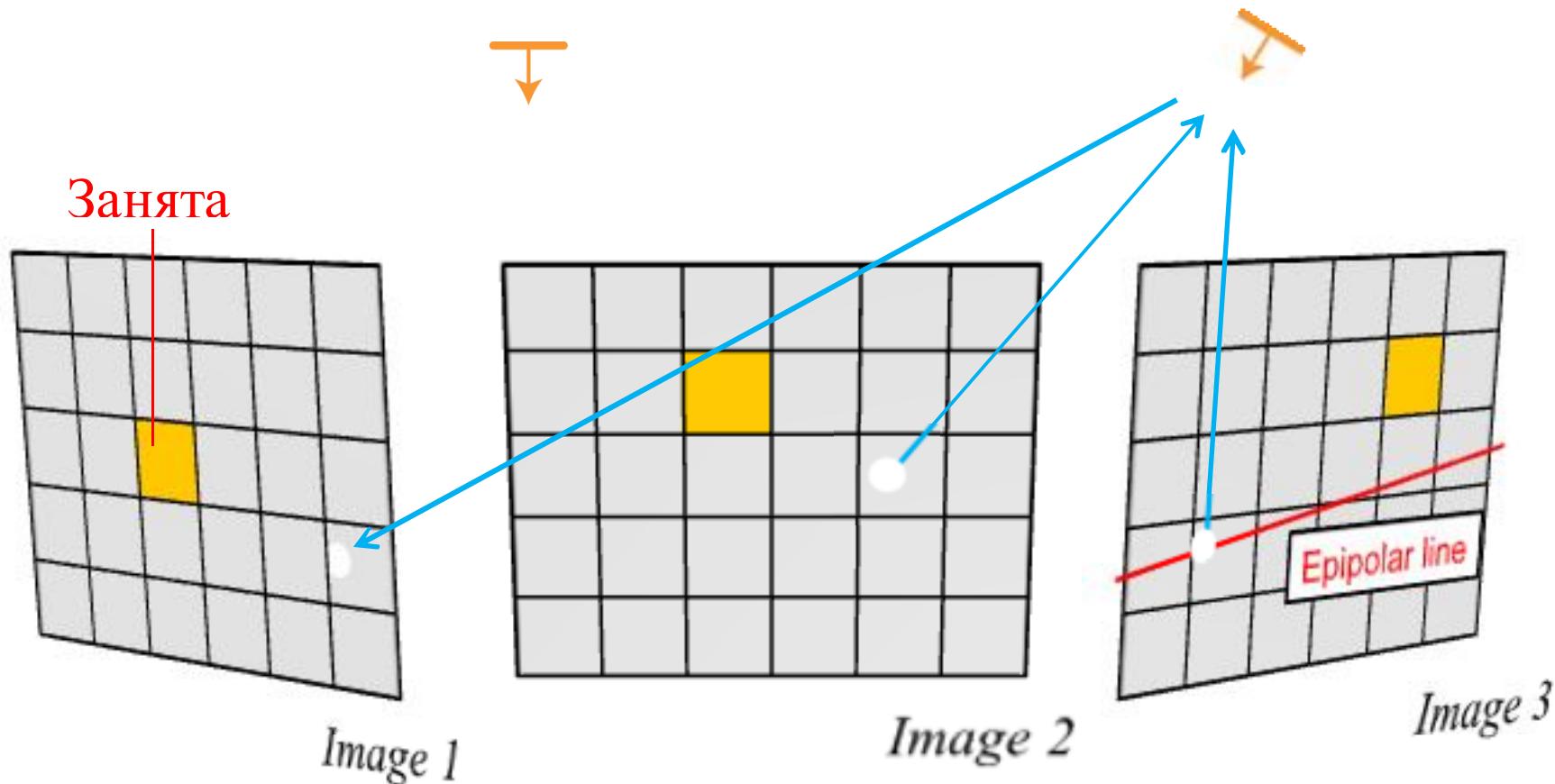
- Повторяем для всех найденных особых точек



Инициализация фрагментов



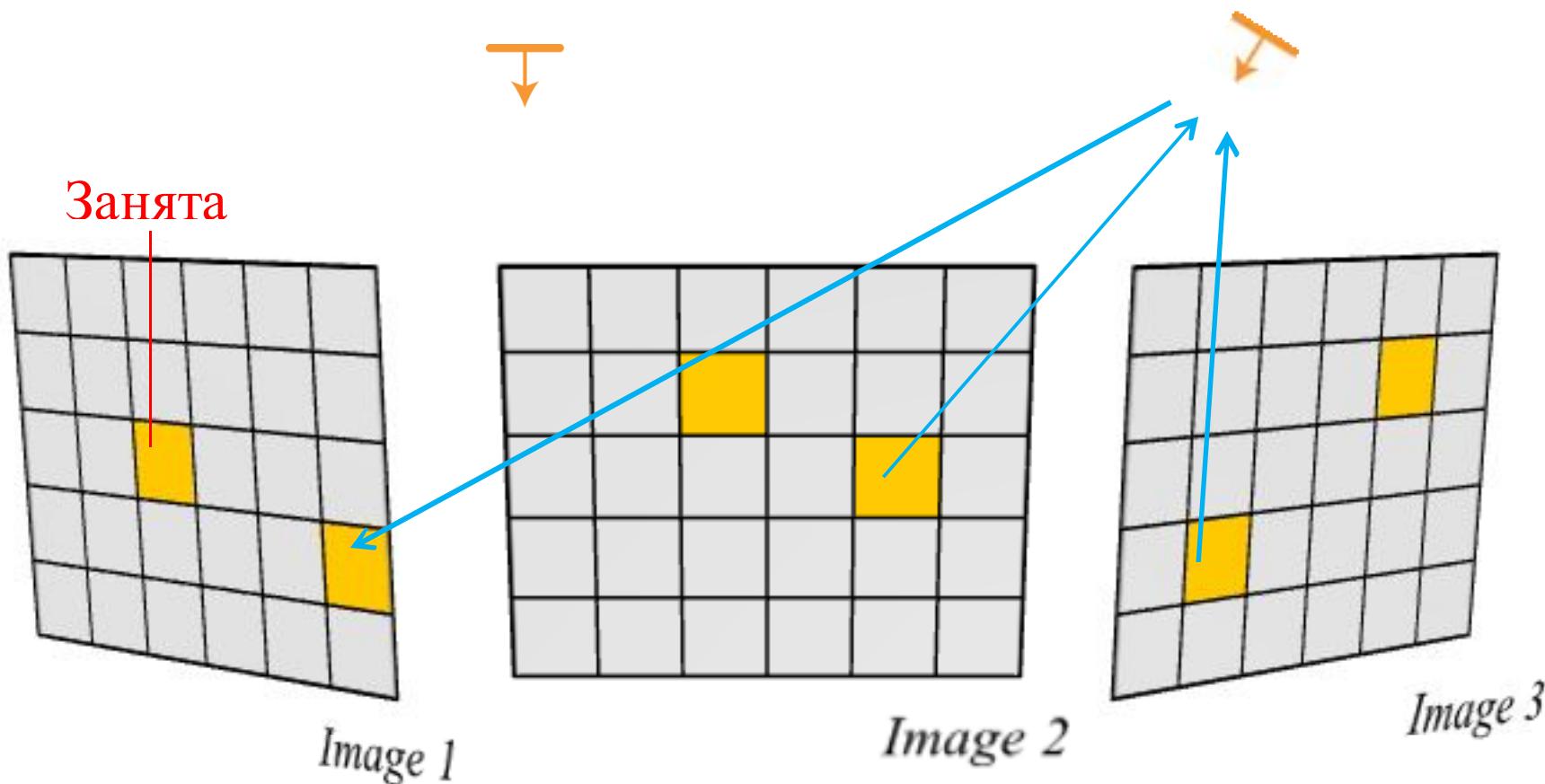
- Повторяем для всех найденных особых точек



Инициализация фрагментов



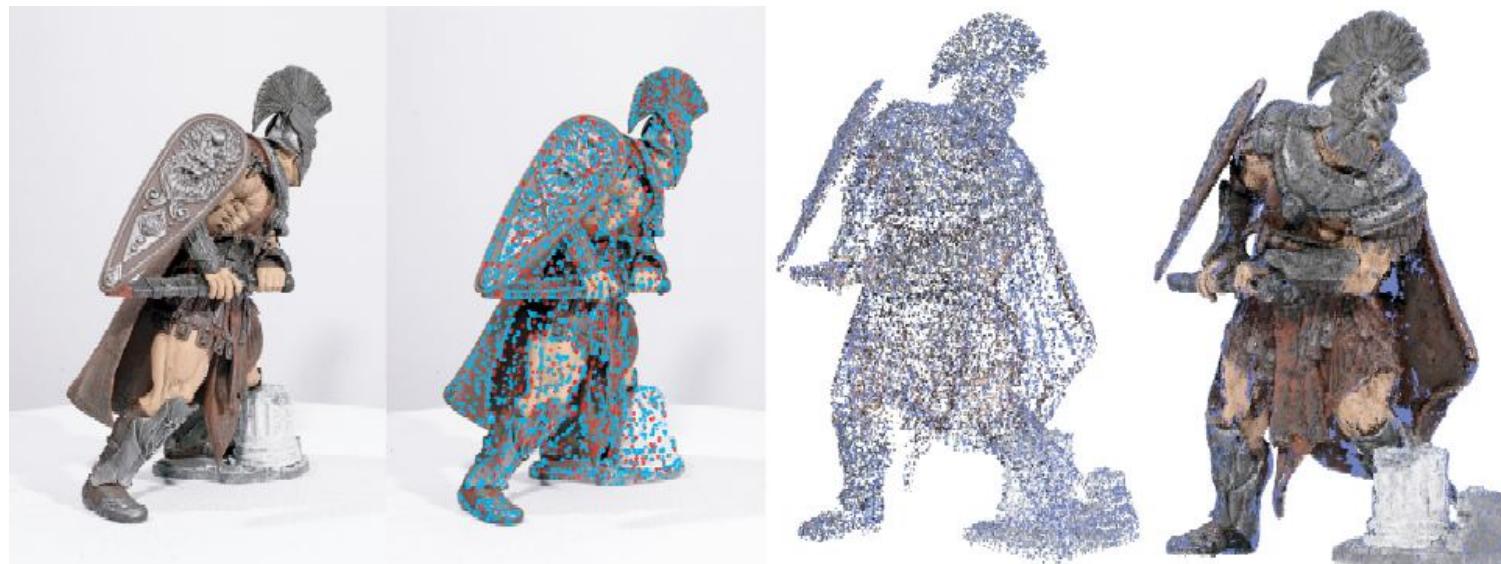
- Повторяем для всех найденных особых точек



Алгоритм на основе фрагментов



- #1. Поиск особых точек
- #2. Инициализация фрагментов
- #3. Расширение фрагментов и фильтрация



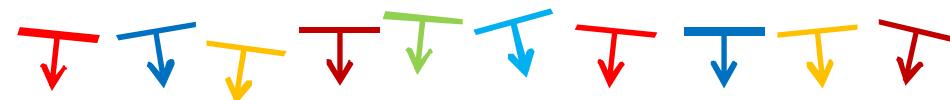
Input image

#1

#2

#3

Расширение фрагментов



Занятый пиксель

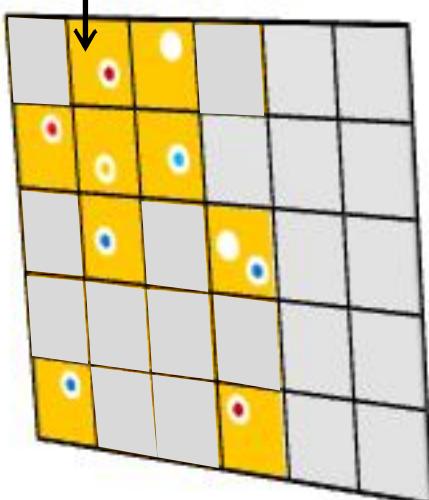


Image 1

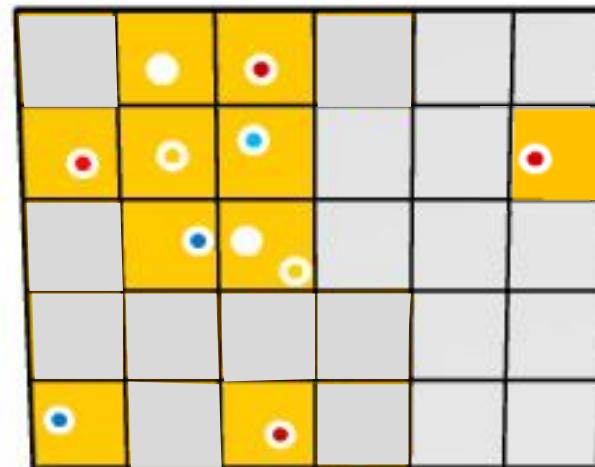


Image 2

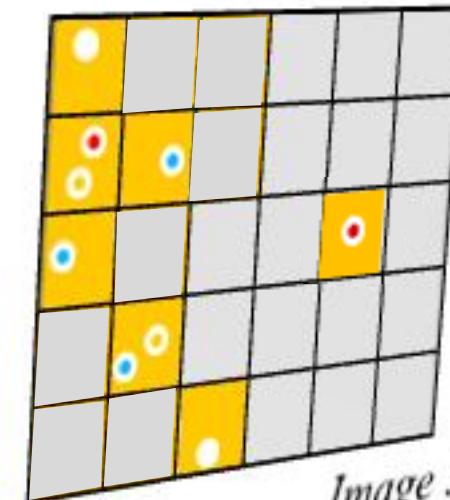
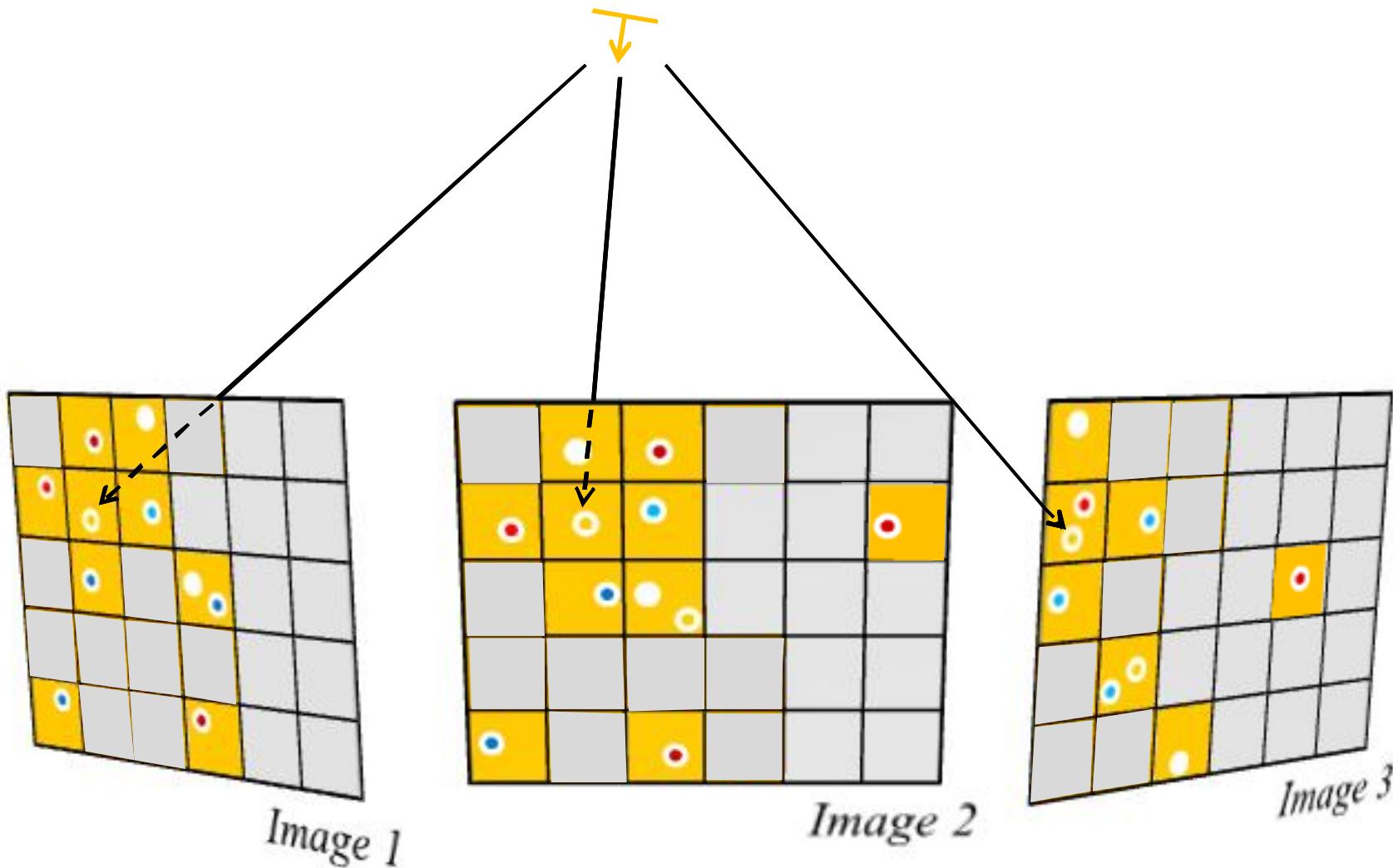


Image 3

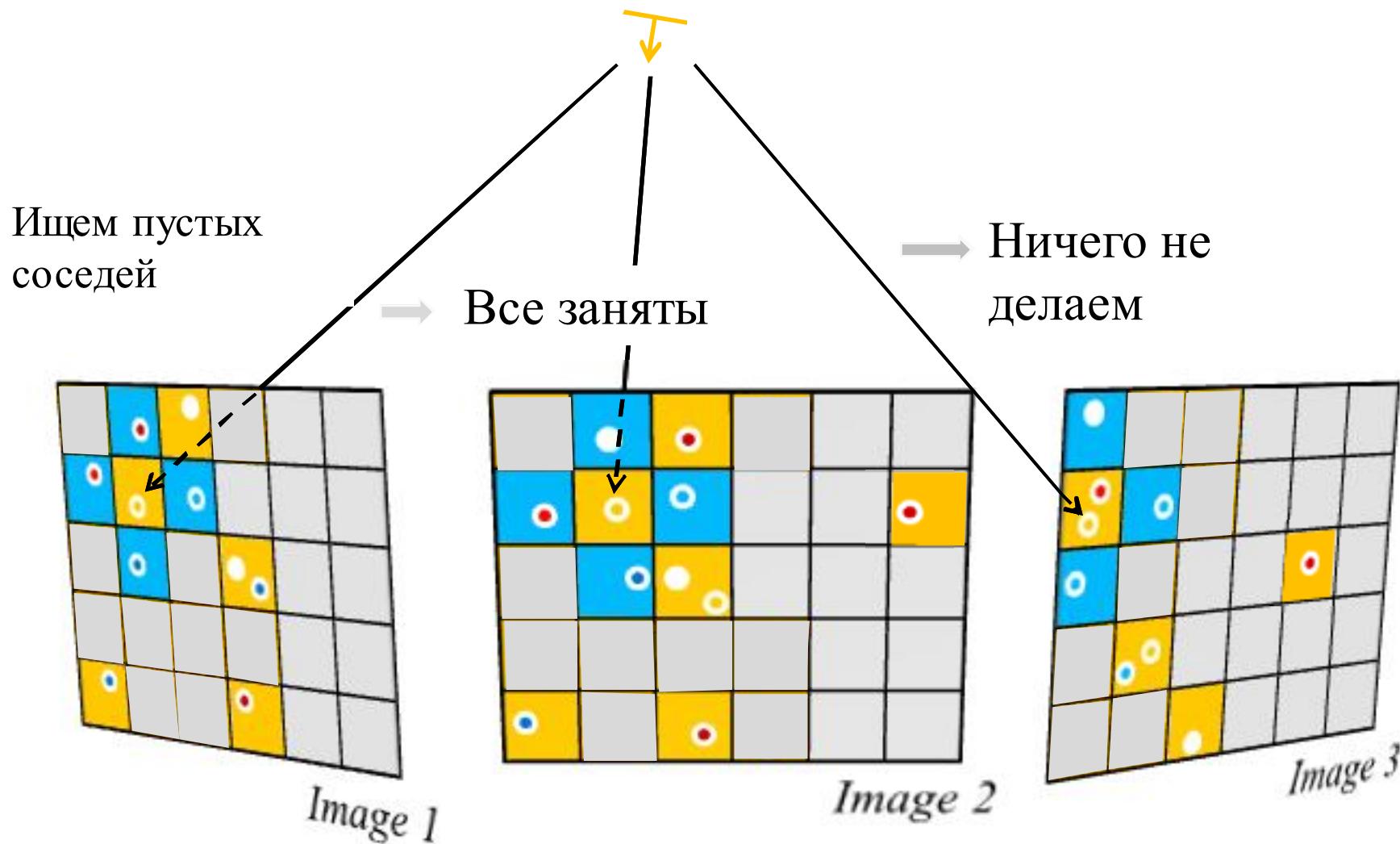
Расширение фрагментов



Выбираем фрагмент



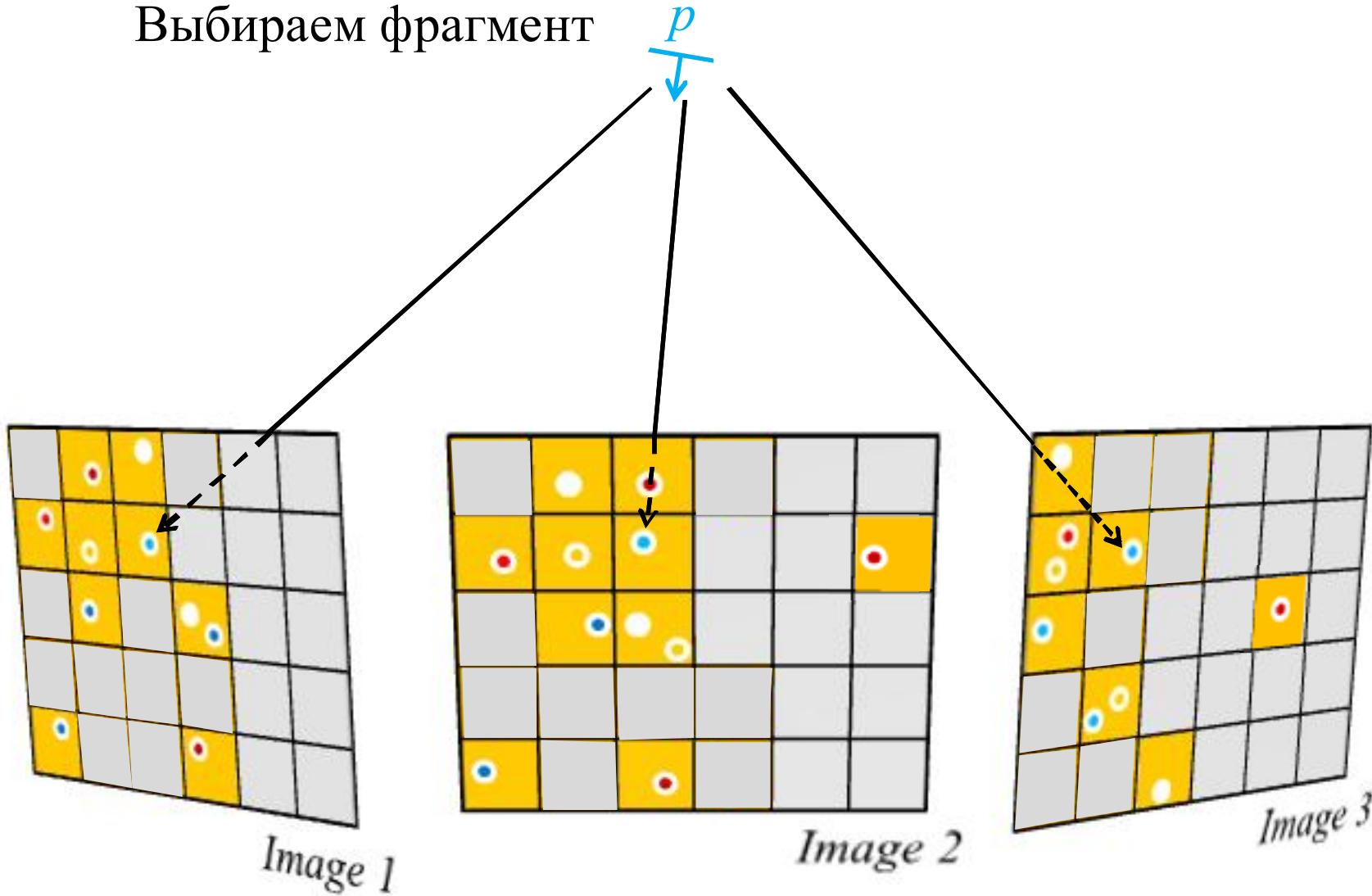
Расширение фрагментов



Расширение фрагментов



Выбираем фрагмент



Расширение фрагментов



Ищем
пустых
соседей

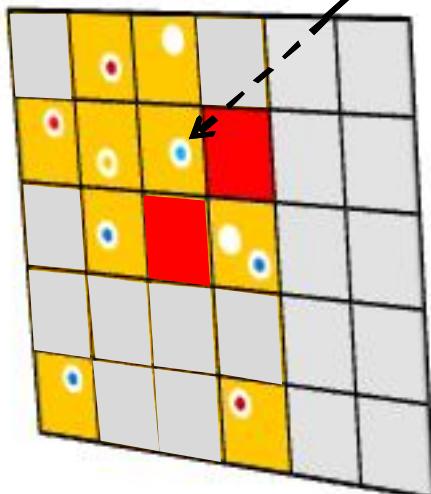


Image 1

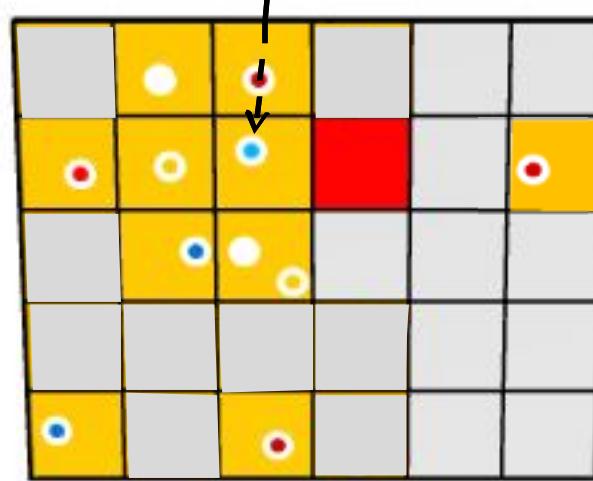


Image 2

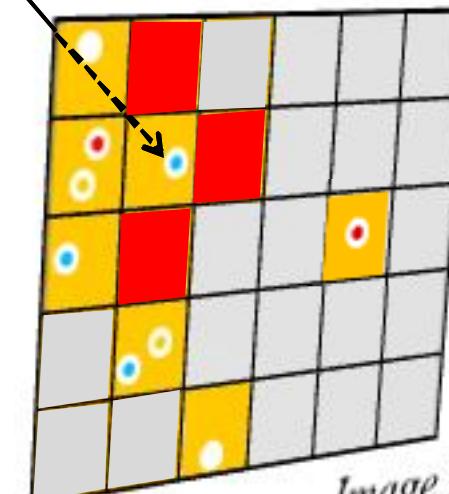


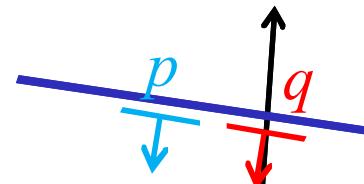
Image 3

p

Расширение фрагментов



Реконструируем
фрагменты во
всех соседних
пикселях



$c(q)$: {пересечение
касательной
плоскости и луча}

$n(q)$:

$V(q)$:

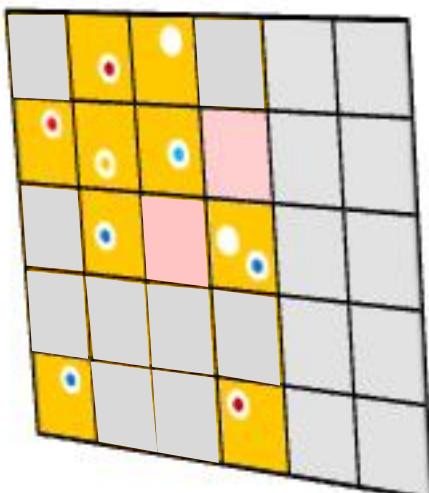


Image 1

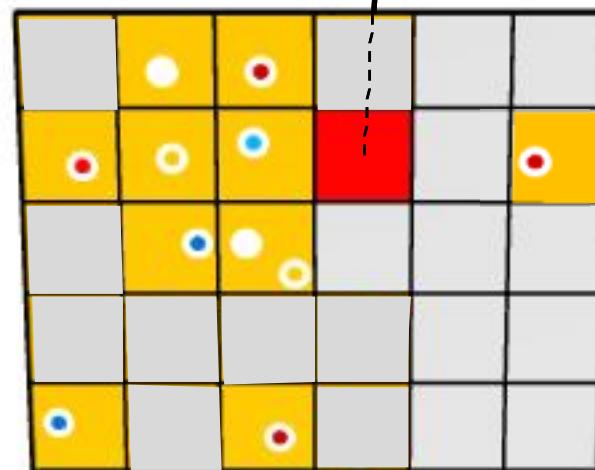


Image 2

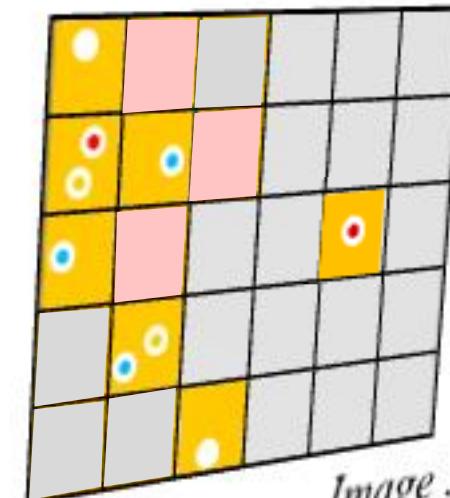
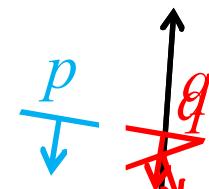


Image 3

Расширение фрагментов



$c(q)$: уточняем

$n(q)$: уточняем

$V(q)$: $V(p)$

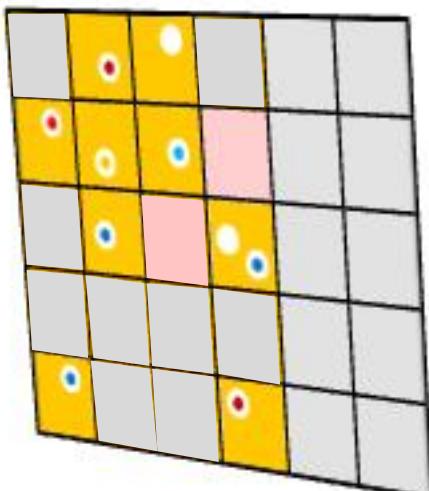


Image 1

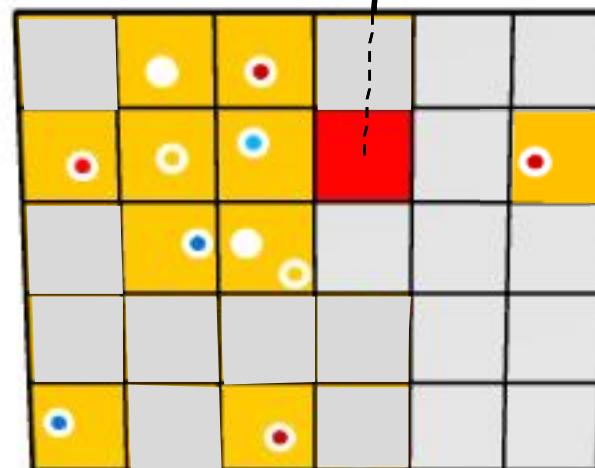


Image 2

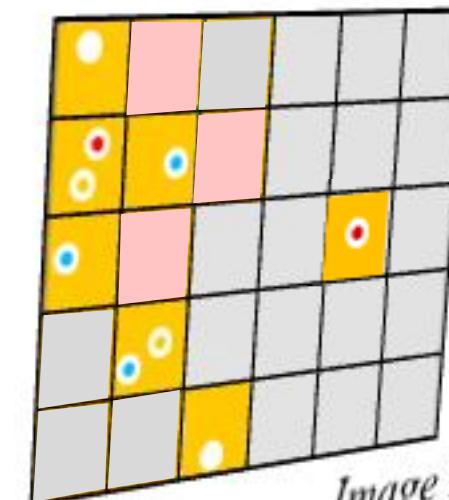
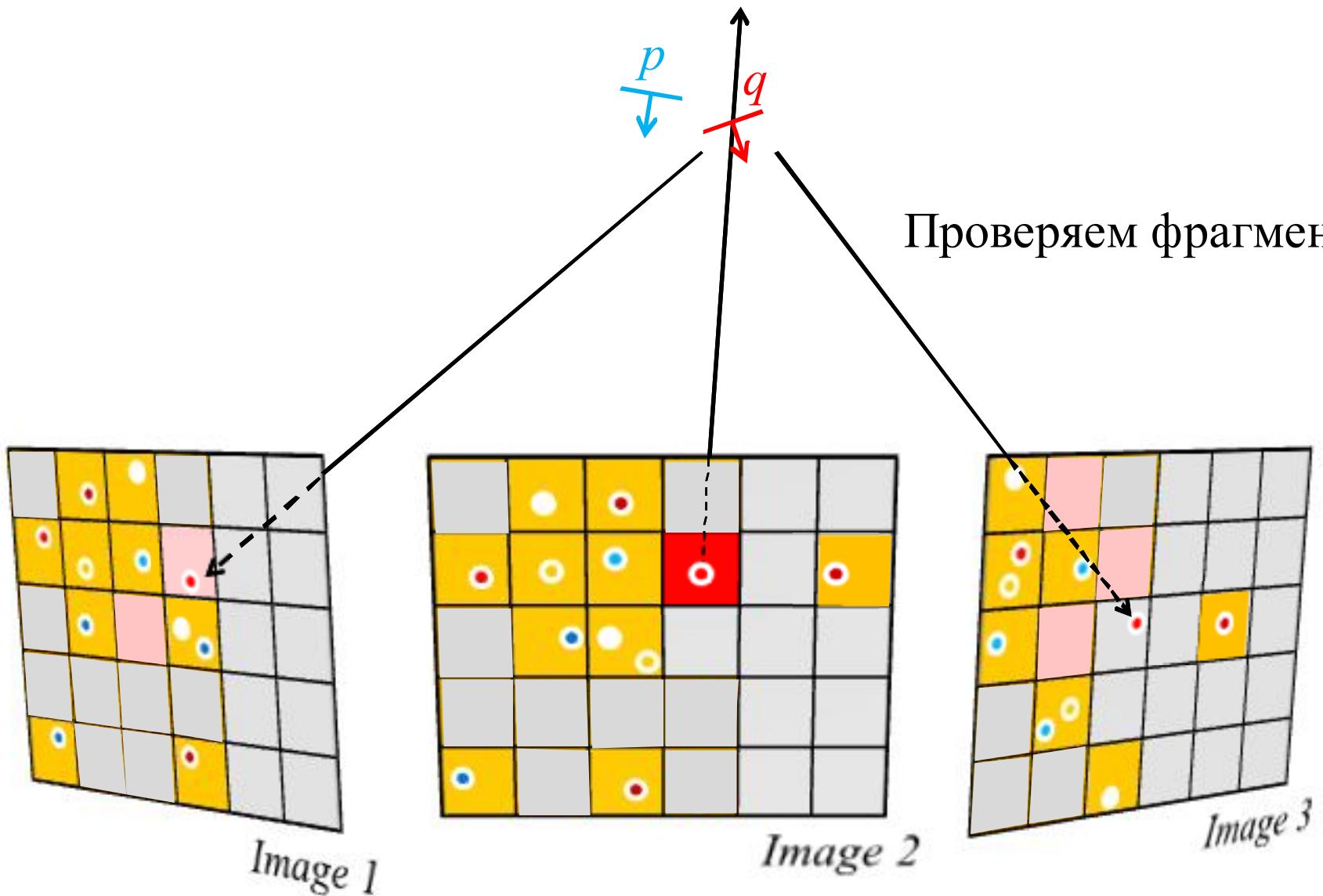
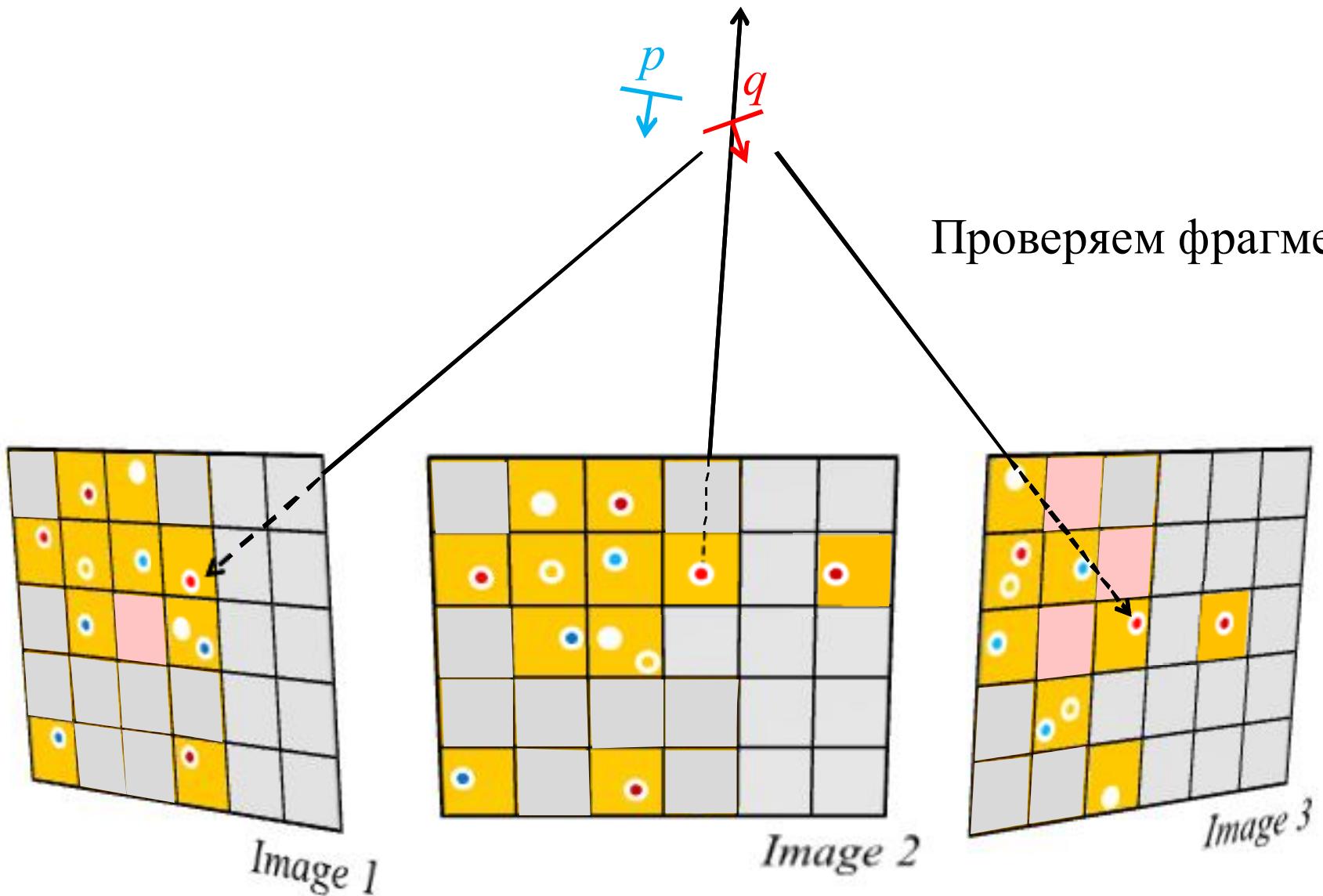


Image 3

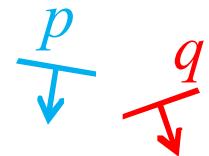
Расширение фрагментов



Расширение фрагментов



Расширение фрагментов



Повторяем

- Для каждого фрагмента
- Для каждого соседнего пустого пикселя

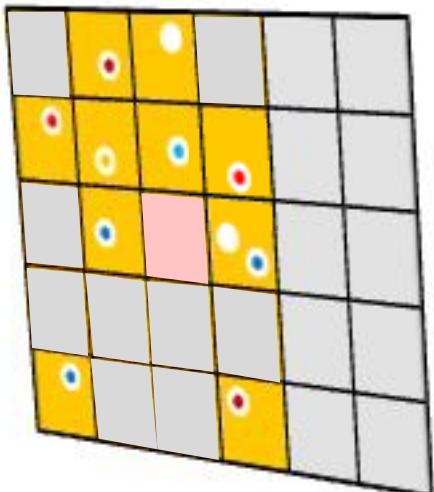


Image 1

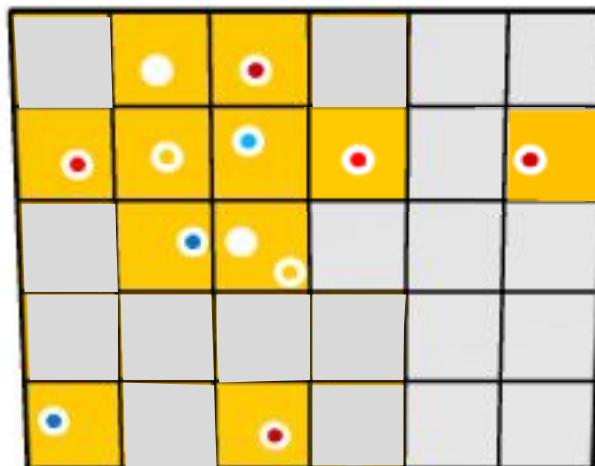


Image 2

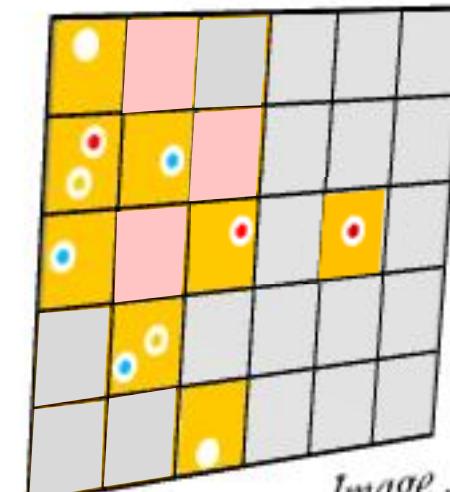


Image 3

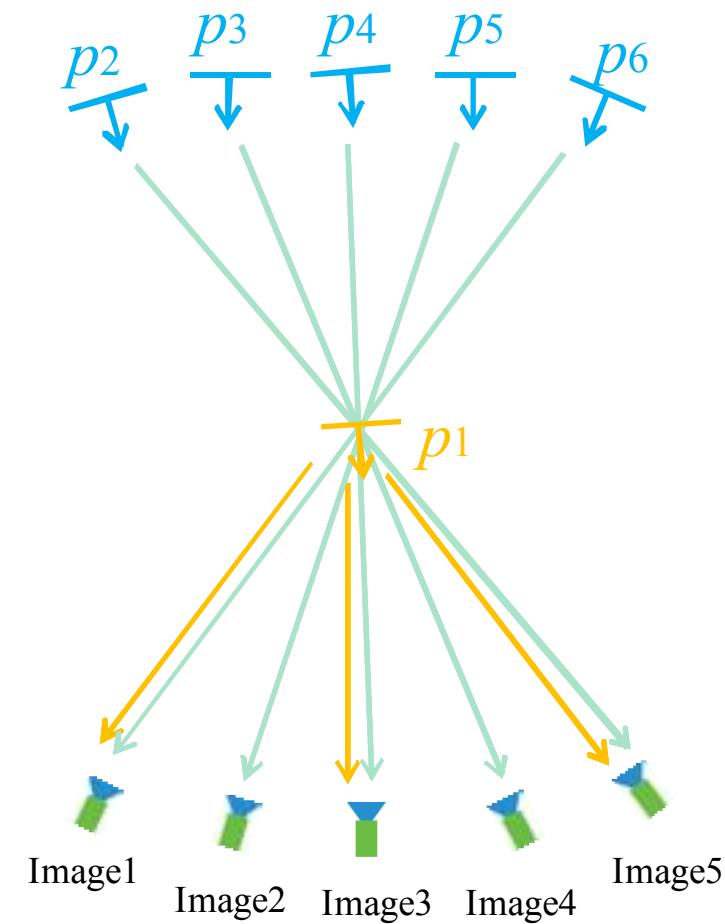
Фильтрация фрагментов



- Согласование видимости

Убираем p_1 если

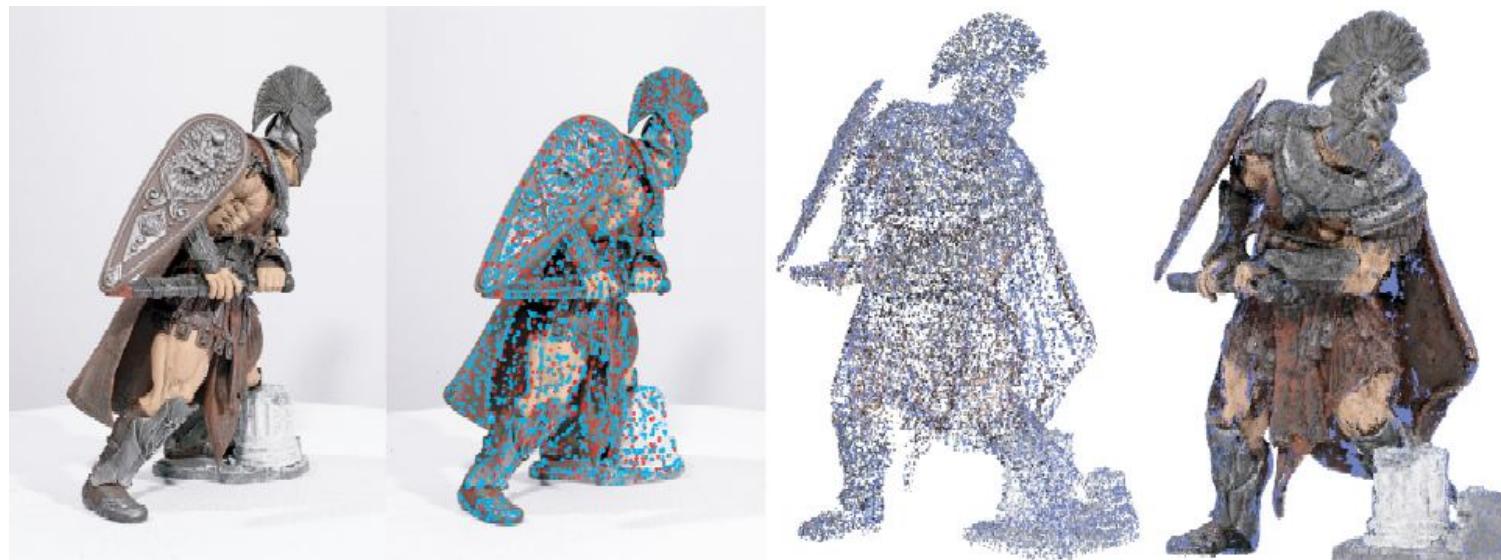
$$|V(p_1)| N(p_1) < \sum_{i=2}^6 N(p_i)$$



Алгоритм на основе фрагментов



- #1. Поиск особых точек
- #2. Инициализация фрагментов
- #3. Расширение фрагментов и фильтрация



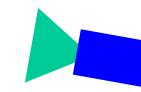
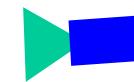
Input image

#1

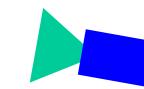
#2

#3

Резюме алгоритма

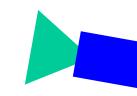
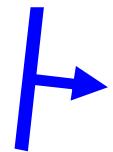
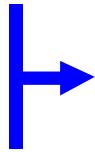


Резюме алгоритма



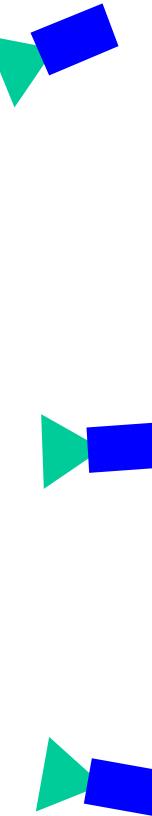
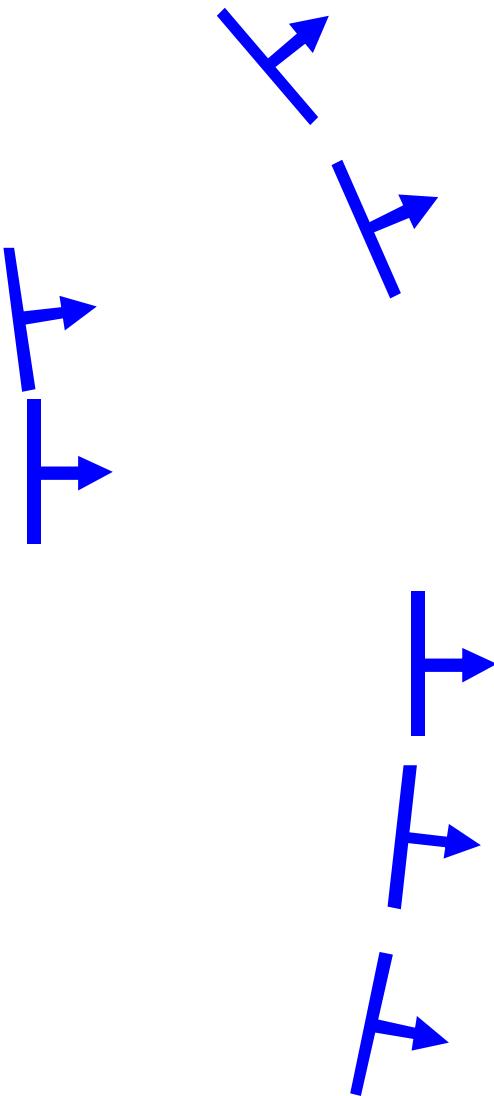
Поиск
особенностей

Резюме алгоритма



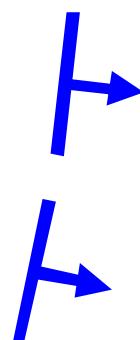
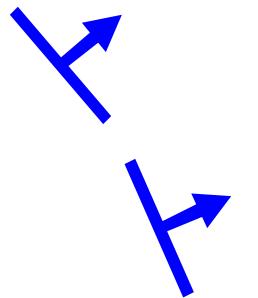
Инициализация
фрагментов

Резюме алгоритма



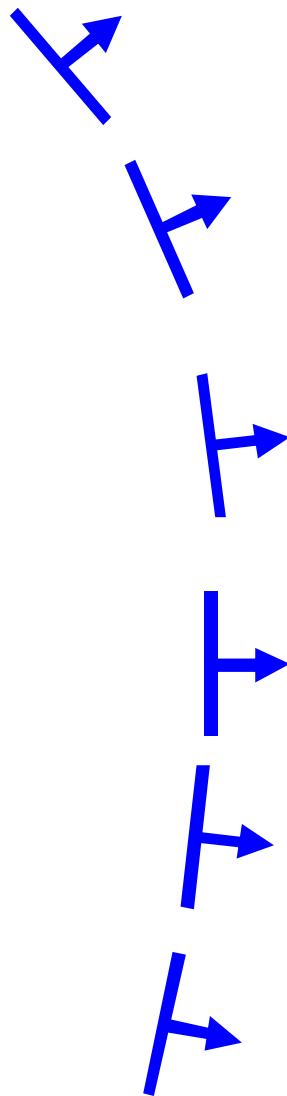
Расширение

Резюме алгоритма



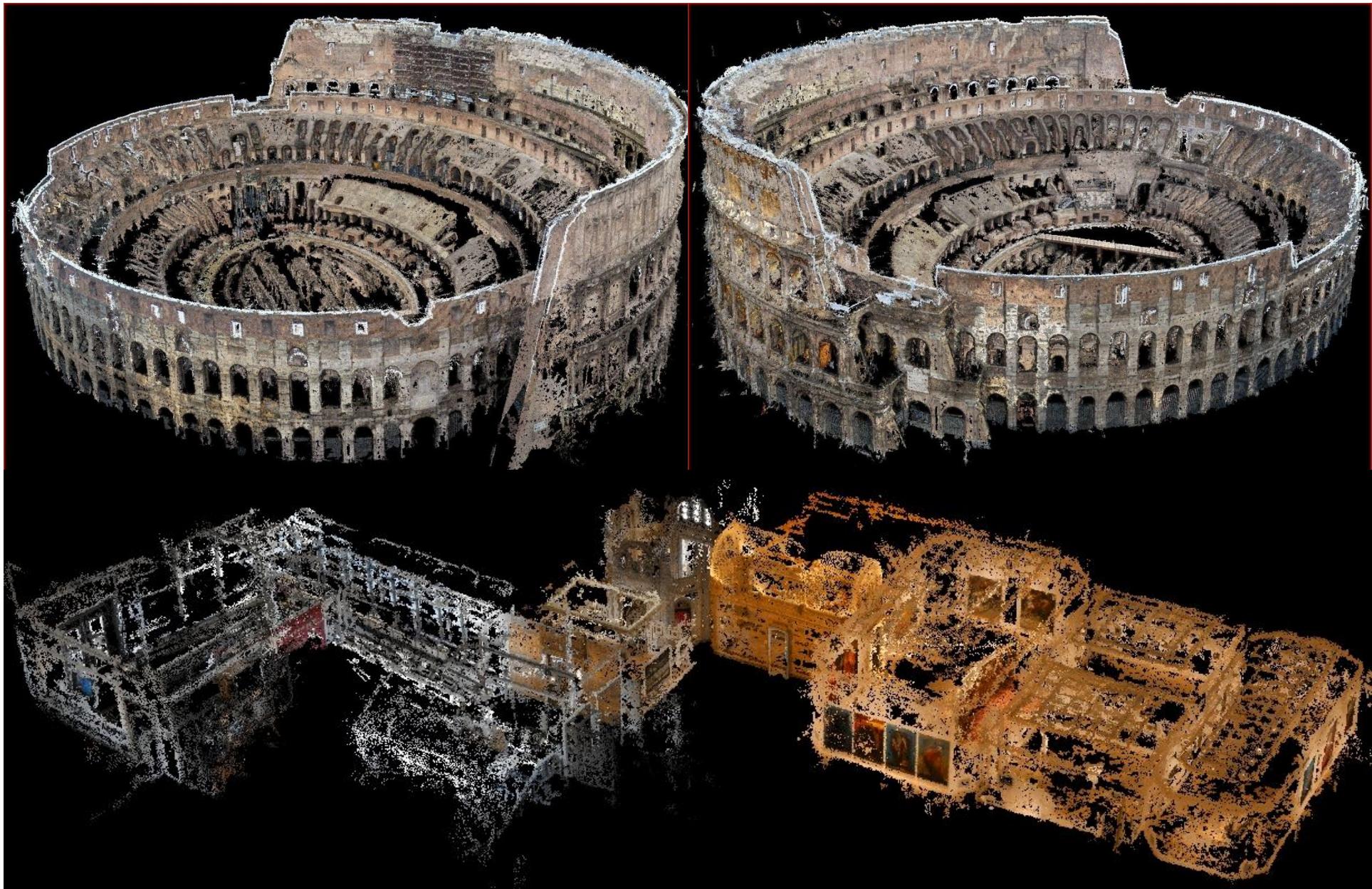
Фильтрация

Резюме алгоритма



Расширение

Примеры работы



Масштабирование



Кластеризация видов и объединение реконструкций

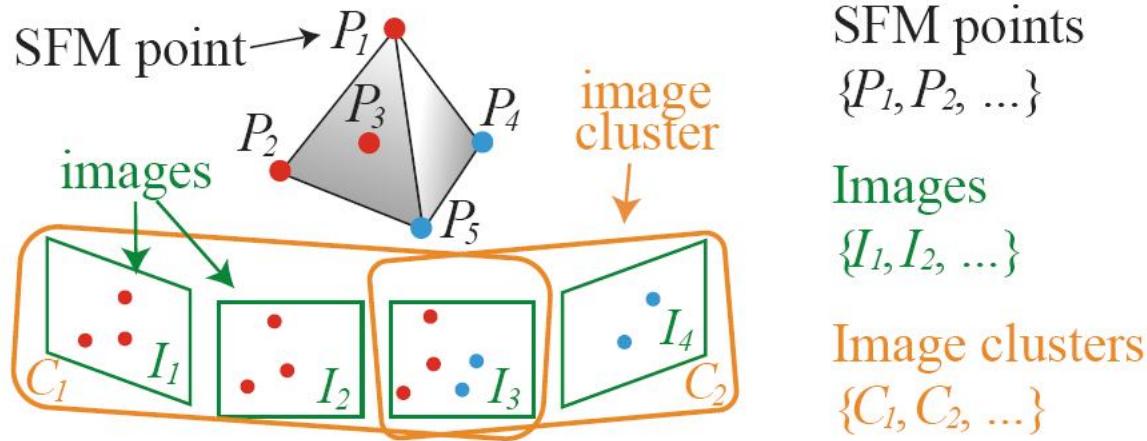


Figure 2. Our view clustering algorithm takes images $\{I_i\}$, SFM points $\{P_j\}$, and their associated visibility information $\{V_j\}$, then produces overlapping image clusters $\{C_k\}$.



Towards Internet-scale Multi-view Stereo

CVPR 2010

Yasutaka Furukawa¹ Brian Curless²

Steven M. Seitz^{1,2} Richard Szeliski³

Google Inc.¹
University of Washington²
Microsoft Research³

Визуальный тест Тьюринга



[The Visual Turing Test for Scene Reconstruction](#). Qi Shan, Riley Adams, Brian Curless, Yasutaka Furukawa, and Steven M. Seitz. 3D Vision 2013, June 2013 (Best Paper Award).



Видео

Учёт контуров



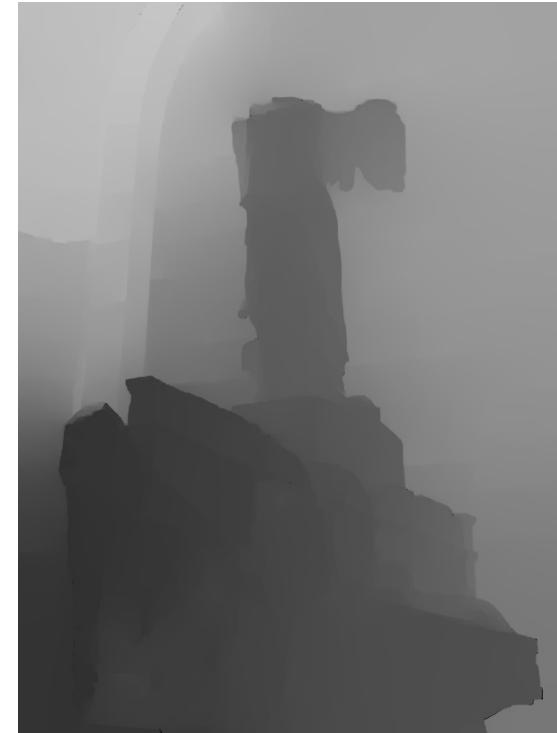
До



После

Qi Shan and Brian Curless and Yasutaka Furukawa and Carlos Hernandez and Steven M. Seitz **Occluding Contours for Multi-View Stereo**. CVPR 2014

Идентификация контуров



Поиск контуров и определение их перекрытия для
идентификации «передних» поверхностей

Дополнение реконструкции



Дополнение реконструкции точками

Учет свободного пространства



PMVS+Poisson



PMVS+Poisson
colored



Proposed approach



with color

План лекции



- Введение
- Многовидовое пассивное стерео
 - Объединение сеток
 - Разрастание регионов
- Приложение: реконструкция головы человека

Моментальное 3Д

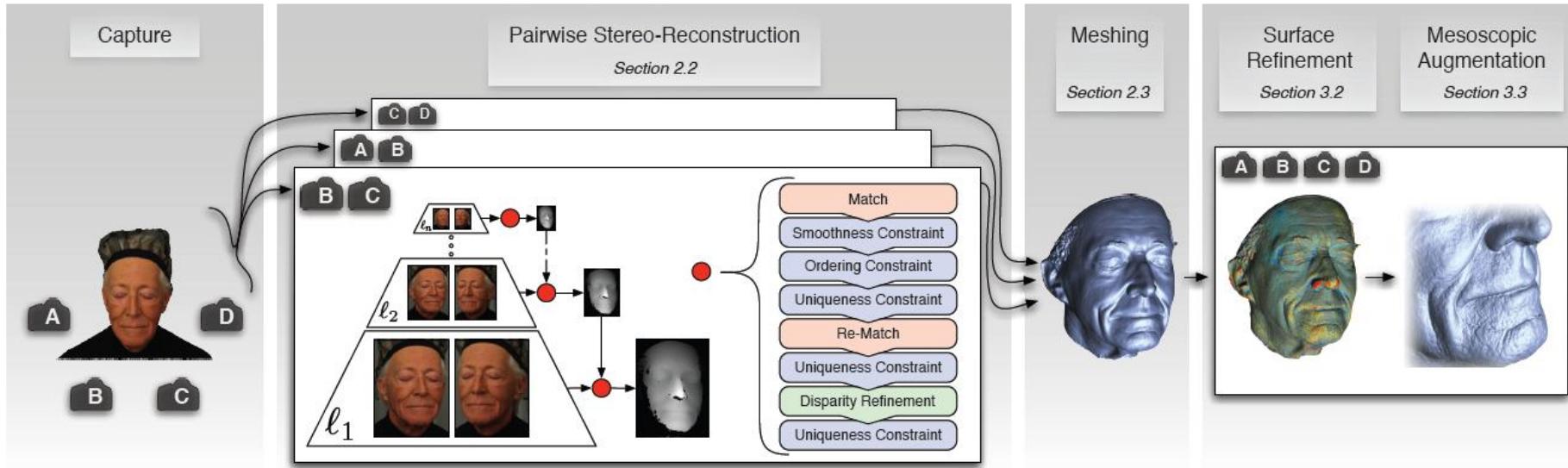


- Построение высокодетализированной 3Д модели лица человека
- От одной стереокамеры до нескольких стереокамер

T. Beeler, B. Bickel, P. Beardsley, R. Sumner, M. Gross [High-Quality Single-Shot Capture of Facial Geometry](#), SIGGRAPH 2010

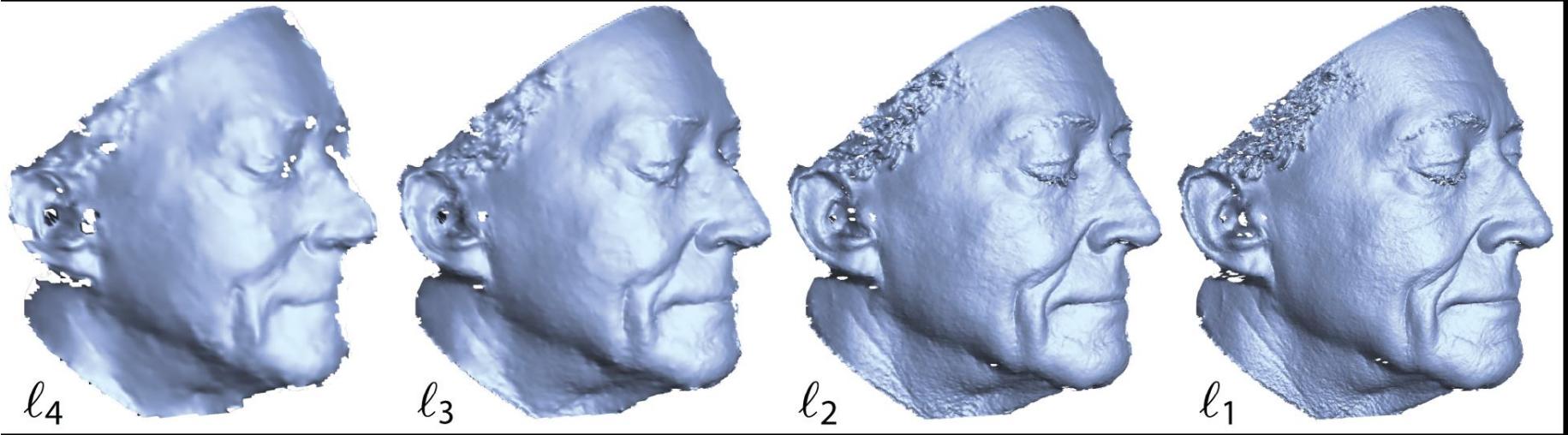


Схема системы



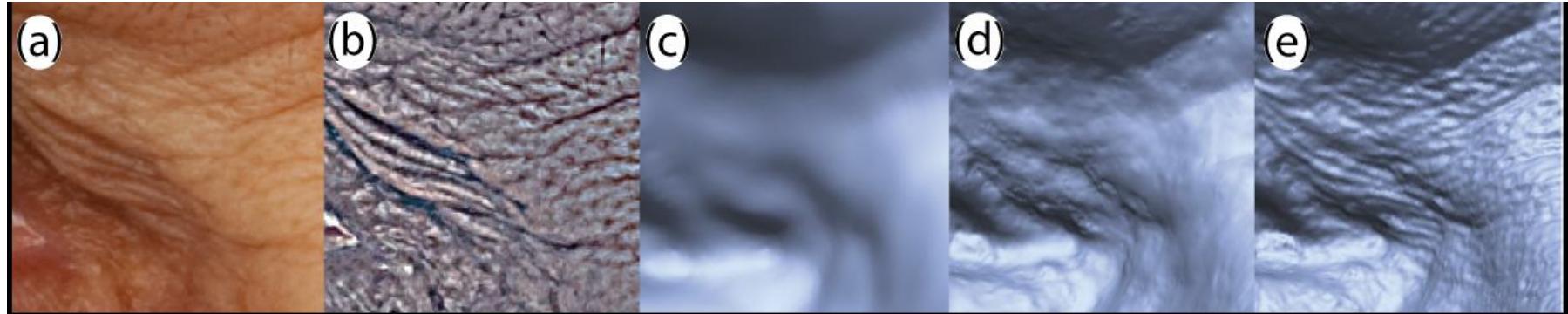
- Иерархическое бинокулярное стерео с итеративным уточнением
 - От разрешения 160*160 до 1280*1280
 - Локальные методы - без использования глобальной оптимизации, сегментации и т.д.

Результат стерео



- Последовательное уточнение геометрии модели
- Субпиксельное уточнение диспаритета

Mesoscopic Augmentation



- Вычисляем высокочастотную составляющую изображения
- Высокие частоты соответствуют мелким деталям на коже человека
- «Искажаем» геометрию согласно специальной формуле

Результат



Results

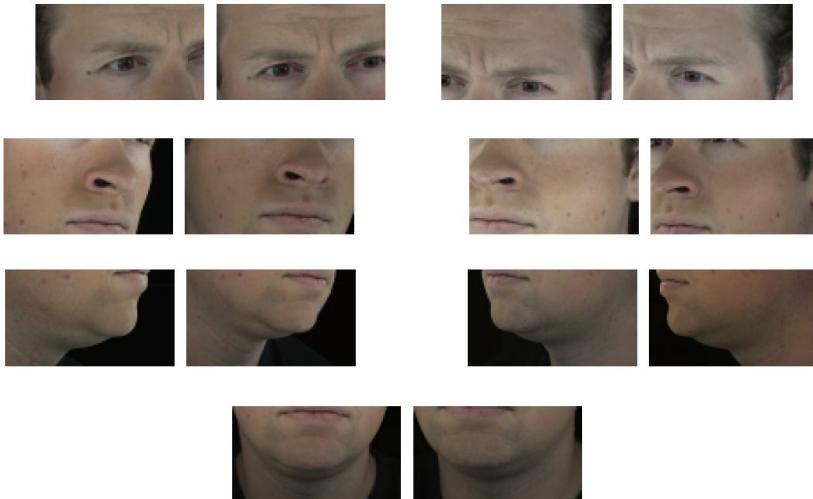
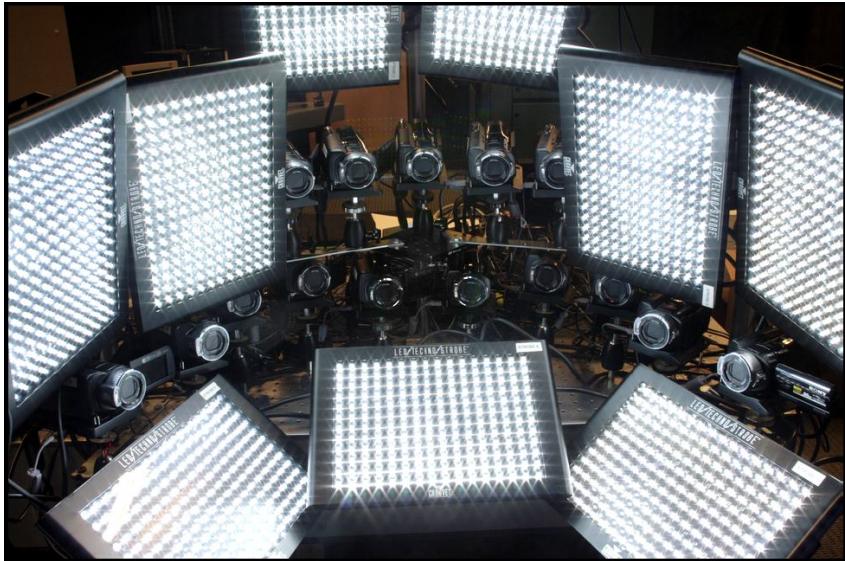
3Д захват лица



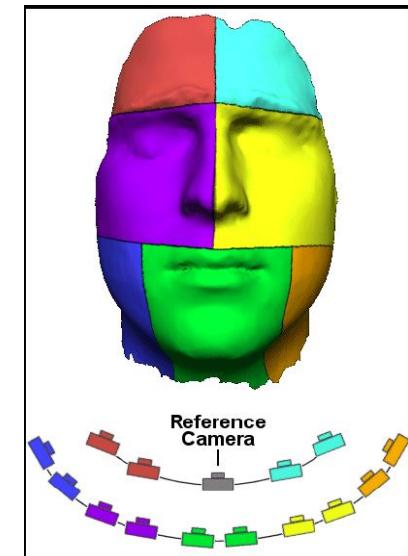
- Пассивная многокамерная система (14 камер)

D. Bradley, W. Heidrich, T. Popa, A. Sheffer [High Resolution Passive Facial Performance Capture](#), SIGGRAPH 2010

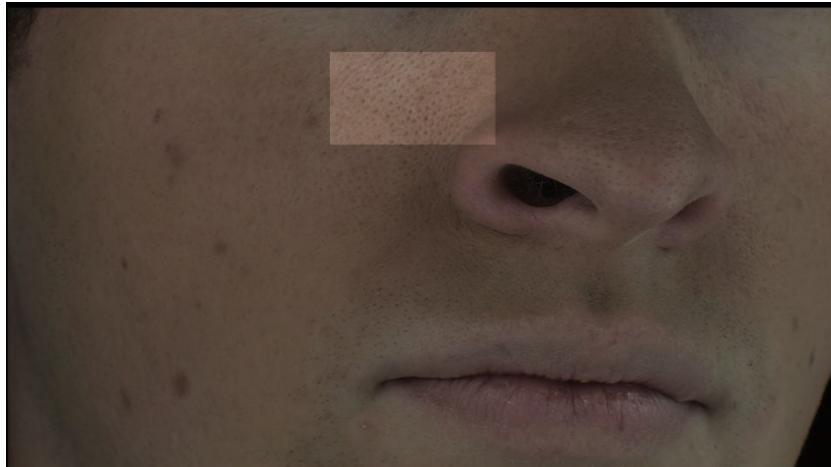
Устройство системы



- 7 стереопар высокого разрешения Sony HDR-SR7
- Система подсветки на основе LED
- Каждая стереопара видит свою зону на лице



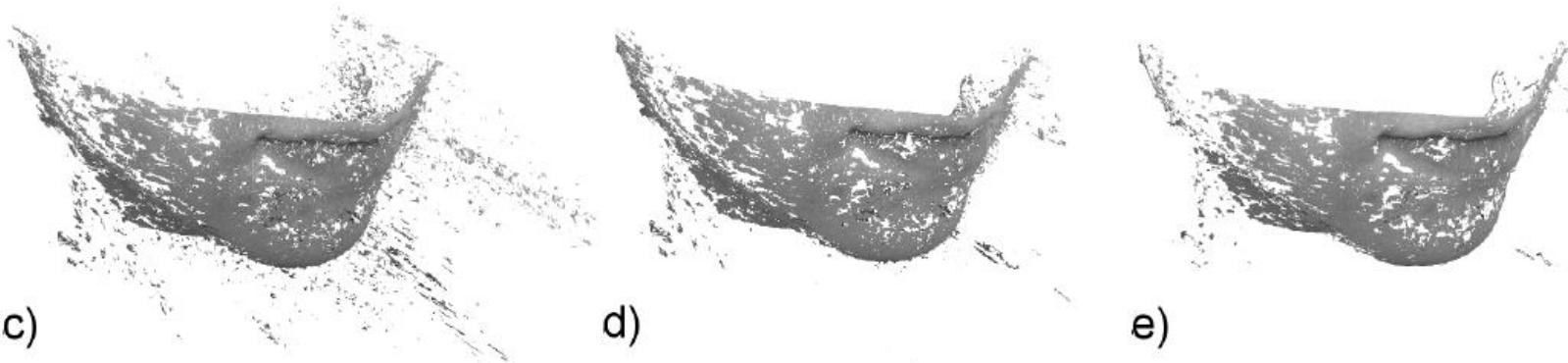
Разрешение изображений



- При таком высоком разрешении и равномерной подсветке поры на коже дают достаточно текстуры для вычисления пассивного стерео

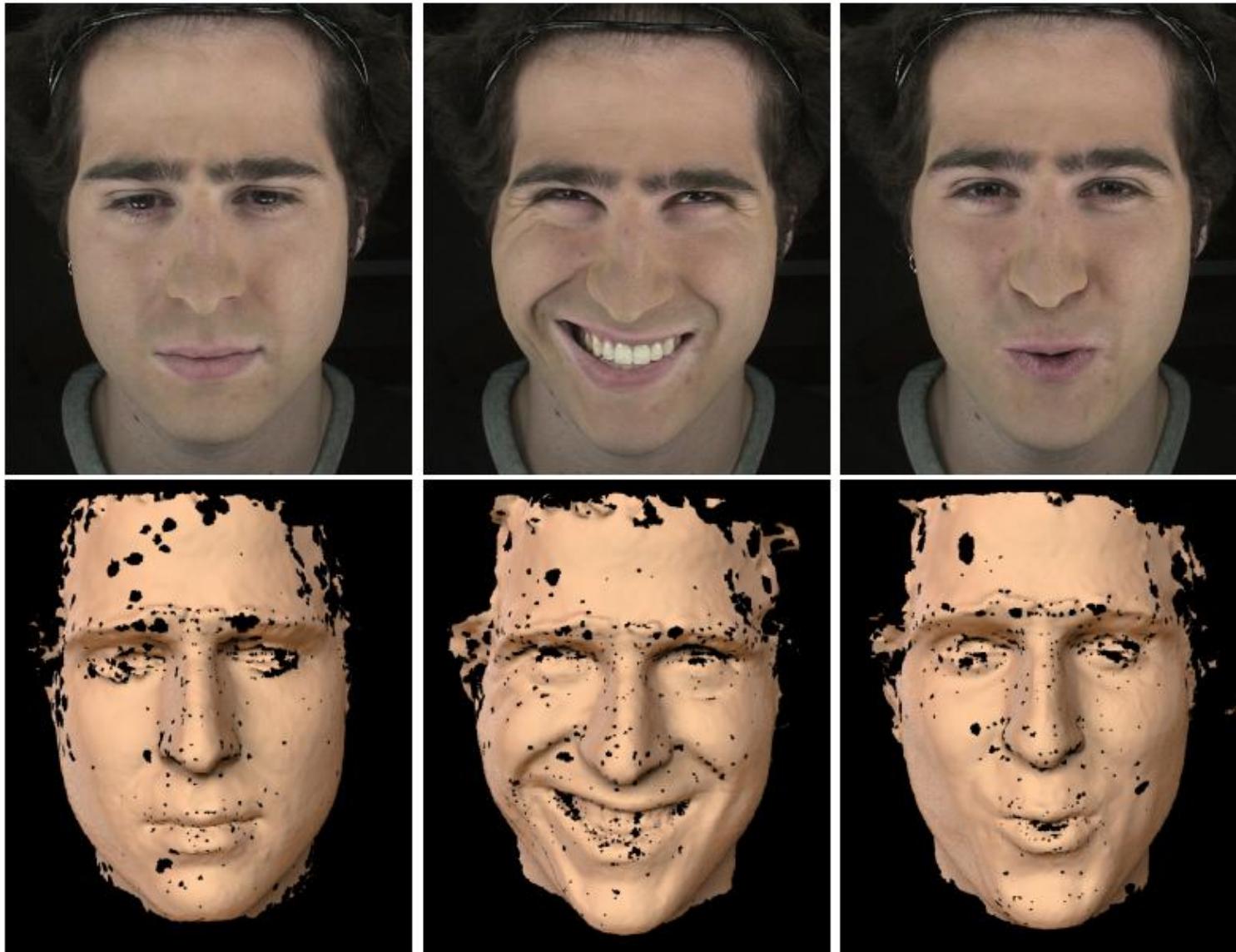
BRADLEY, D., BOUBEKEUR, T., AND HEIDRICH, W. Accurate multi-view reconstruction using robust binocular stereo and surface meshing. CVPR. 2008.

Пассивное стерео и обработка



- Каждая стереопара даёт карту глубины и облако точек (шумное)
- Итеративное сглаживание и вычисление порогов для отбрасывания шумов

Результат по 1 кадру



Отслеживание

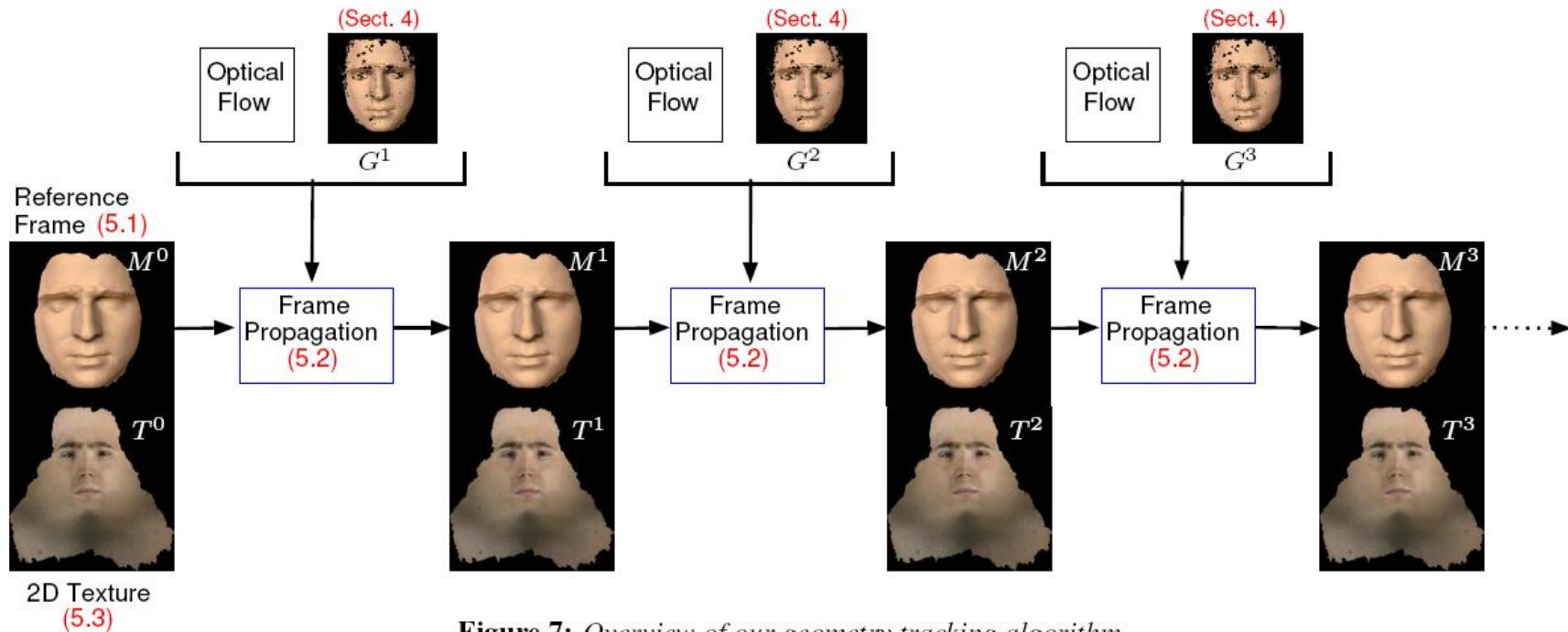
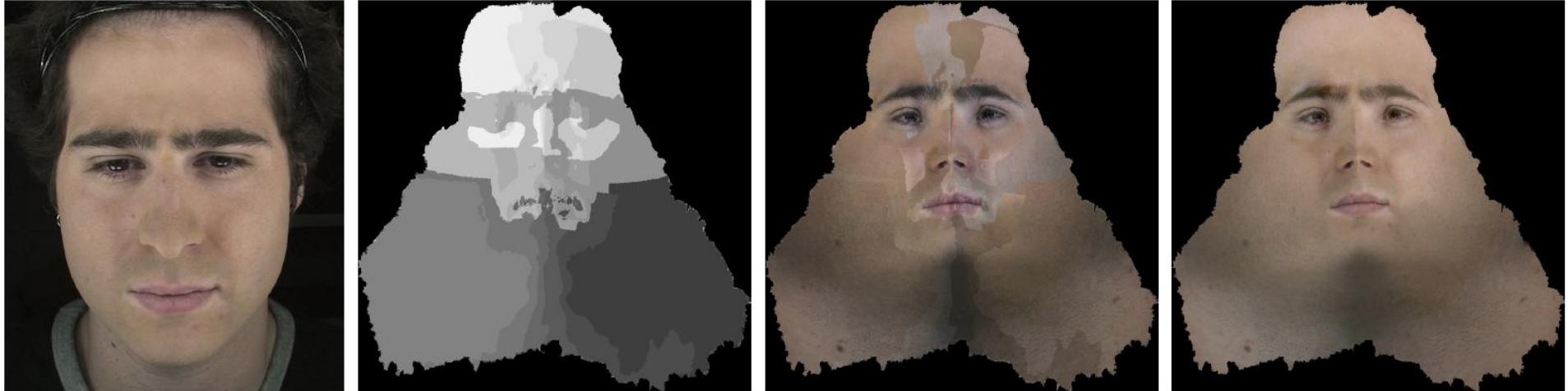


Figure 7: Overview of our geometry tracking algorithm.

- Первое изображение выбирается как опорная модель
- Нужно построить такую деформацию опорной модели, чтобы она совпадала с моделями по следующим кадрам
- Для направления используется оптический поток

Синтез текстуры



- Участки текстуры копируются со своих камер
- Поскольку камеры без цветовой калибровки, при прямом копировании текстуры видны резкие цветовые переходы (артефакты)
- Для обработки используется композиция по Пуассону

Результат



Results

Применение на практике



L.A. Noire (2011)

Резюме



- Методы на основе пассивного стерео по качеству начали приближаться к методам активного стерео и тоже начинают находить практическое применение
- Все методы многовидового пассивного стерео опираются на принцип фотосогласования
- Два основных современных подхода:
 - Объединение сеток
 - Распространение фрагментов