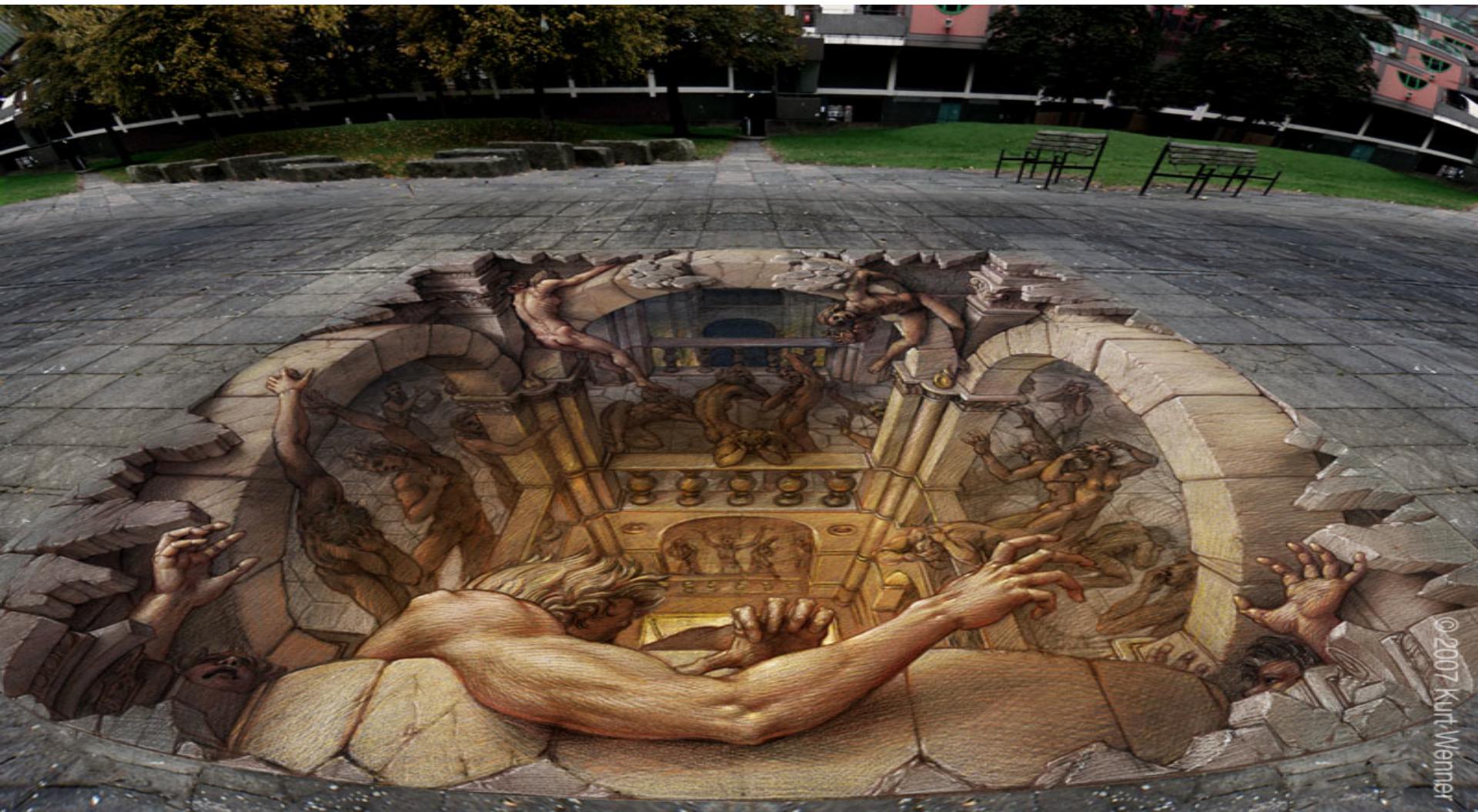


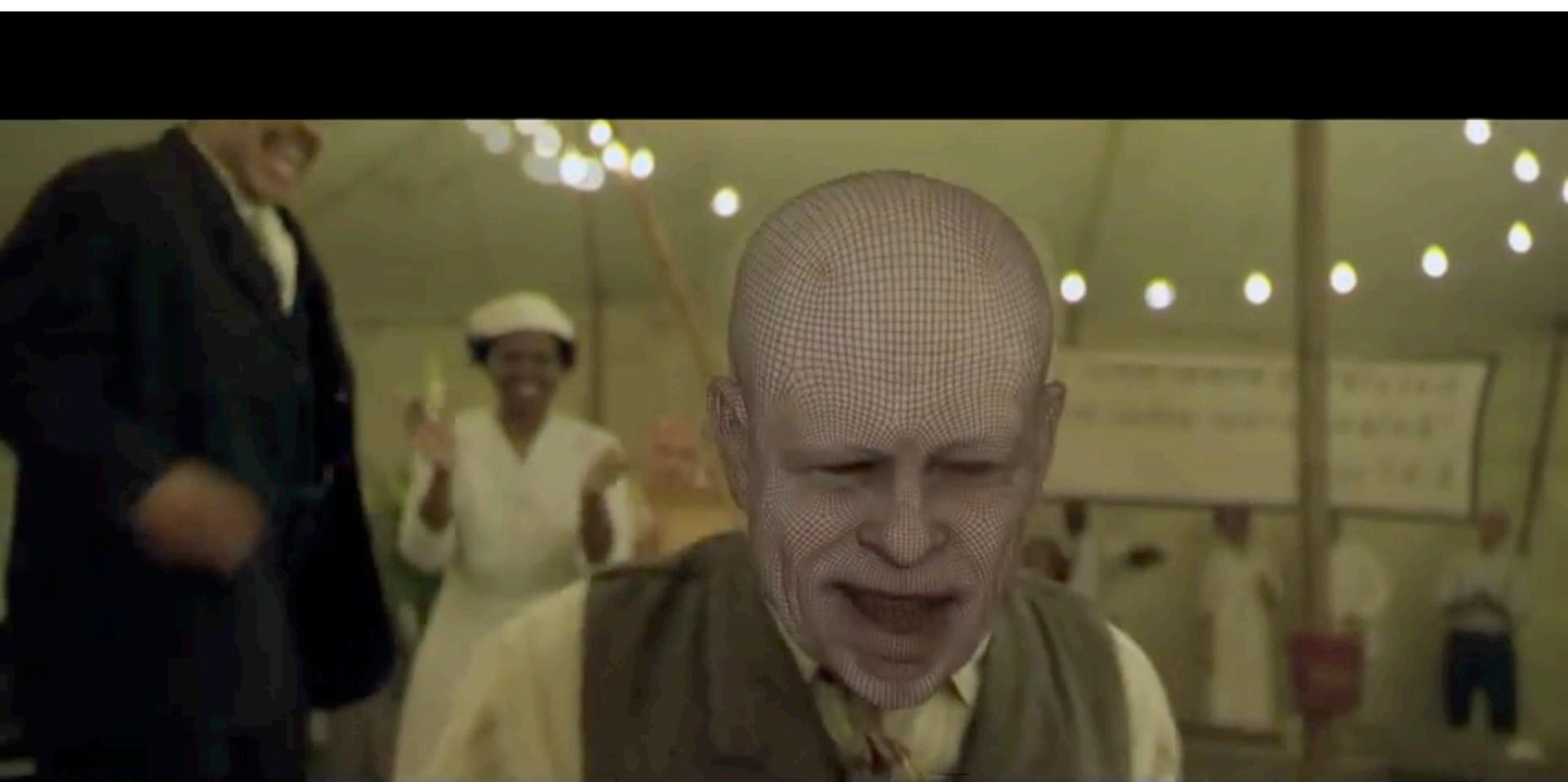
Проективная геометрия в компьютерном зрении

Ерухимов Виктор



Зачем нужна геометрия в
компьютерном зрении?



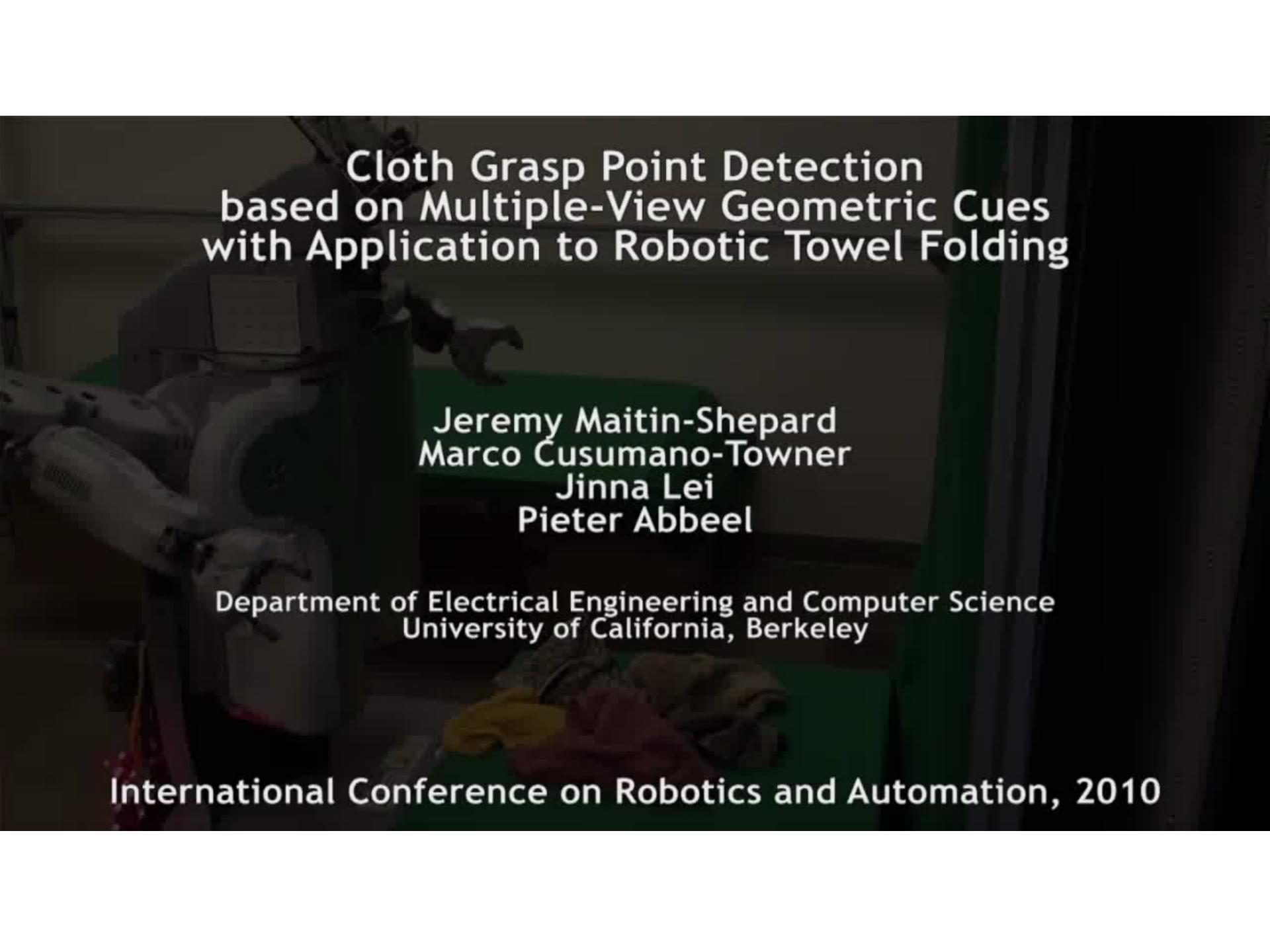


 Scan page to see more



of coordinated covers
sofa. Cover: 50% viscose/
cotton. Sigmund grey/check.
Sofa: W180cm, D84cm, H74cm.
Sessel: W100cm, H80cm, D82cm.
Each piece of furniture is unique.

Sofa £80 Hand made, gives soft
feels. Each piece of furniture is unique.
Sessel: W100cm, H80cm, D82cm. 202.016.82

A dark, low-light photograph showing a robotic arm's grippers holding a piece of light-colored, possibly yellow or white, fabric. The background is dark and out of focus.

Cloth Grasp Point Detection based on Multiple-View Geometric Cues with Application to Robotic Towel Folding

Jeremy Maitin-Shepard
Marco Cusumano-Towner
Jinna Lei
Pieter Abbeel

Department of Electrical Engineering and Computer Science
University of California, Berkeley

International Conference on Robotics and Automation, 2010

Gogle™

Copyrighted Material

SECOND EDITION

Multiple View Geometry in computer vision



Richard Hartley and Andrew Zisserman

Copyrighted Material

CAMBRIDGE

ДЖ. ГОЛУБ, Ч. ВАН ЛОУН

МАТРИЧНЫЕ ВЫЧИСЛЕНИЯ



ИЗДАТЕЛЬСТВО «МИР»

ozon.ru

Singular Value Decomposition

$$A = UDV^T$$

$$\begin{aligned} A\mathbf{v}_i &= \sigma_i \mathbf{u}_i \\ A^T \mathbf{u}_i &= \sigma_i \mathbf{v}_i \end{aligned} \quad D \text{ уникальна}$$

Сложность $4m^2n + 8mn^2 + 9n^3$

Использование SVD

- $\min_{\|x\|=1} \|AX\|^2$
- Ядро матрицы определяется последними правыми сингулярными векторами
- ранг – ненулевыми сингулярными числами

Использование SVD

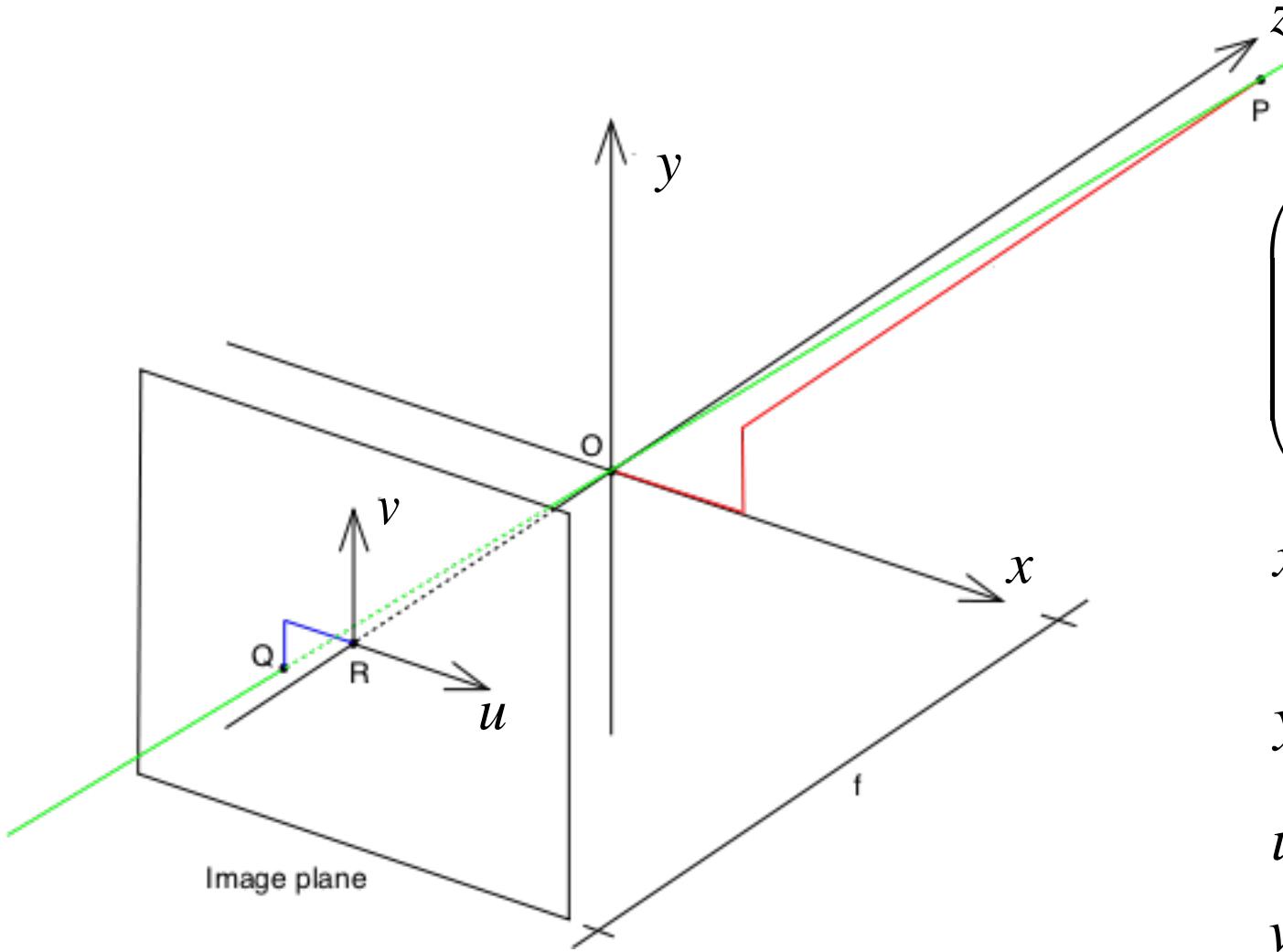
- Ортогональная проблема Прокруста

$$\underset{B^T B = I}{\arg \min} \|A - B\|$$

- Теорема Эккарта-Янга

$$\underset{\text{rank}(B) = r < \text{rank}(A)}{\arg \min} \|A - B\|$$

Pinhole camera model



$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = R \begin{pmatrix} X_0 \\ Y_0 \\ Z_0 \end{pmatrix} + t$$

$$x' = \frac{X}{Z}$$

$$y' = \frac{Y}{Z}$$

$$u = f_x x' + c_x$$

$$v = f_y y' + c_y$$

Проективное пространство

- Состоит из множества прямых линий, проходящих через нулевую точку линейного пространства
- Включает в себя линию на бесконечности
- Проективное пространство, соответствующее трехмерному линейному пространству, называют проективной плоскостью \mathbb{P}^2

Однородные координаты

Точка на проективной плоскости $p = \begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} \in P^2$

ставится в соответствие всем точкам из трехмерного пространства, проектирующимся в нее:

$$u = \begin{pmatrix} x \\ y \\ z \end{pmatrix} = z \begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} \in R^3$$

(x,y,z) называют однородными координатами точки p .

Однородные координаты

Уравнение прямой:

$$p^T l = 0 \quad p_x l_x + p_y l_y + p_z l_z = 0$$

Прямая, проходящая через две точки:

$$l = p_1 \times p_2 \quad p_i^T p_1 \times p_2 = 0$$

Пересечение прямых:

$$p = l_1 \times l_2 \quad l_i^T l_1 \times l_2 = 0$$

Где пересекаются прямые линии

$$l_1 = \begin{pmatrix} a \\ b \\ c \end{pmatrix}, l_2 = \begin{pmatrix} a \\ b \\ \tilde{c} \end{pmatrix}$$

Точка на бесконечности

Прямые линии

$$l_1 = \begin{pmatrix} a \\ b \\ c \end{pmatrix}, l_2 = \begin{pmatrix} a \\ b \\ \tilde{c} \end{pmatrix}$$

пересекаются в точке

$$p = l_1 \times l_2 = \begin{pmatrix} \hat{a} \\ \hat{b} \\ 0 \end{pmatrix}$$

Прямая на бесконечности

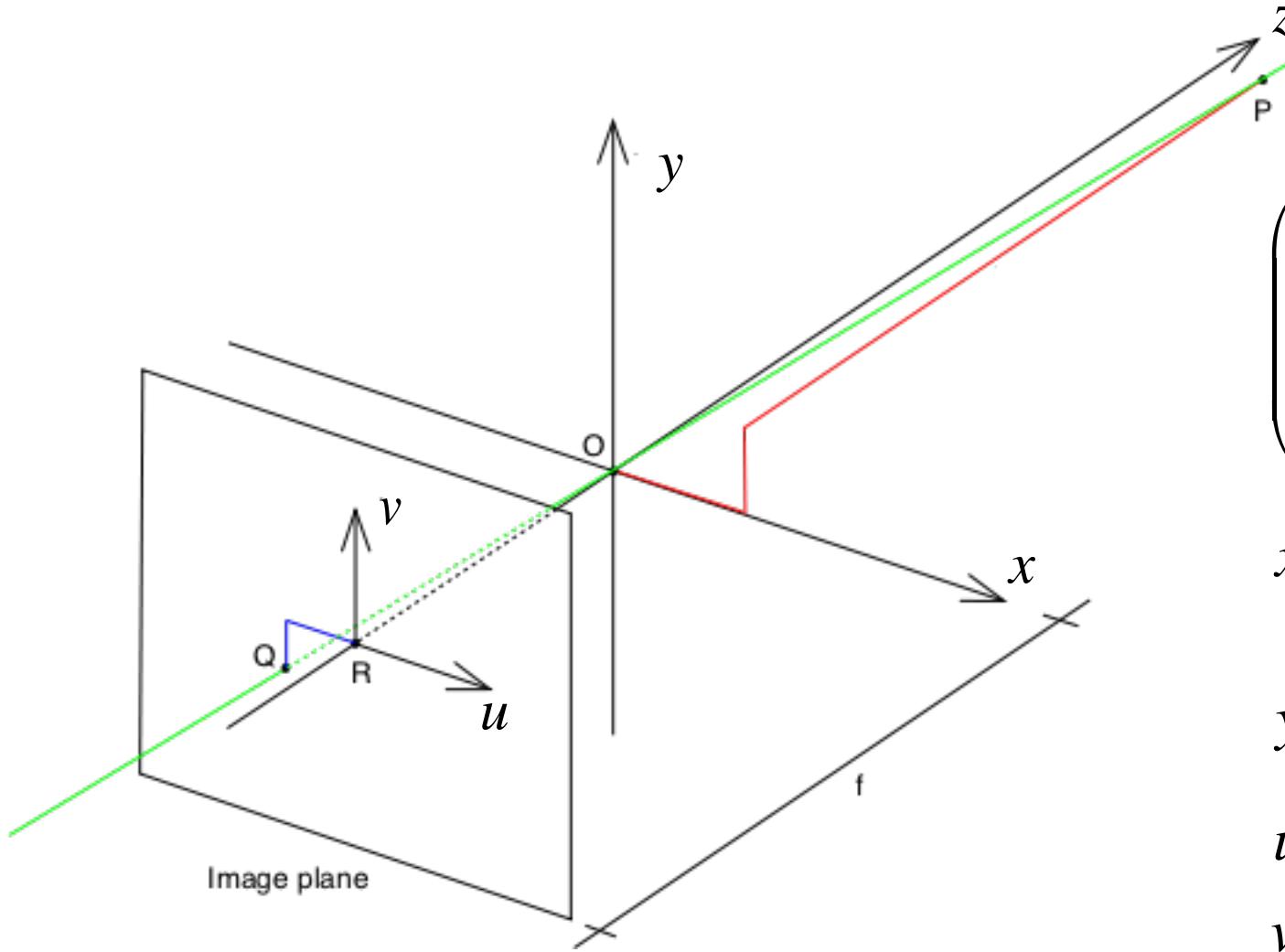
Все точки на бесконечности лежат на одной прямой:

$$l_{\infty} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

$$\begin{pmatrix} x \\ y \\ 0 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = 0$$

Group	Matrix	Distortion
Projective 8 dof	$\begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix}$	
Affine 6 dof	$\begin{bmatrix} a_{11} & a_{12} & t_x \\ a_{21} & a_{22} & t_y \\ 0 & 0 & 1 \end{bmatrix}$	
Similarity 4 dof	$\begin{bmatrix} sr_{11} & sr_{12} & t_x \\ sr_{21} & sr_{22} & t_y \\ 0 & 0 & 1 \end{bmatrix}$	
Euclidean 3 dof	$\begin{bmatrix} r_{11} & r_{12} & t_x \\ r_{21} & r_{22} & t_y \\ 0 & 0 & 1 \end{bmatrix}$	

Pinhole camera model



$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = R \begin{pmatrix} X_0 \\ Y_0 \\ Z_0 \end{pmatrix} + t$$

$$x' = \frac{X}{Z}$$

$$y' = \frac{Y}{Z}$$

$$u = f_x x' + c_x$$

$$v = f_y y' + c_y$$

Проективное преобразование (гомография)

- Обратимое отображение $H^l: P^2 \rightarrow P^2$, такое, что любые три точки x_1, x_2, x_3 лежат на одной прямой тогда и только тогда, когда на одной прямой лежат их образы $H^l(x_1), H^l(x_2), H^l(x_3)$, называется гомографией.

Теорема о представлении гомографии

- Отображение $H^l: P^2 \rightarrow P^2$ является гомографией тогда и только тогда, когда существует обратимая матрица 3×3 H , такая, что для каждой точки из P^2 , представляемой $3 \times$ -мерным вектором x , выполняется соотношение $H^l(x) = Hx$
- Доказать

Соотношение гомографии для координат пикселей

$$w \begin{pmatrix} \tilde{u} \\ \tilde{v} \\ 1 \end{pmatrix} = H \begin{pmatrix} u \\ v \\ 1 \end{pmatrix}$$

$$\begin{aligned}\tilde{u} &= \frac{h_{11}u + h_{12}v + h_{13}}{h_{31}u + h_{32}v + h_{33}} \\ \tilde{v} &= \frac{h_{21}u + h_{22}v + h_{23}}{h_{31}u + h_{32}v + h_{33}}\end{aligned}$$

Свойства гомографии

$$\tilde{p} = H^I(p) \quad p = \begin{pmatrix} u \\ v \end{pmatrix} \in I$$

- Суперпозиция двух гомографий -- гомография

$$p_1 = H_1^I(p), \quad p_2 = H_2^I(p_1) \xrightarrow{\text{blue arrow}} p_2 = (H_1 \cdot H_2)^I(p)$$

- Гомографии образуют группу в \mathbb{P}^2

Матрица проекции камеры

$$w \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = P \begin{pmatrix} X_0 \\ Y_0 \\ Z_0 \\ 1 \end{pmatrix}$$

$$P = K[R \mid T]$$
$$K = \begin{pmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{pmatrix}$$

Матрица проекции камеры

$$q = \begin{pmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{pmatrix} [R \mid T] Q$$

$K \quad P$

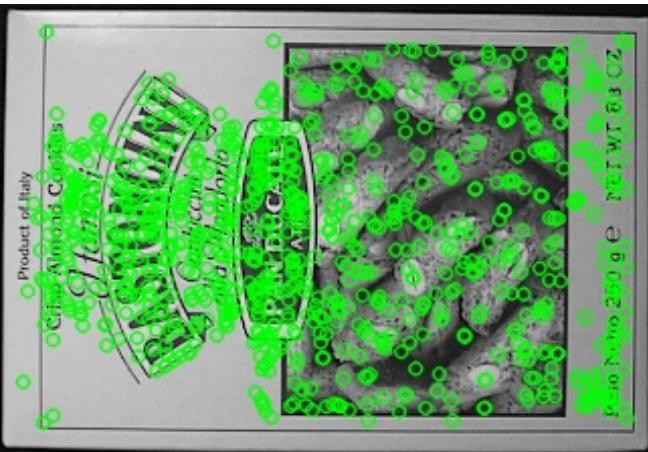
$q \in P^2$
 $Q \in P^3$

Отображение гомографии

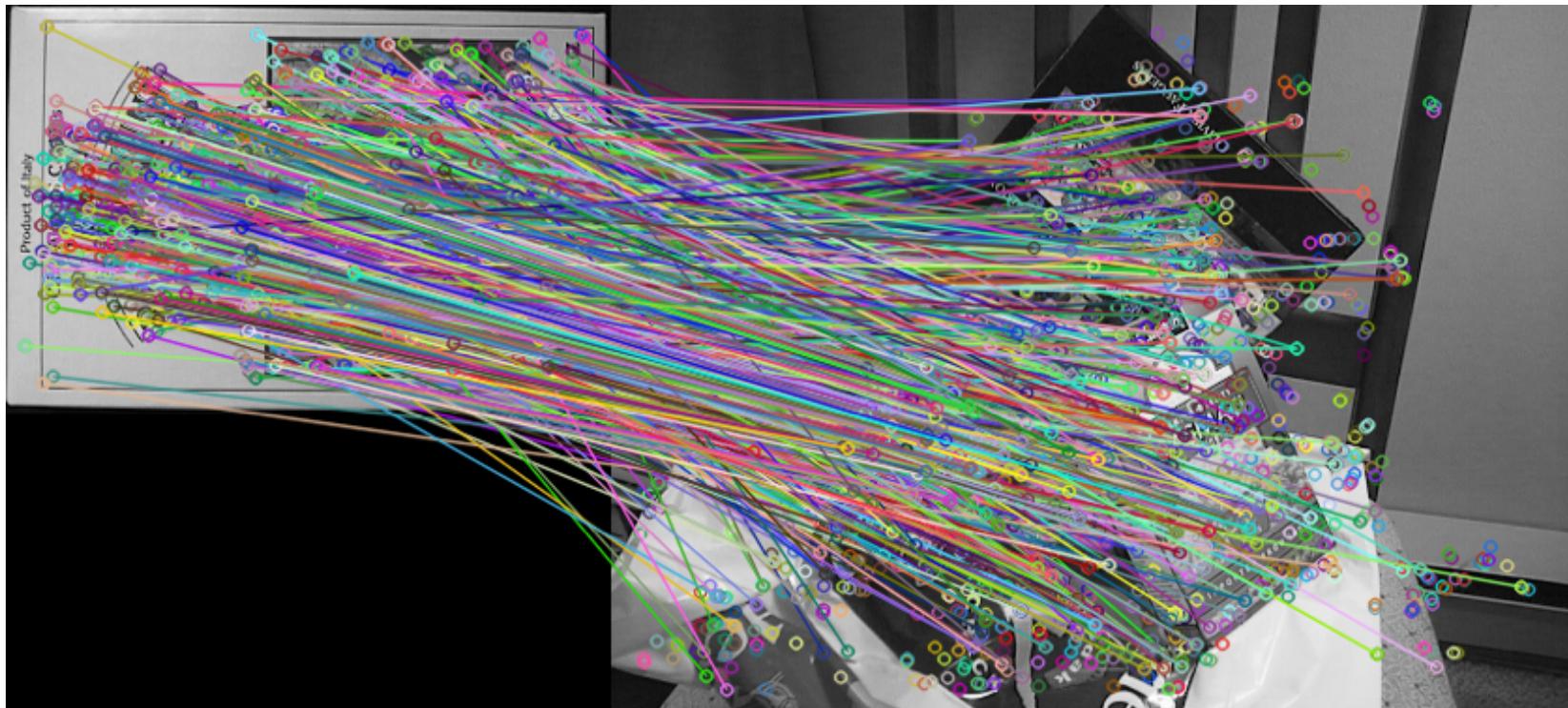
$$S_2 = RS_1 + T \quad n_1^T S_1 = d_1 \quad S_1, S_2 \in \Re^3$$

$$\begin{aligned} & S_2 = RS_1 + T \frac{n_1^T S_1}{d} \\ & S_2 = \left(R + \frac{Tn_1^T}{d} \right) S_1 \quad \xrightarrow{\text{blue arrow}} \quad q_2 = \left(R + \frac{Tn_1^T}{d} \right) q_1 \end{aligned}$$

Keypoints example



Matching descriptors example

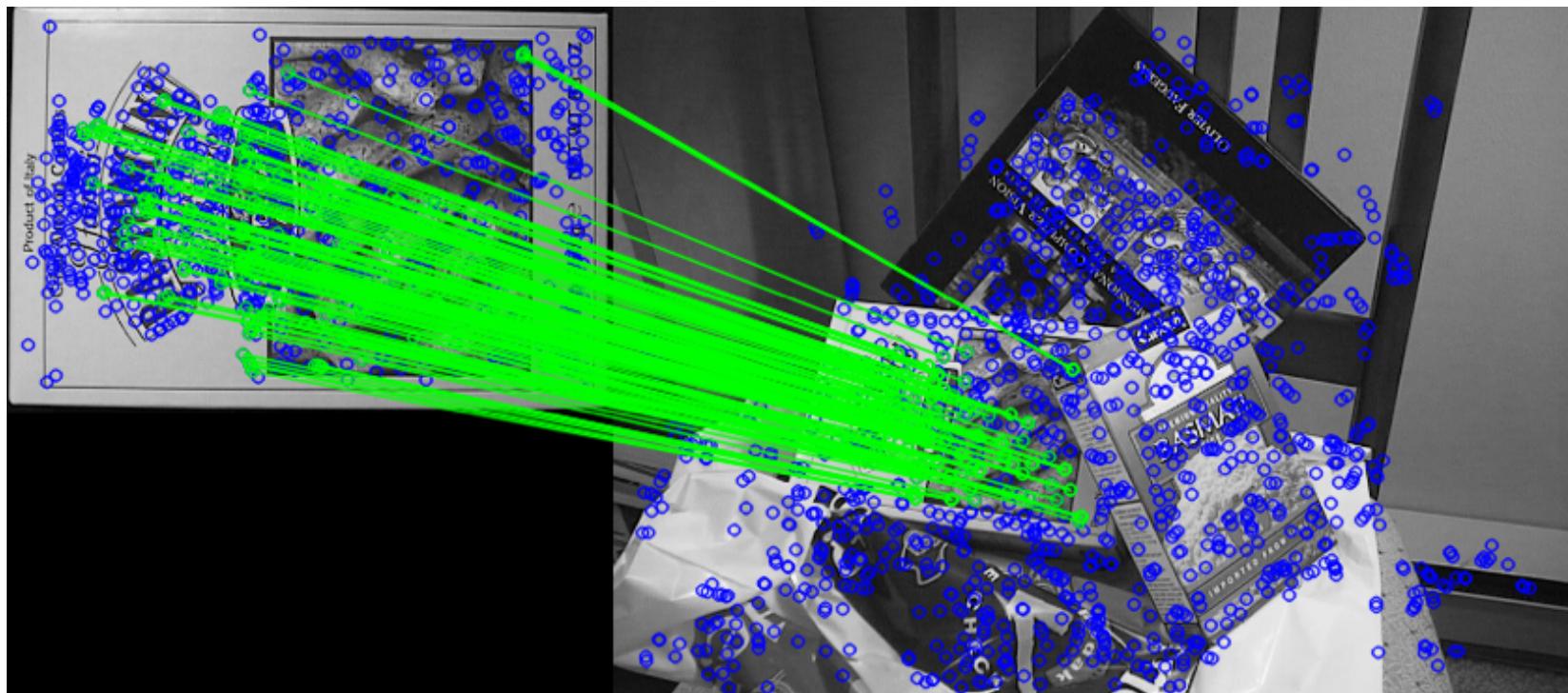


Random Sample Consensus

- Do n iterations until #inliers > inlierThreshold
 - Draw k matches randomly
 - Find the transformation
 - Calculate inliers count
 - Remember the best solution

The number of iterations required $\sim 10 * \left(\frac{\# \text{ matches}}{\# \text{ inliers}} \right)^k$

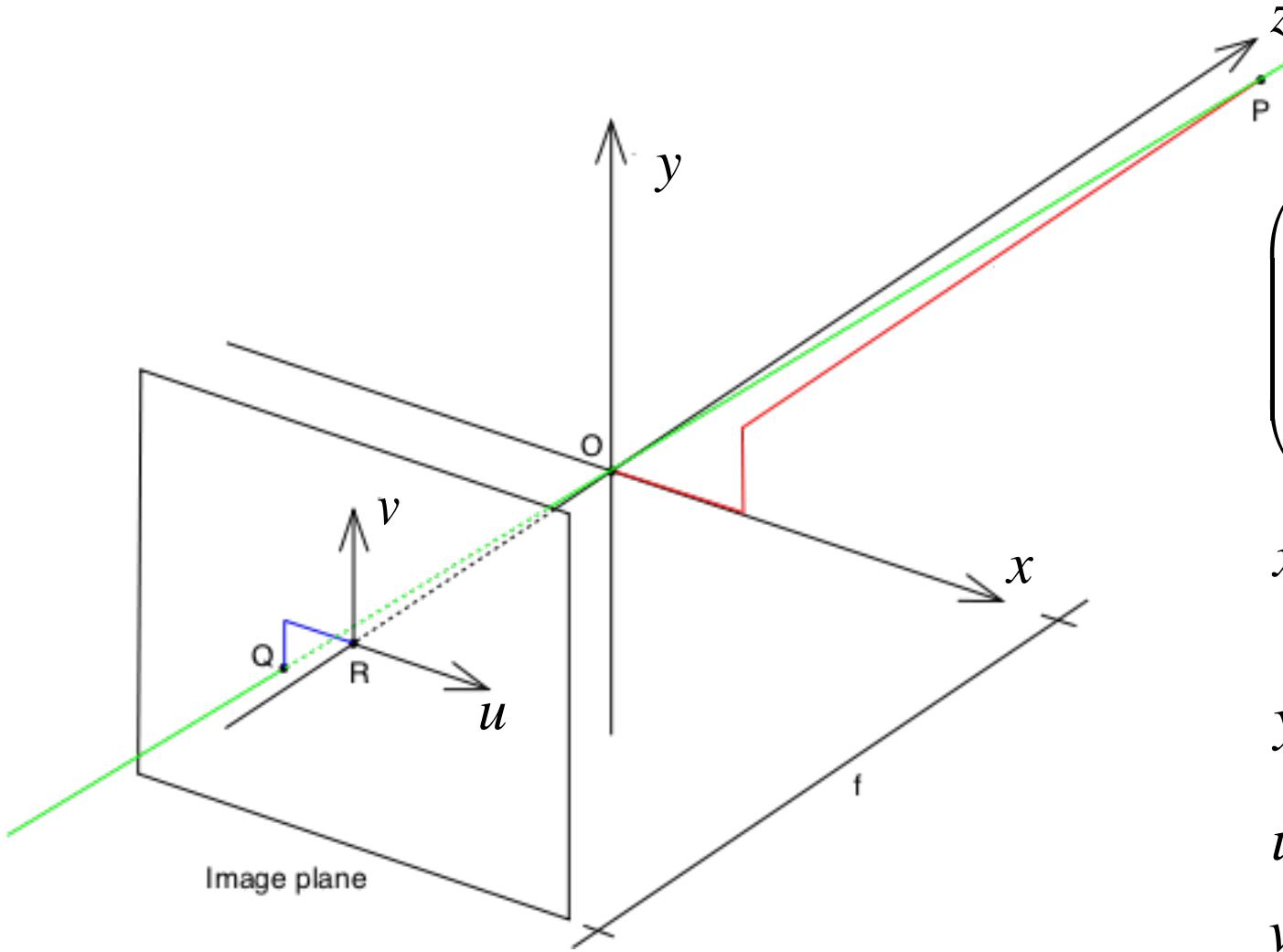
Geometry validation



Когда между изображениями существует гомография?

- Плоская сцена
- Бесконечно удаленная сцена

Pinhole camera model



$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = R \begin{pmatrix} X_0 \\ Y_0 \\ Z_0 \end{pmatrix} + t$$

$$x' = \frac{X}{Z}$$

$$y' = \frac{Y}{Z}$$

$$u = f_x x' + c_x$$

$$v = f_y y' + c_y$$

Модель искажений линзы

- Искажения, вносимые линзой, можно компенсировать, не используя предположения о трехмерной сцене

Pose estimation

Detected
image points:

$$\left\{ \begin{pmatrix} u_i \\ v_i \end{pmatrix} \right\}_{i=1..n}$$

Train object
points:

$$\left\{ \begin{pmatrix} X_i \\ Y_i \\ Z_i \end{pmatrix} \right\}_{i=1..n}$$

A class of object
transformations:

$$f \left[\begin{pmatrix} X \\ Y \\ Z \end{pmatrix}, R, T \right] = R \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} + T$$

Reprojection error

$$w_i \begin{pmatrix} u_i^p \\ v_i^p \\ 1 \end{pmatrix} = P \begin{pmatrix} X_i \\ Y_i \\ Z_i \\ 1 \end{pmatrix}$$

$$error(P) = \sum_i \left[\begin{pmatrix} u_i \\ v_i \end{pmatrix} - \begin{pmatrix} u_i^p \\ v_i^p \end{pmatrix} \right]^2$$

Perspective-n-Points problem

$$\min_{R, T} \text{error}(K, R, T)$$

Direct Linear Transformation

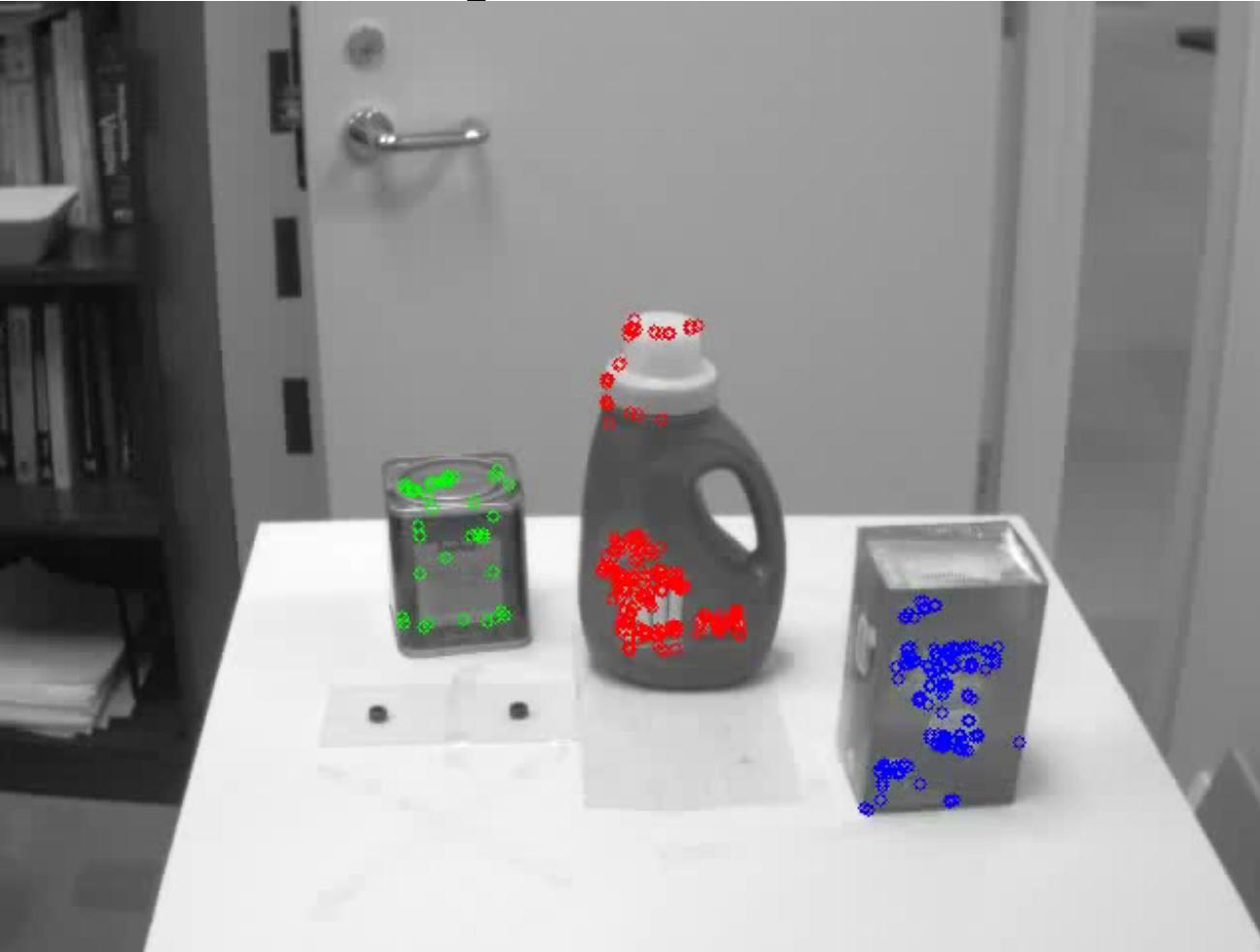
$$\begin{pmatrix} u_i^p w_i \\ v_i^p w_i \\ w_i \end{pmatrix} = P \begin{pmatrix} X_i \\ Y_i \\ Z_i \\ 1 \end{pmatrix} \quad \xrightarrow{\text{blue arrow}} \quad \begin{pmatrix} u_i^p \\ v_i^p \\ 1 \end{pmatrix} \times P \begin{pmatrix} X_i \\ Y_i \\ Z_i \\ 1 \end{pmatrix} = 0$$

$$[R \mid T] = K^{-1}P$$

Other methods

- $O(N)$ closed-form methods
- Levenberg-Marquardt
- P3P, P4P
- RANSAC

Object detection example

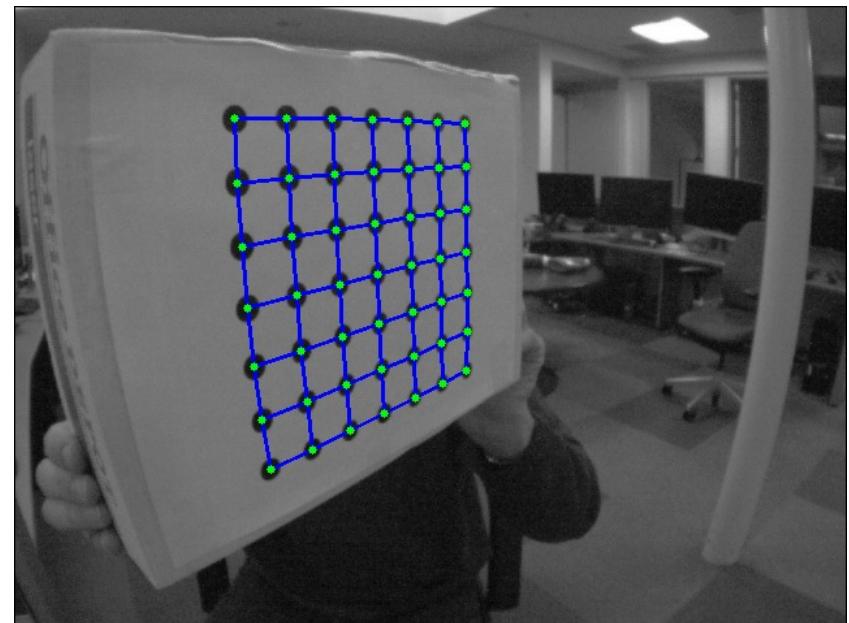


Iryna Gordon and David G. Lowe,
"What and where: 3D object
recognition with accurate pose," in
*Toward Category-Level Object
Recognition*, eds. J. Ponce, M.
Hebert, C. Schmid, and A. Zisserman,
(Springer-Verlag, 2006), pp. 67-82.

Manuel Martinez Torres, Alvaro
Collet Romea, and Siddhartha
Srinivasa, MOPED: A Scalable and
Low Latency Object Recognition
and Pose Estimation
System, Proceedings of ICRA 2010,
May, 2010.

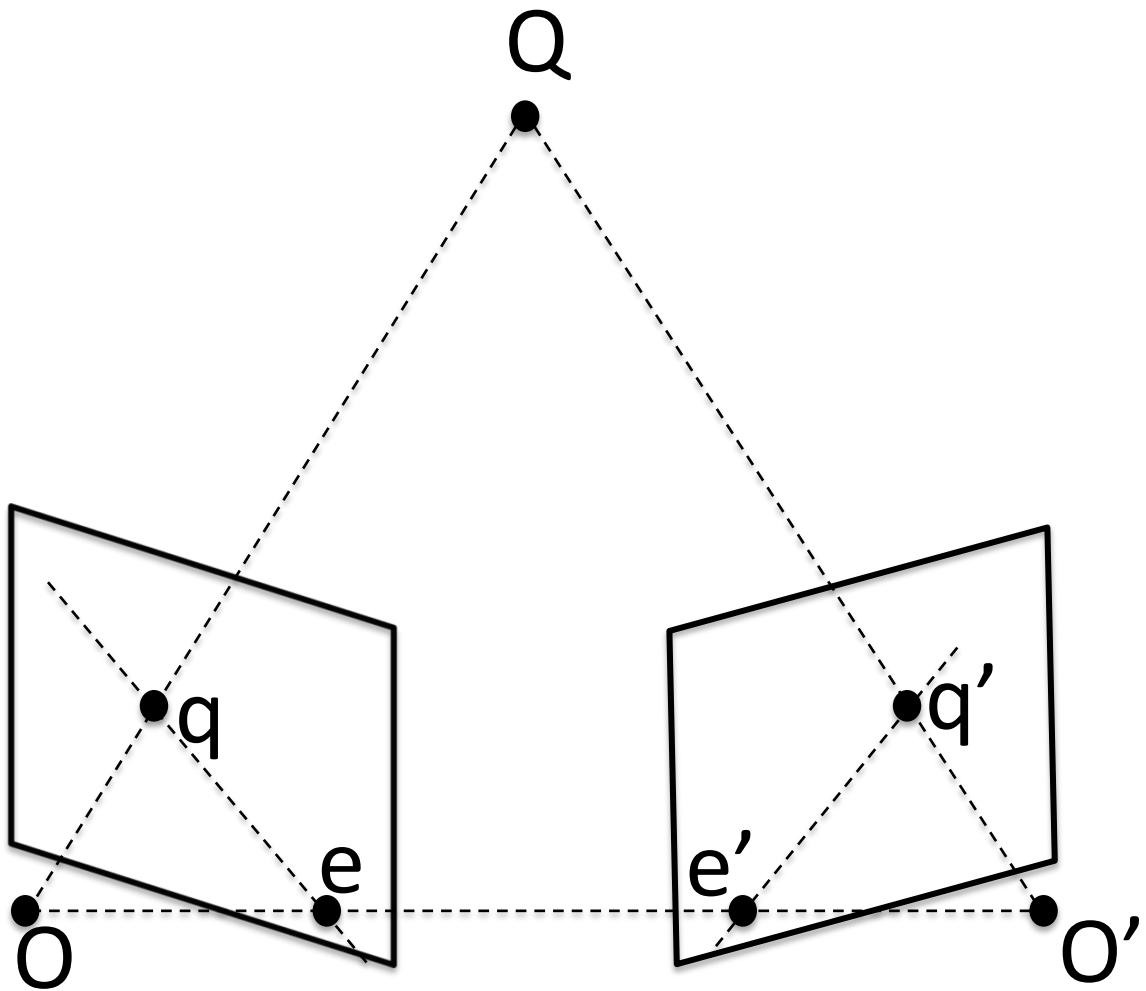
Camera calibration

- Estimate camera parameters given images of a known template



- QR разложение для обратимой матрицы
единственно, если диагональные элементы
 R выбраны положительными.

Stereo: epipolar geometry



Fundamental matrix
constraint

$$(u', v', 1) \cdot F \cdot \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = 0$$

Эпиполярные линии

- Линия в плоскости изображения одной из двух камер, проходящая через проекцию трехмерной точки и проекцию центра другой камеры, называется эпиполярной

$$l' = Fq$$

$$l = q'^T F$$

- Проекция центра одной камеры на другую называется эпиполем

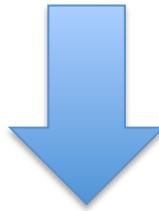
Свойства

- Если F – фундаментальная матрица для пары камер (P, P') , то F^T – фундаментальная матрица для пары камеры (P', P)
- Эпиполярные линии
- Эпиполь
- F имеет 7 степеней свободы

Определение трехмерных координат

$$P = K[I|0]$$

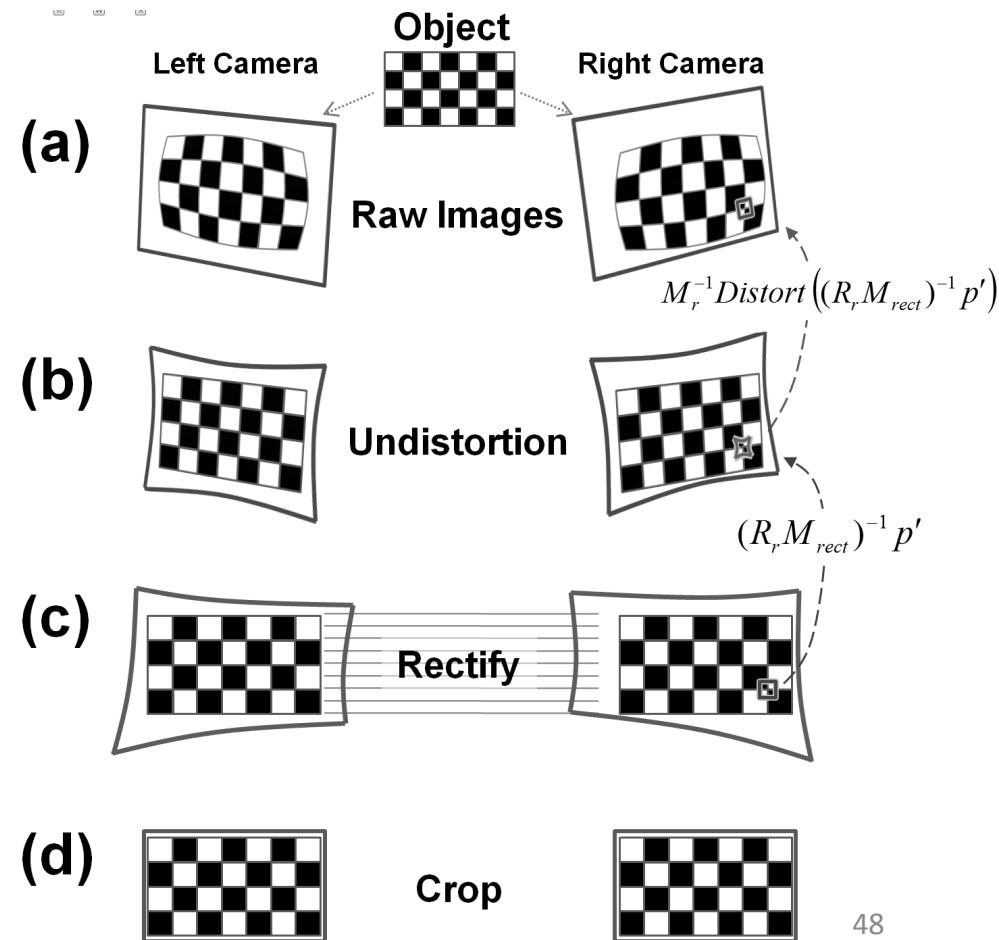
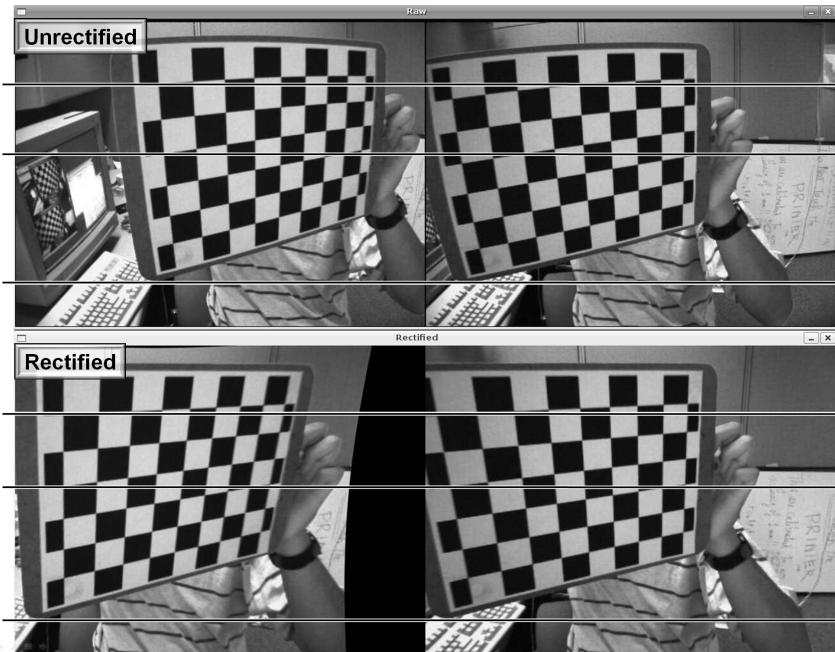
$$P' = K[I|T]$$



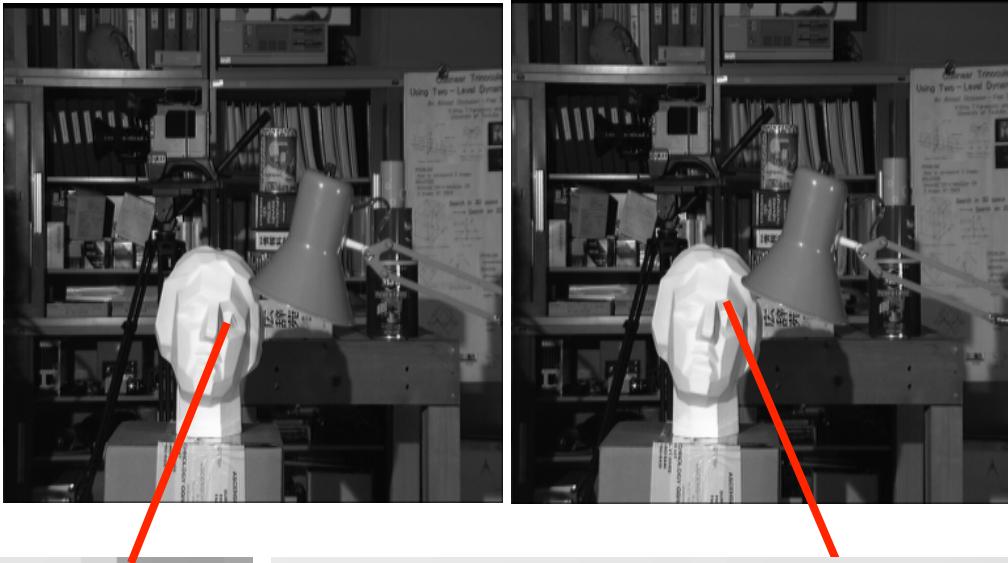
$$Z = \frac{f_x T}{x' - x}$$

Stereo Rectification

- Algorithm steps are shown at right:
- Goal:
 - Each row of the image contains the same world points
 - “Epipolar constraint”

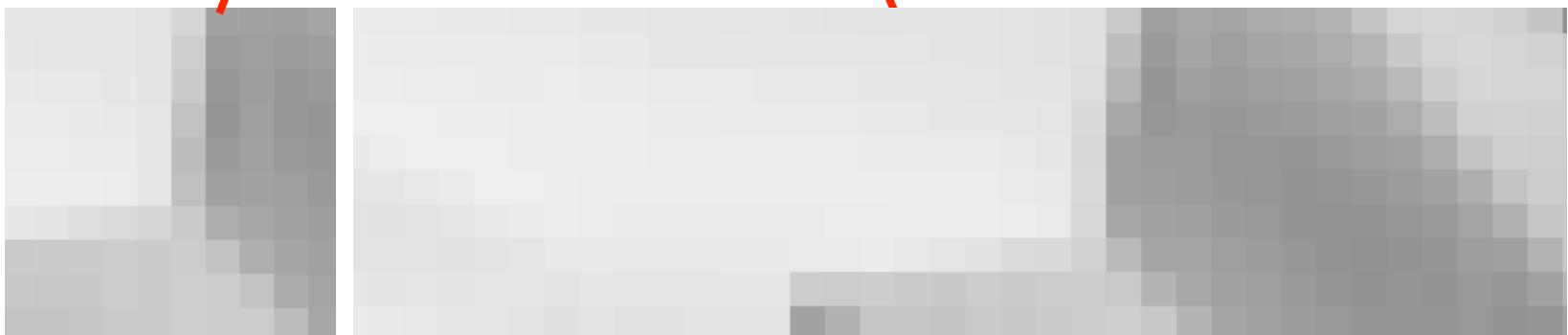


Stereo correspondence block matching

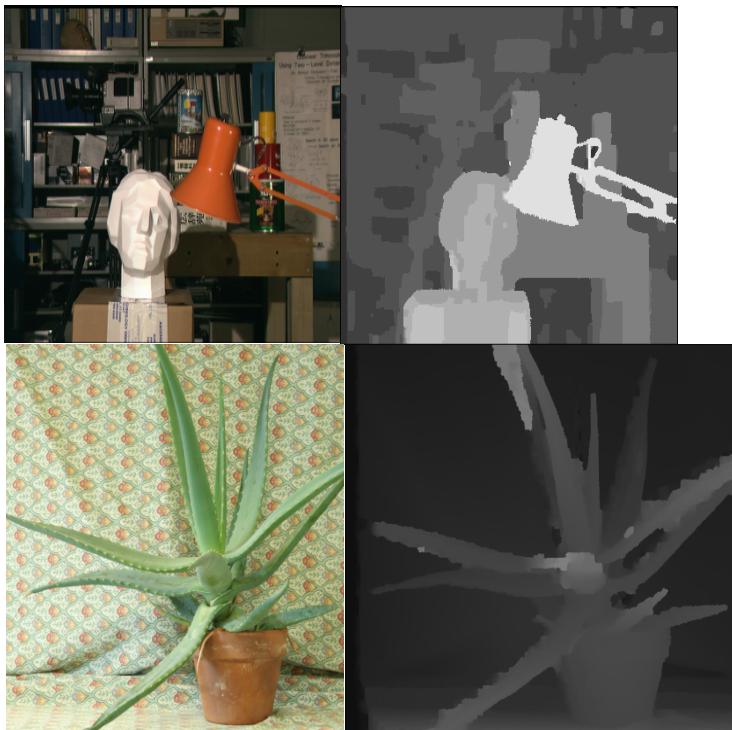
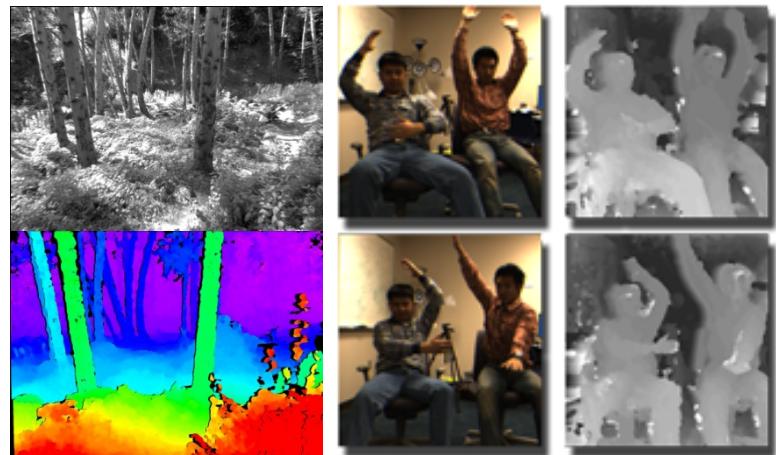


For each block in left image:

Search for the corresponding block in the right image such that SSD or SAD between pixel intensities is minimum

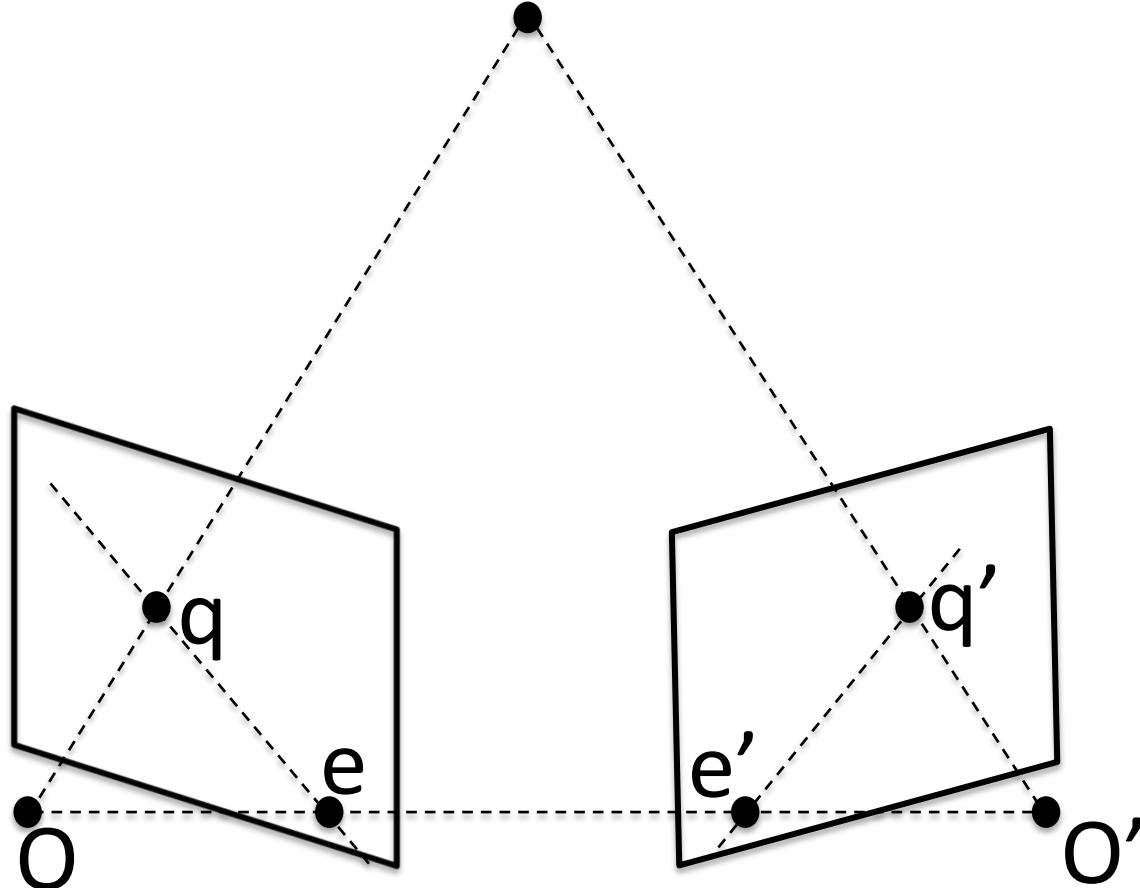


Stereo Matching



$$Z = f_x \frac{T}{D}$$

Structure from motion



Случай 1: некалиброванная камера

Projective reconstruction theorem

- Пусть (q_i, q'_i) – пары соответствующих точек на двух изображениях, и существует единственная F такая, что $q'_i F q_i = 0$ для всех i . Пусть $(P_1, P'_1, \{Q_{1i}\})$ и $(P_2, P'_2, \{Q_{2i}\})$ – две реконструкции, соответствующие (q_i, q'_i) . Тогда существует несингулярная матрица H такая, что $P_2 = P_1 H^{-1}$, $P'_2 = P'_1 H^{-1}$ и $Q_{2i} = H Q_{1i}$, кроме точек, для которых верно $F q_i = q'^T_i F = 0$

Случай 2: калиброванная камера

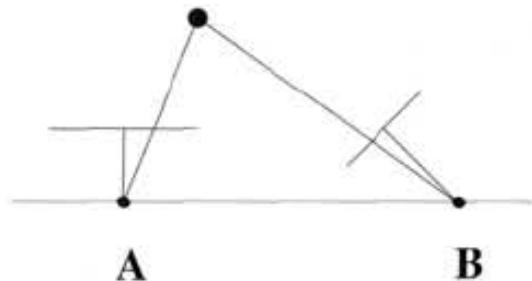
- Для матрицы $F=U\text{diag}(1,1,0)V^T$ и матрицы проекции первой камеры $P=[E|0]$, существуют 4 возможных значения матрицы проекции второй камеры:

$$P' = [UWV^T|+u_3] \text{ или } P' = [UWV^T| -u_3] \text{ или}$$

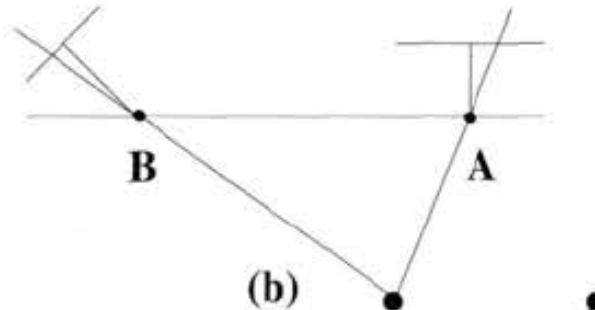
$$P' = [UW^TV^T|+u_3] \text{ или } P' = [UW^TV^T| -u_3]$$

$$W = \begin{bmatrix} 0 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

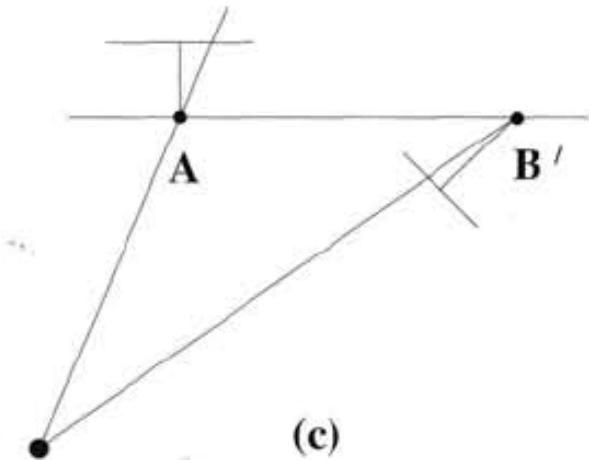
4 решения для матрицы проекции второй камеры



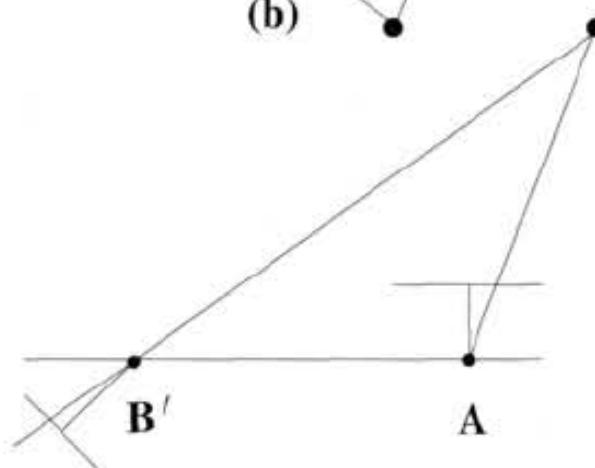
(a)



(b)



(c)

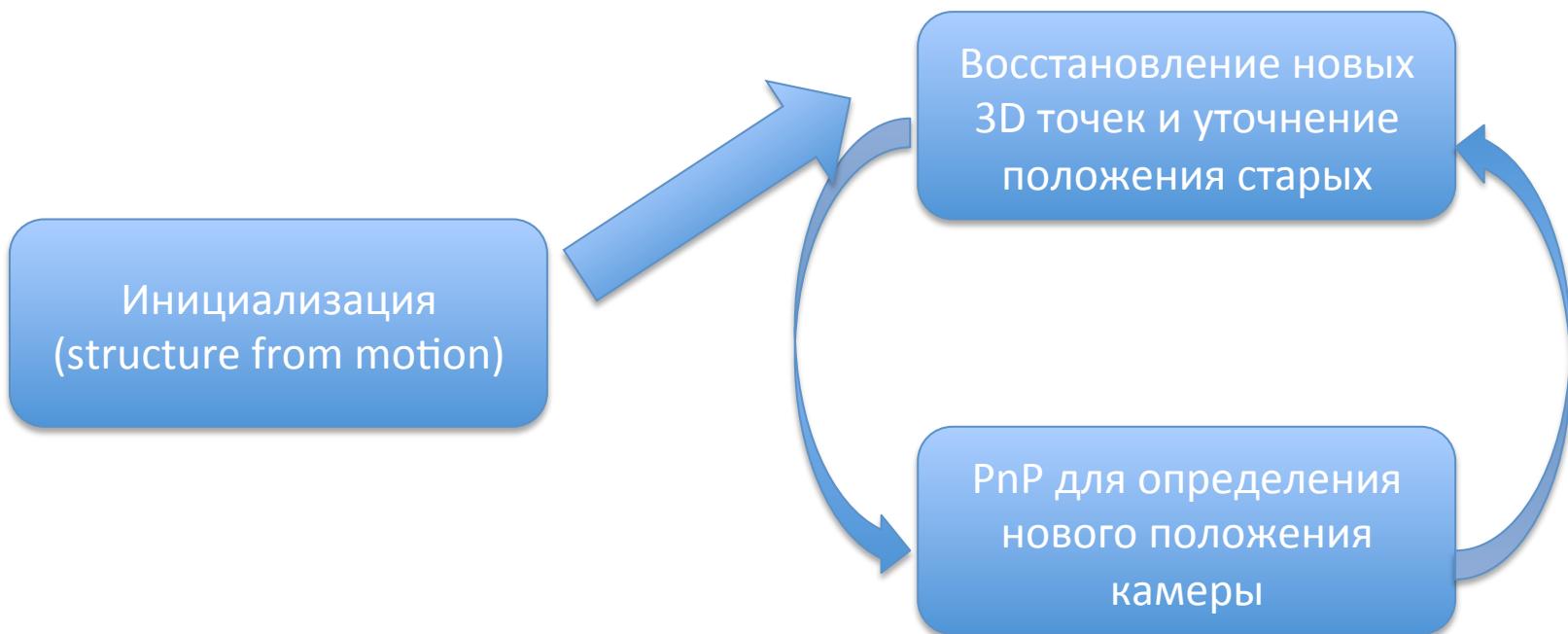


(d)

Содержание

- Pinhole модель камеры
- Формализм однородных координат для проективного пространства
- Задача PnP (найти позу известного объекта по его проекции)
- Фундаментальная матрица для пары камер
- Стерео рекалибрация
- Structure from motion (восстановление трехмерной сцены при помощи одной камеры)

Структура алгоритма визуальной одометрии



OpenCV functions: single view

- SVD
- projectPoints
- undistort
- solvePnP
- calibrateCamera
- findChessboardCorners/findCirclesGrid

OpenCV functions: two views

- findHomography
- findFundamentalMat
- stereoCalibrate
- stereoRectify/reprojectImageTo3D/
initUndistortRectifyMap

Откуда берутся такие искажения?



Модель искажений линзы

$$x'' = x'(1 + k_1 r^2 + k_2 r^4 + k_3 r^6) + 2p_1 x' y' + p_2(r^2 + 2x'^2)$$

$$y'' = y'(1 + k_1 r^2 + k_2 r^4 + k_3 r^6) + p_1(r^2 + 2y'^2) + 2p_2 x' y'$$

$$\text{where } r^2 = x'^2 + y'^2$$

$$u = f_x * x'' + c_x$$

$$v = f_y * y'' + c_y$$