



Геометрия камеры



Many slides adopted from Svetlana Lazebnik, Steve Seitz, Richard Hartley and Noah Snavely



План лекции

- Устройство фотоаппарата
- Модель перспективной проекции
- Гомография
- Геометрия двух камер

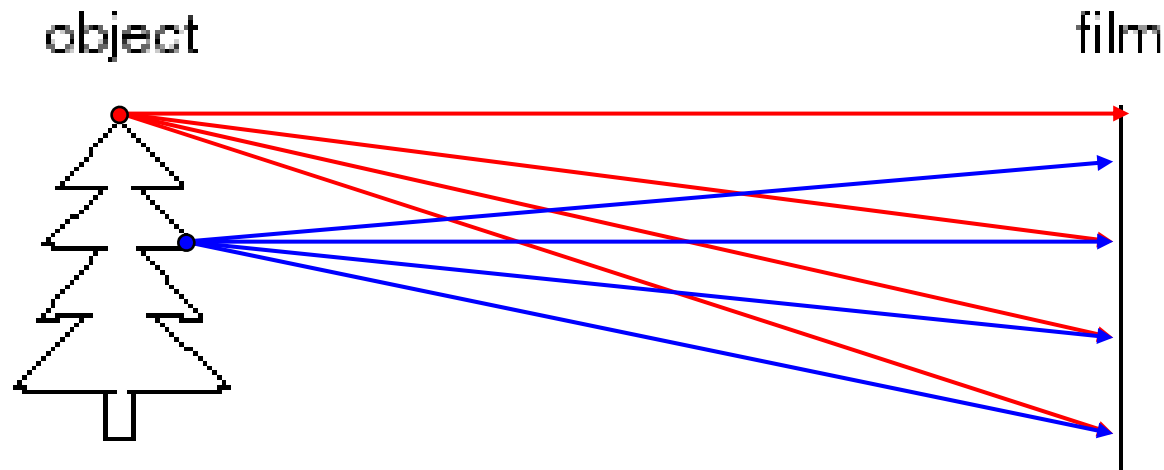


Геометрия камеры

- В первой лекции курса мы рассматривали, как изображение кодируется в глазу человека и на матрице цифровой фотокамеры
- Теперь рассмотрим устройство камеры с точки зрения траекторий лучей света от объекта до его изображения
- Построим геометрическую модель формирования изображения



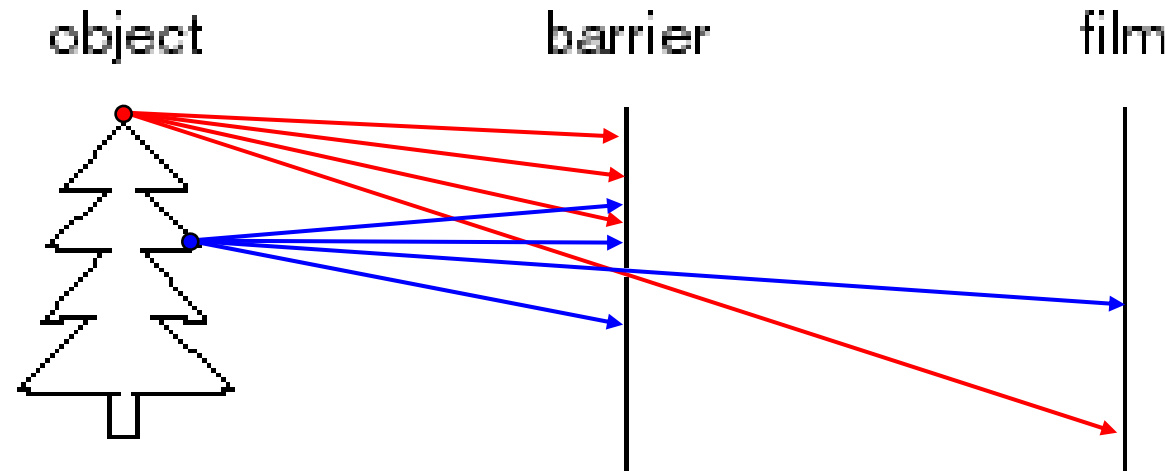
Сделаем простейшую камеру



- Положим кусок пленки перед объектом
- На каждую точку пленки падают лучи от всех точек объекта
- Получается бессмысленное изображение



Камера-обскура

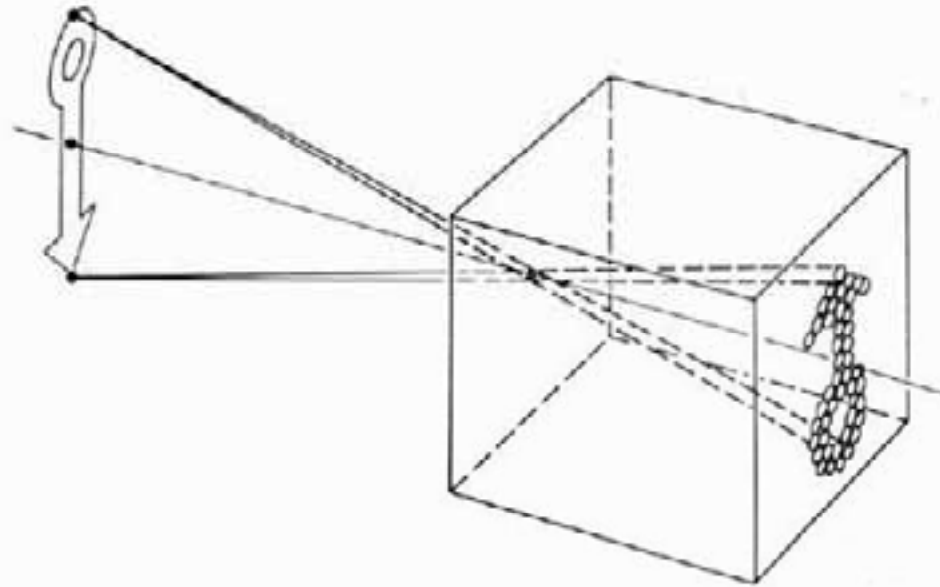


Добавим преграду с отверстием, чтобы отсечь лишние лучи

- На каждую точку пленки падает один луч!
- Картинка должна быть чёткой
- Отверстие в преграде называется апертурой или диафрагмой



Модель камеры-обскуры



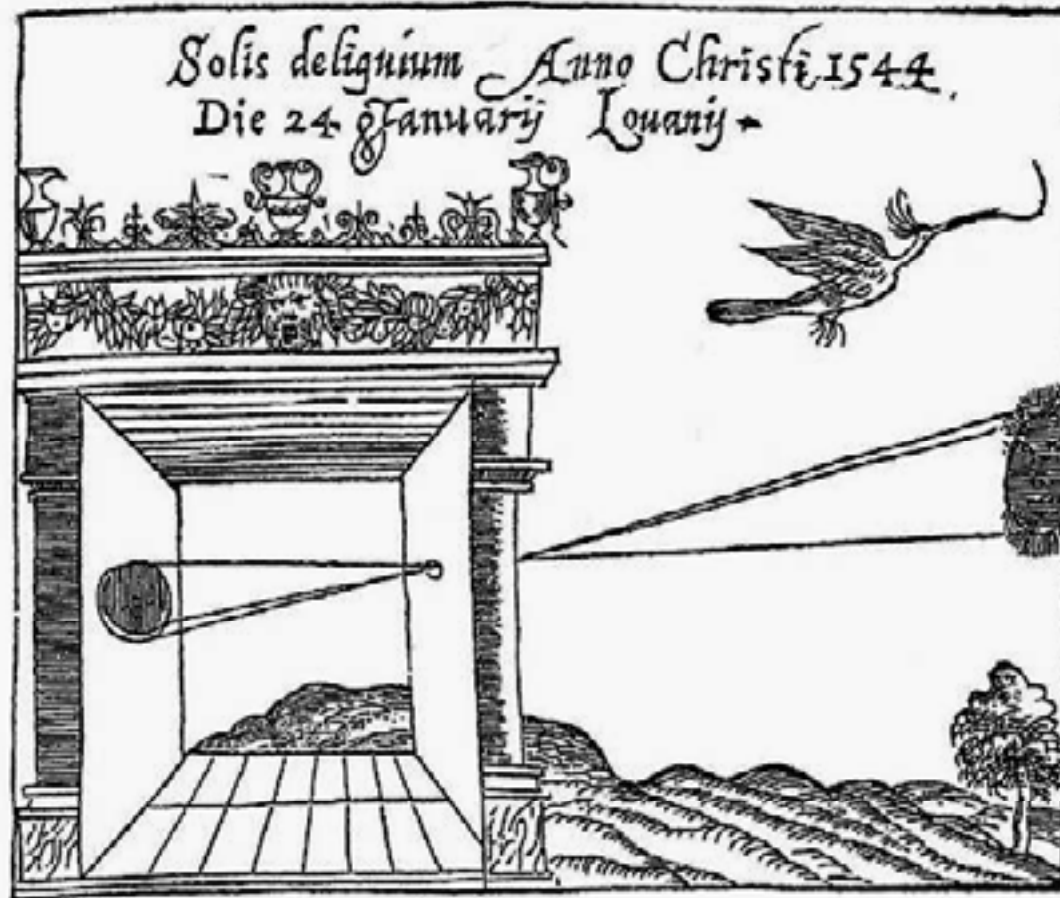
Модель:

- В преграде отверстие размеров в одну точку
- Все лучи проходят через одну точку
- Эта точка называется **Центром Проекции (ЦП)**
- Реальный размер отверстия – **Апертура**
- Изображение формируется на **Картинной плоскости**
- **Фокусным расстоянием f** называется расстояние от ЦП до Картинной плоскости



Камера-Обскура

Camera Obscura, Gemma Frisius, 1558



- Самая Первая Камера
- Была известная еще Аристотелю
- Глубина комнаты и есть Фокусное расстояние



В домашних условиях



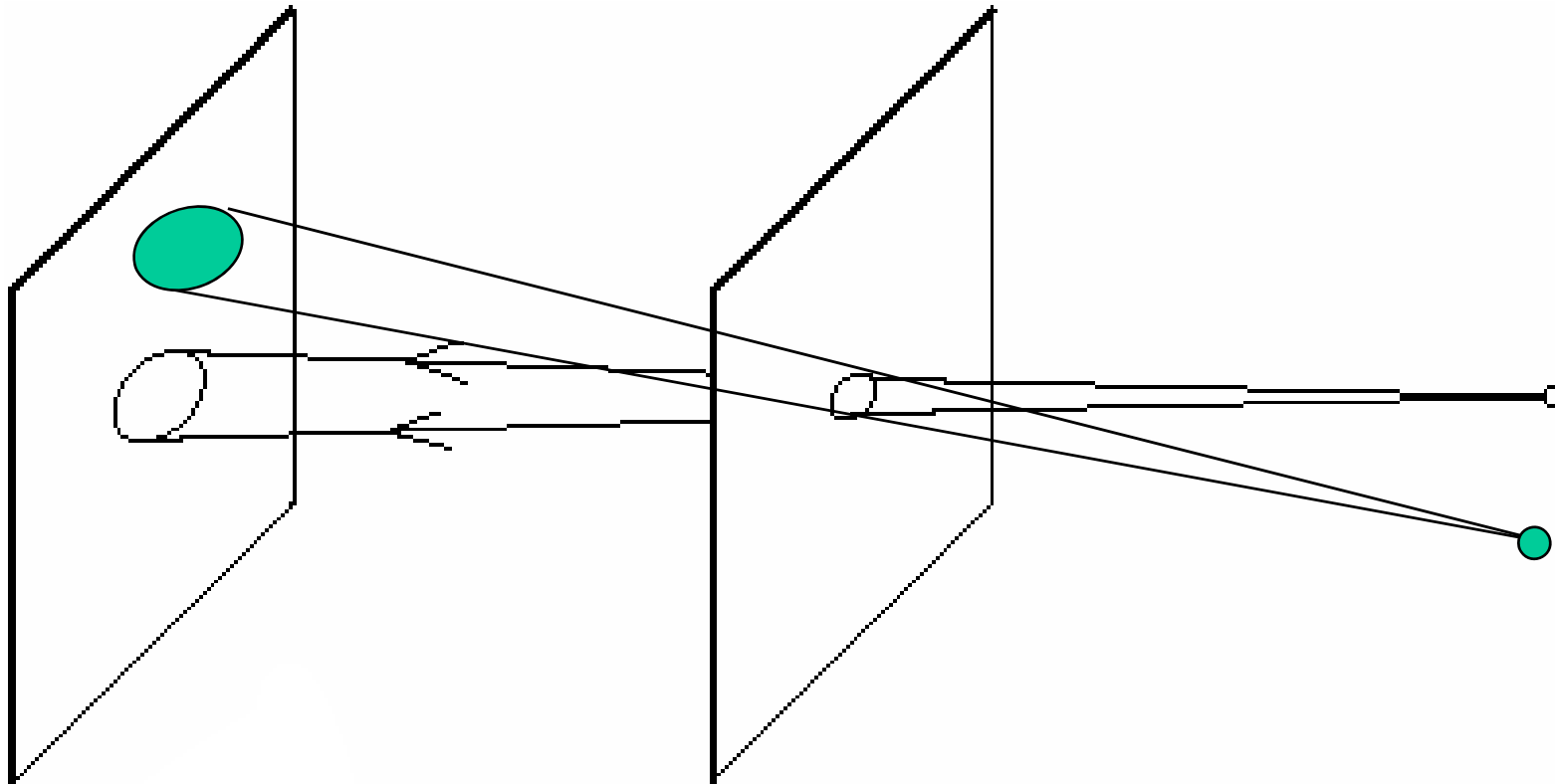
Почему
картинка
такая
размытая?



<http://www.debevec.org/Pinhole/>



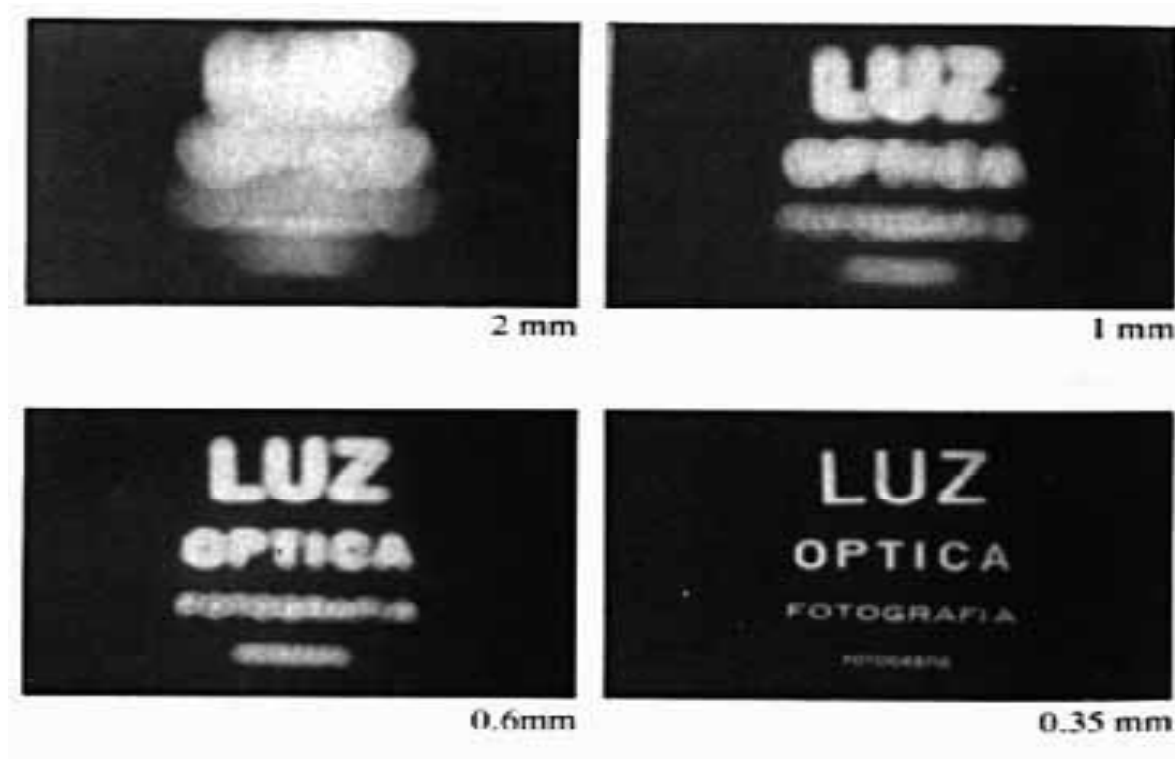
Апертура – не точка!



- Целый пучок лучей проходит через отверстие в преграде
- Изображение одной точки – небольшой кружок
- Размер кружка зависит от размера апертуры



Уменьшаем апертуру

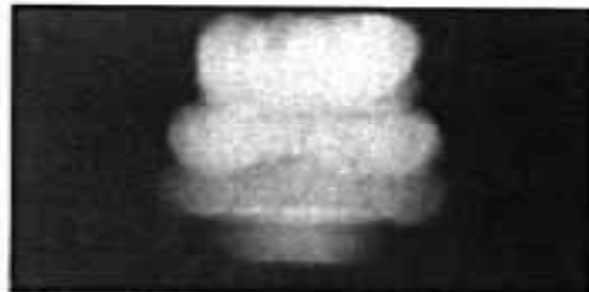


Почему бы не сделать апертуру как можно меньше?

- Меньше апертура – меньше света проходит
- При малых апертурах начинаются дифракционные эффекты



Уменьшаем диафрагму



2 mm



1 mm



0.6 mm



0.35 mm



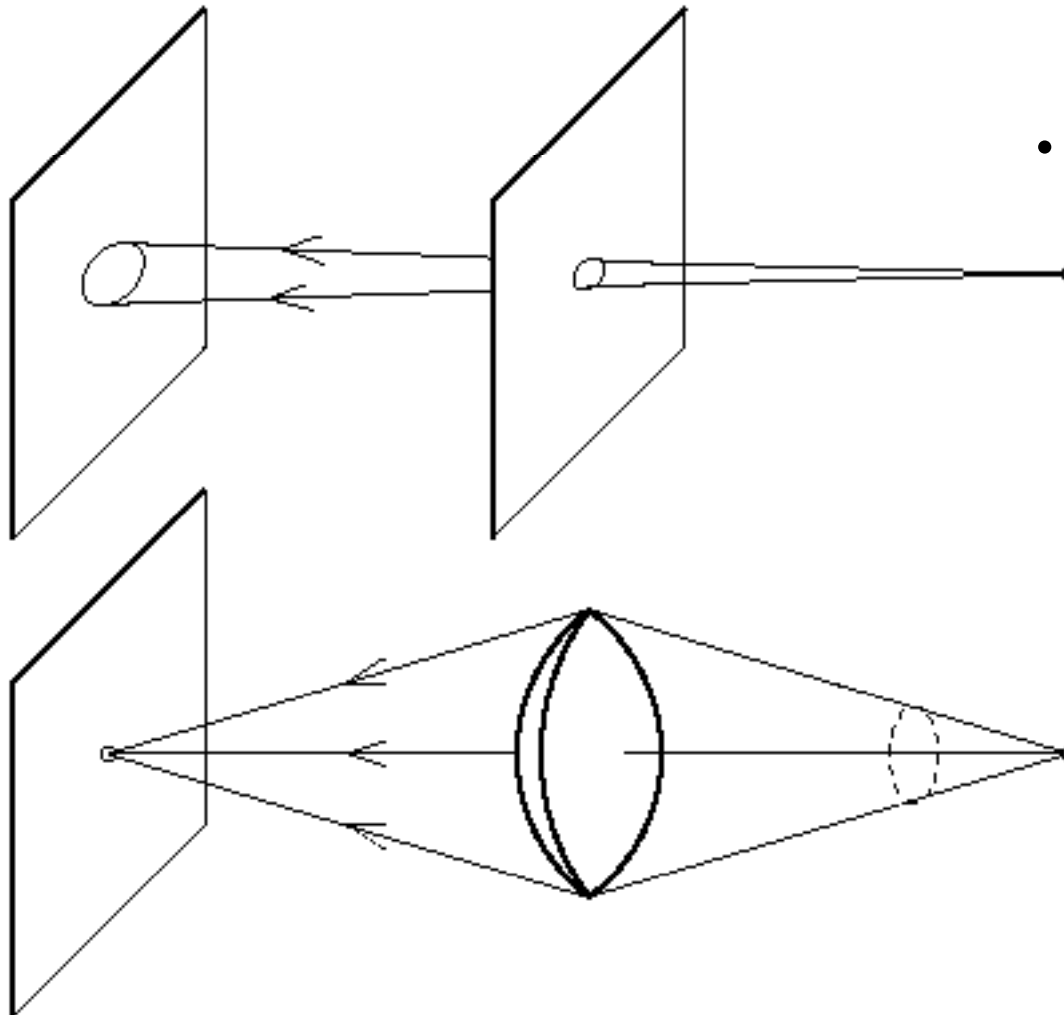
0.15 mm



0.07 mm



Линза!

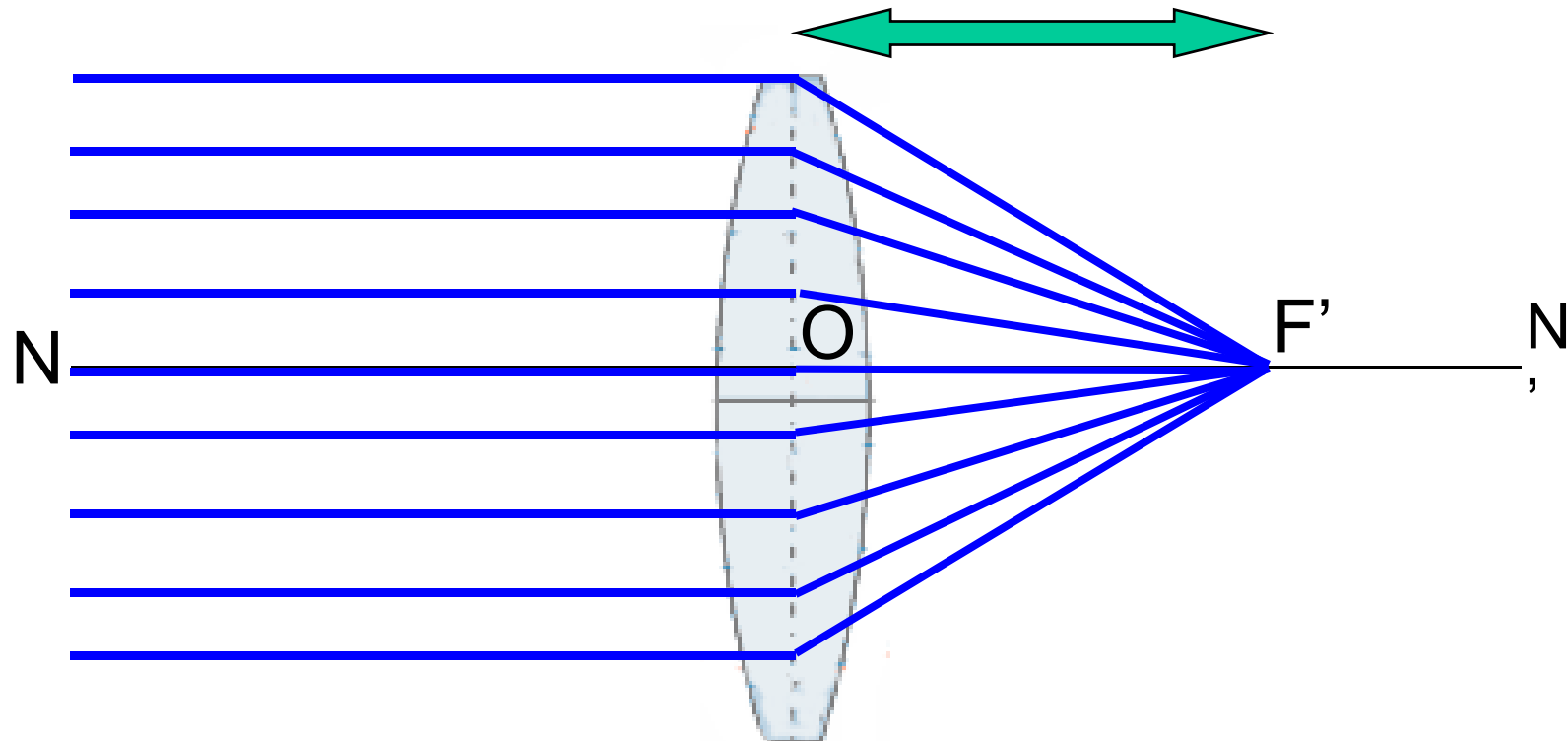


- Линза:
диск из прозрачного
однородного материала,
ограниченный двумя
полированными
сферическими, или
плоской и сферической,
поверхностями

Линза позволяет увеличить поток света от каждой точки



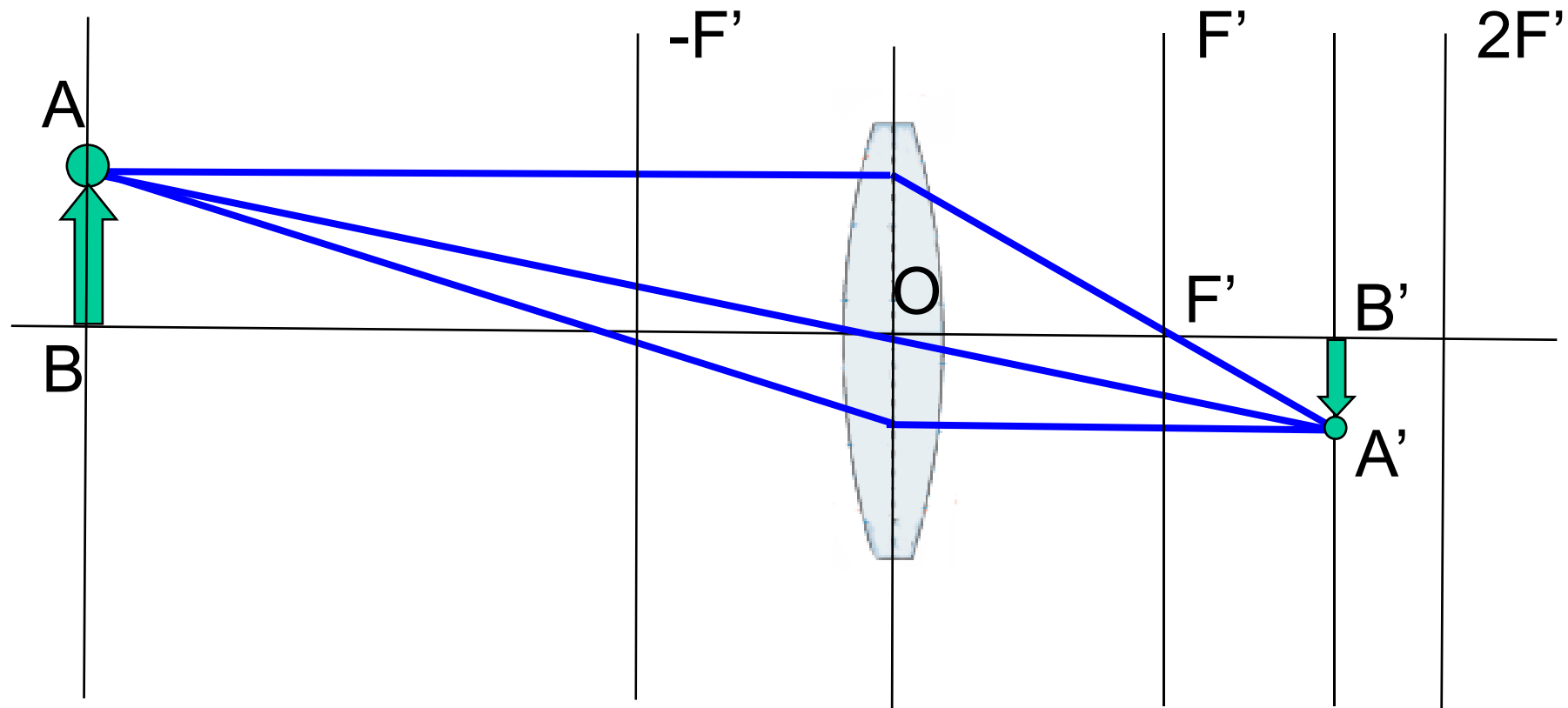
Линза



- NN' – главная оптическая ось, пересекающая центры сферических поверхностей
- Пучок параллельных прямых пересекается в **главном фокусе F'**
- $OF' = f$ – главное фокусное расстояние



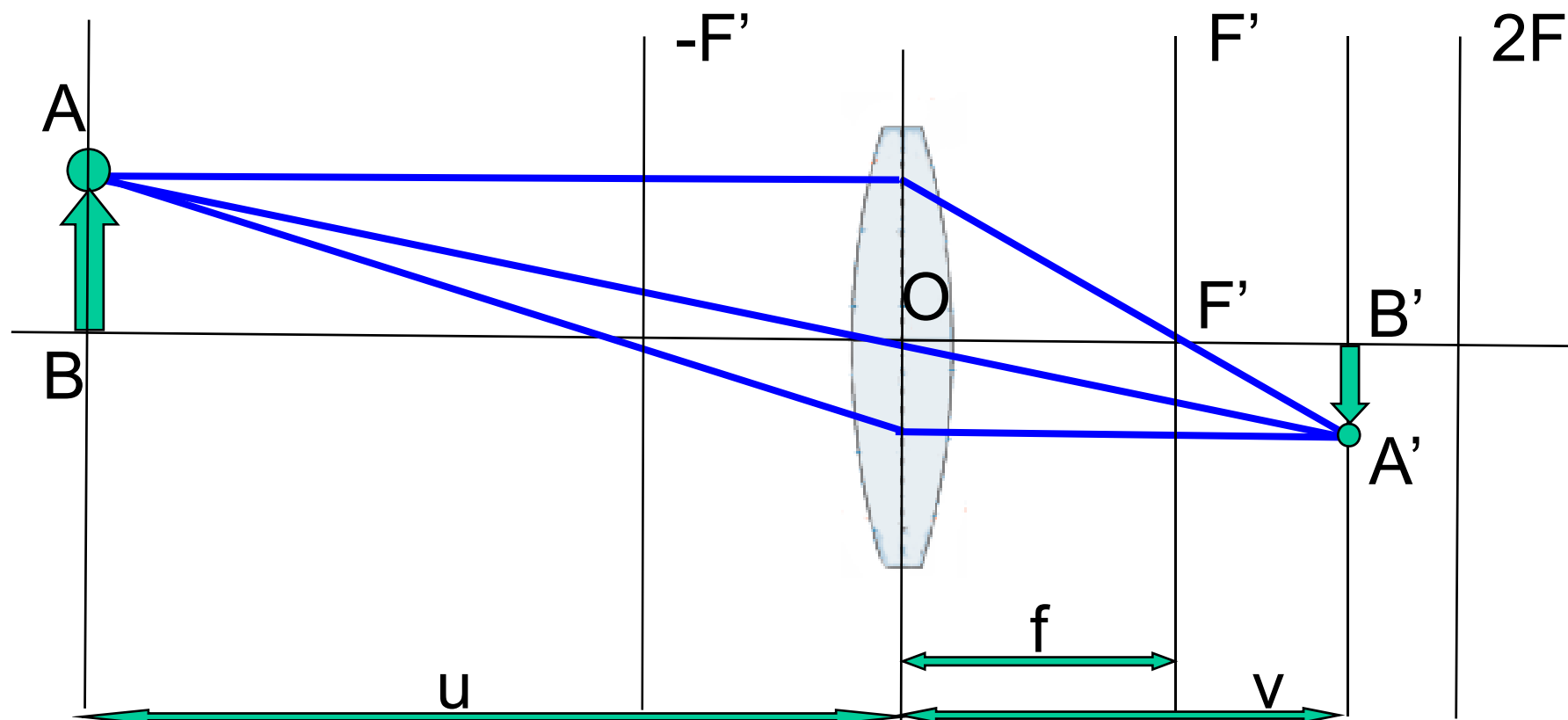
Линза



- Лучи от одной точки объекта, преломляясь линзой, фокусируются в одной точке позади линзы
- Луч, проходящий через центр линзы не преломляется
- Система точно как камера-обскура, но собирает больше света
 - O – центр проекции (линзы)



Линза



$$\frac{1}{u} + \frac{1}{v} = \frac{1}{f} \quad - \text{Основное уравнение линзы}$$

- u, v – сопряженные фокусные расстояния



Фотоаппарат



Figure from US Navy Manual of Basic Optics and Optical Instruments, prepared by Bureau of Naval Personnel. Reprinted by Dover Publications, Inc., 1969.

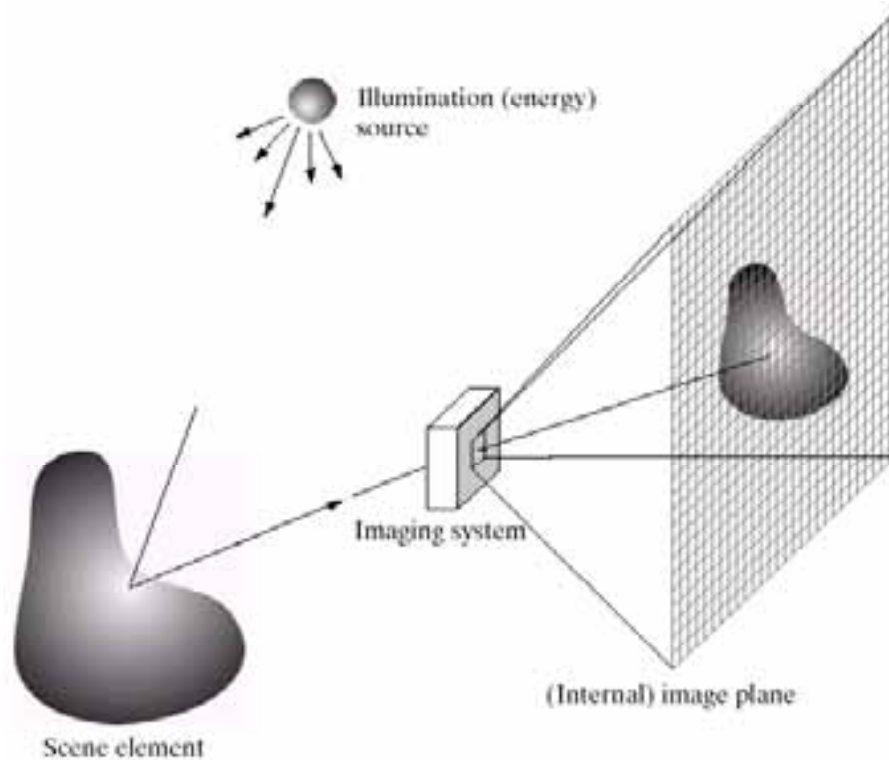


Все фотоаппараты устроены на описанном принципе

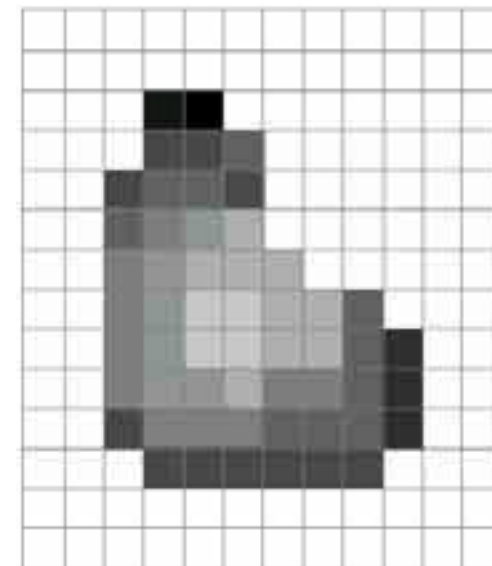
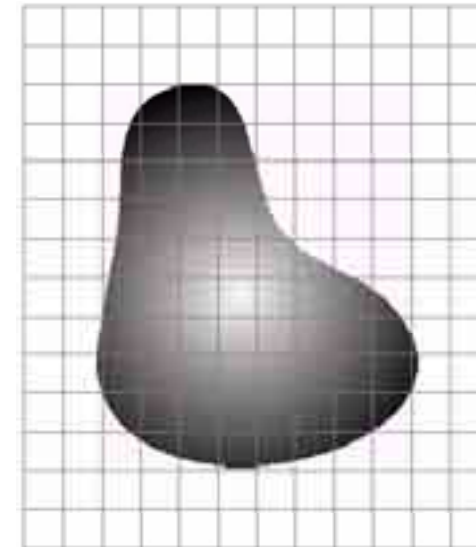
- Вместо линзы более сложная оптическая система (объектив)
- Изображение формируется на плёнке или матрице



Пиксельное изображение

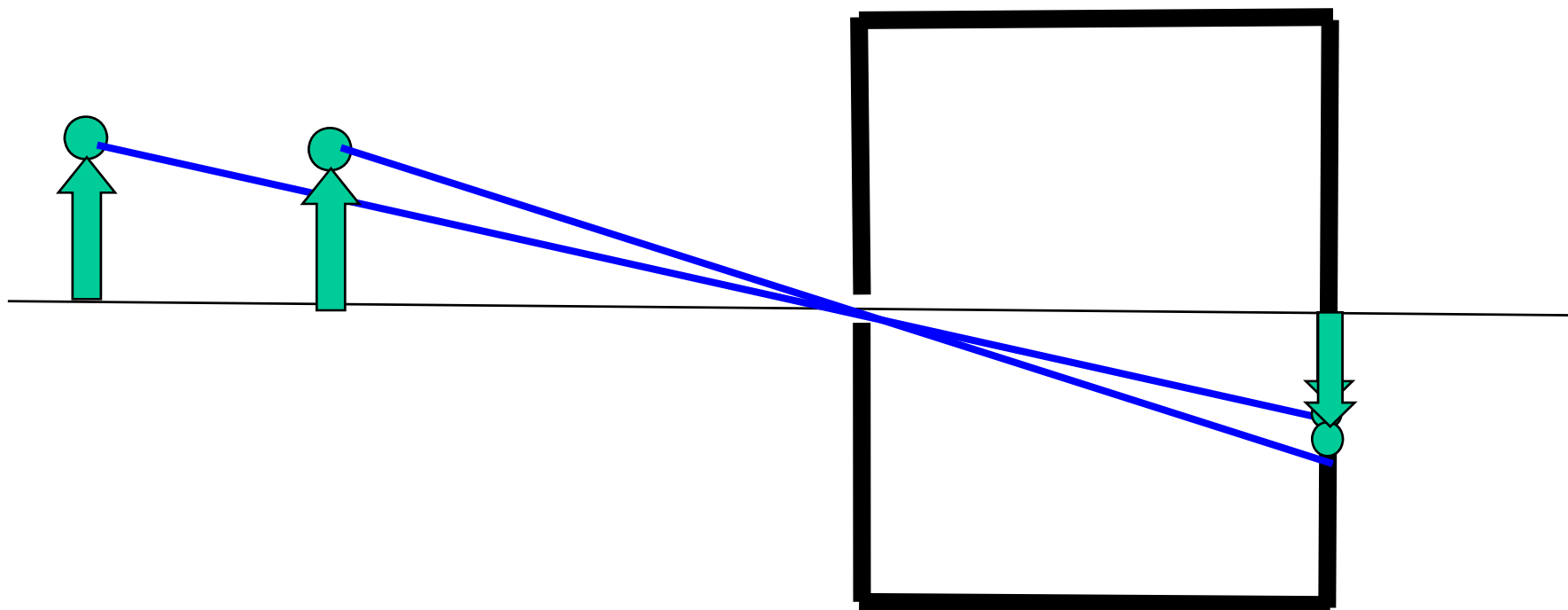


- Спроецированное изображение непрерывное, с гладкими границами
- На матрице оно дискретизируется
 - По пространству (пиксельная решетка)
 - По цвету





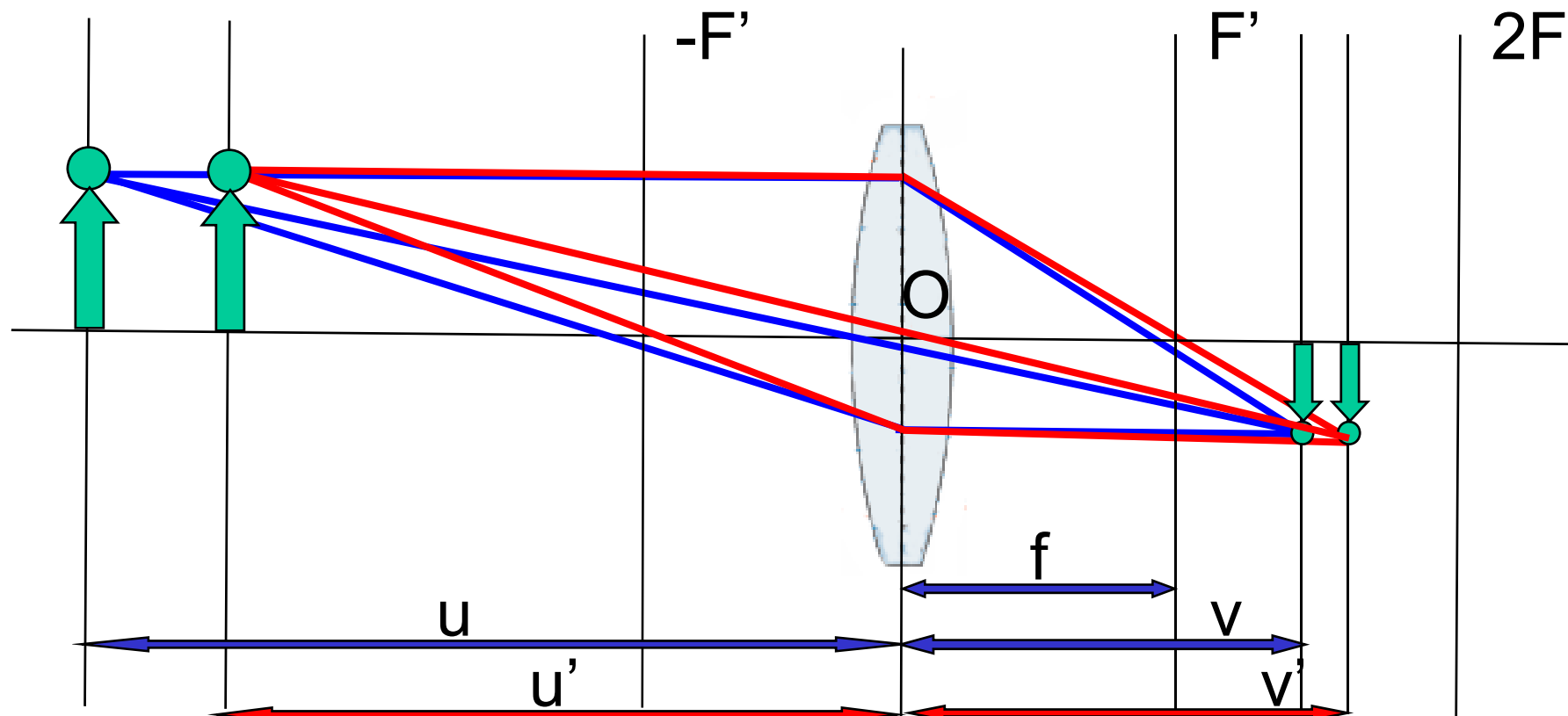
Идеальная камера-обскура



- От любой точки объекта, в не зависимости от расстояния, через отверстие проходит только один луч
- Изображение получается чётким



Фокусировка камеры

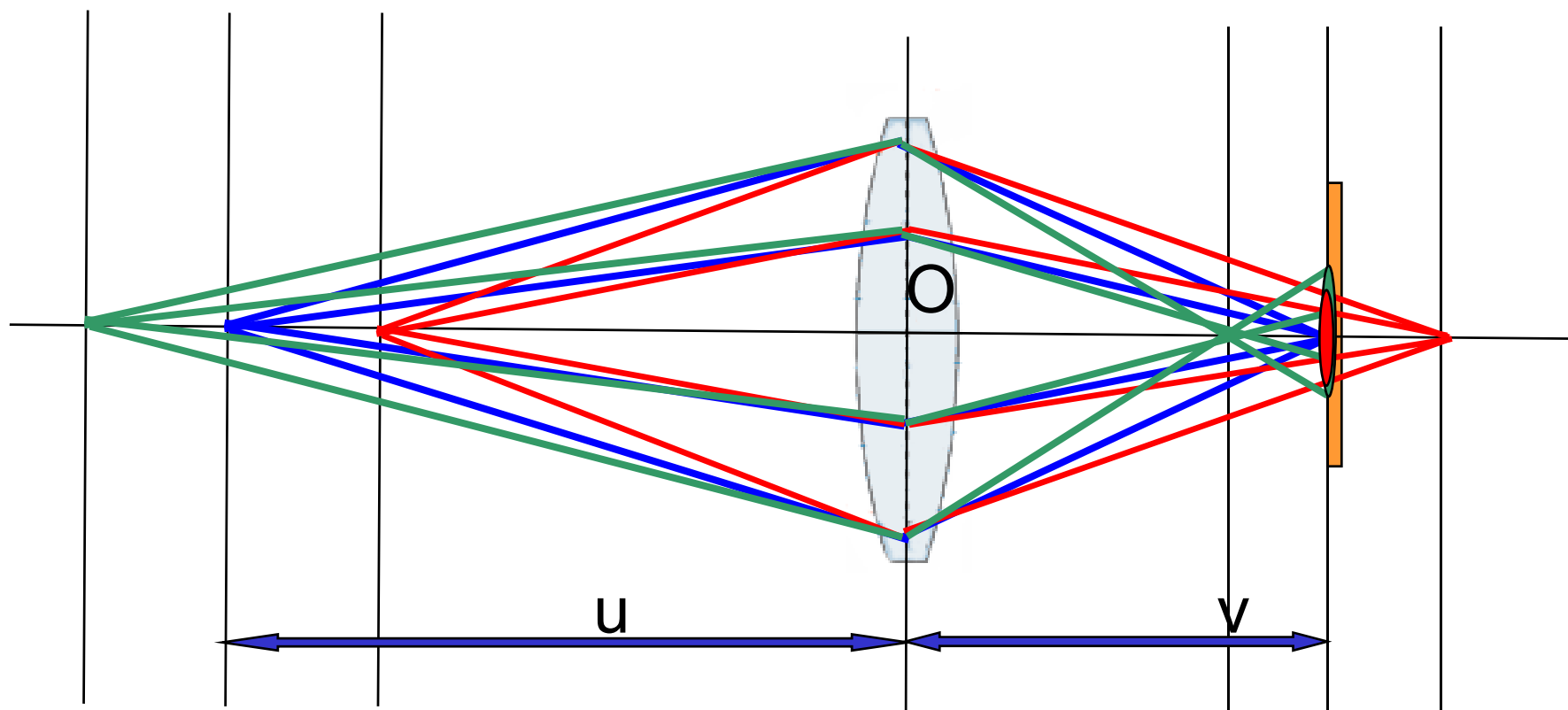


$$\frac{1}{u} + \frac{1}{v} = \frac{1}{f} \quad \frac{1}{u'} + \frac{1}{v'} = \frac{1}{f} \quad (u' \neq u) \Rightarrow (v' \neq v)$$

- u, v – сопряженные фокусные расстояния



Пятно рассеяния



- Пусть матрица поставлена так, чтобы сошлись лучи от объекта на расстоянии u
- Пучок лучей не сходится в одну точку, а образует на пленке «кружок рассеяния» или «пятно рассеяния»



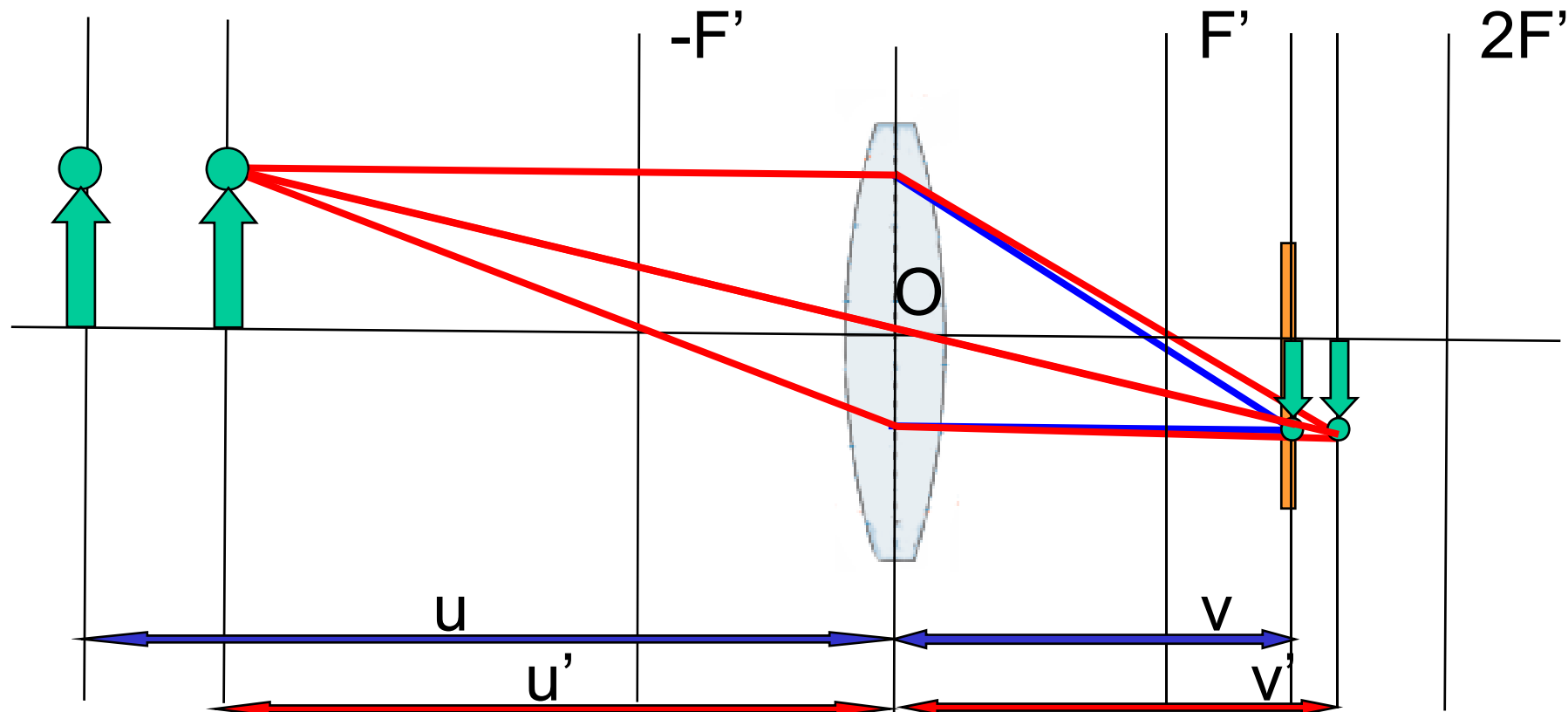
Фокусировка



Только часть объектов оказываются «в фокусе»



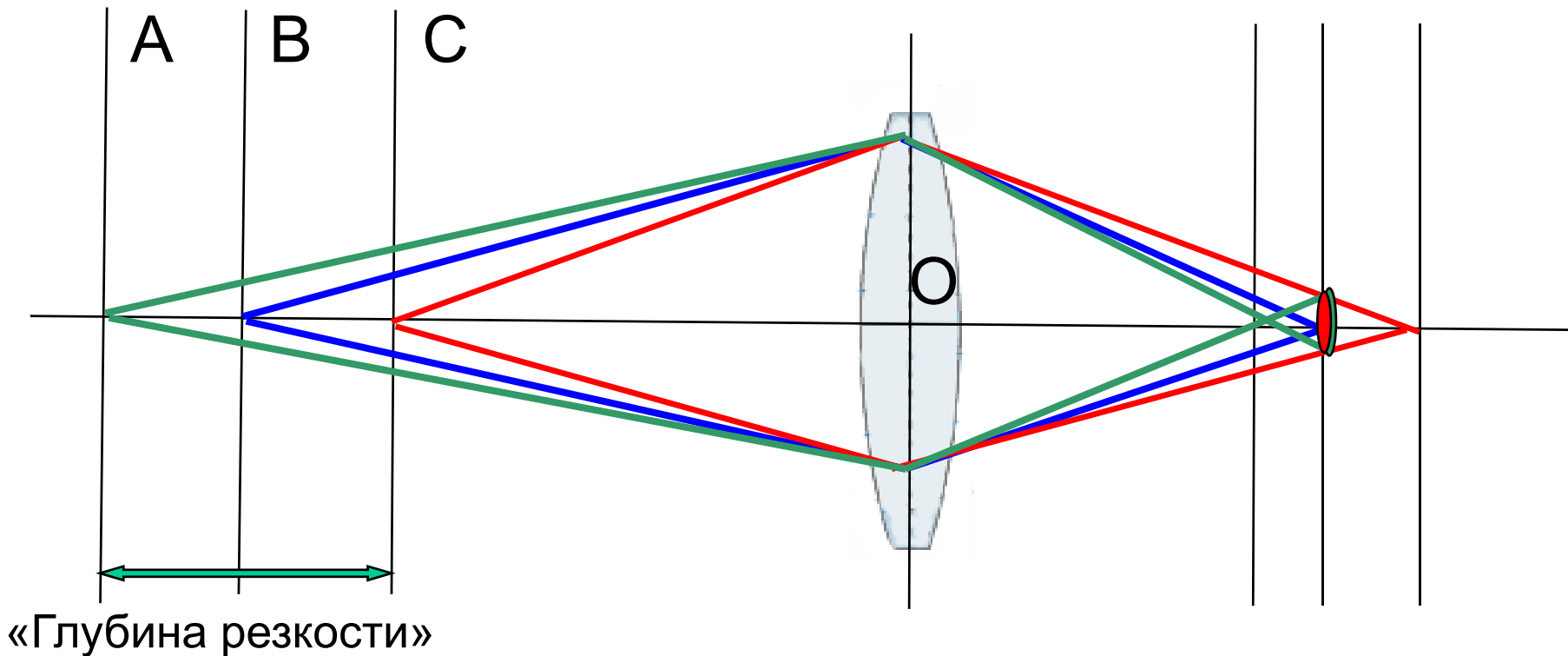
Фокусировка камеры



- Либо смещение матрицы относительно линзы (изменение фокусного расстояния)
- Либо изменения степени преломления в линзе (объективе)



Глубина резкости



- «Пятно рассеяния» маленькое – объект «в фокусе»
- Интервал в котором объекты попадают «в фокус» называется «глубиной резкости»



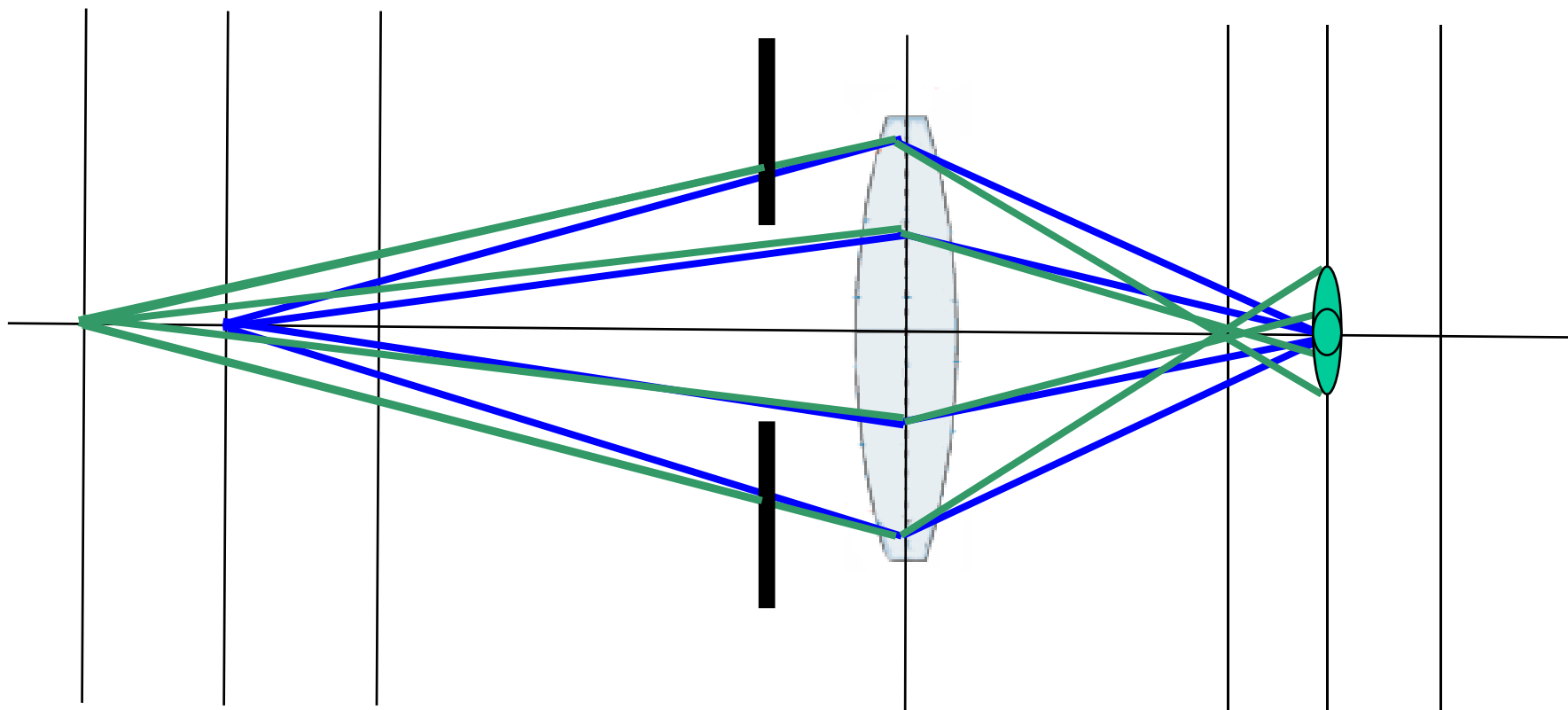
Глубина резкости



Глубина резкости = Depth of field (DOF)



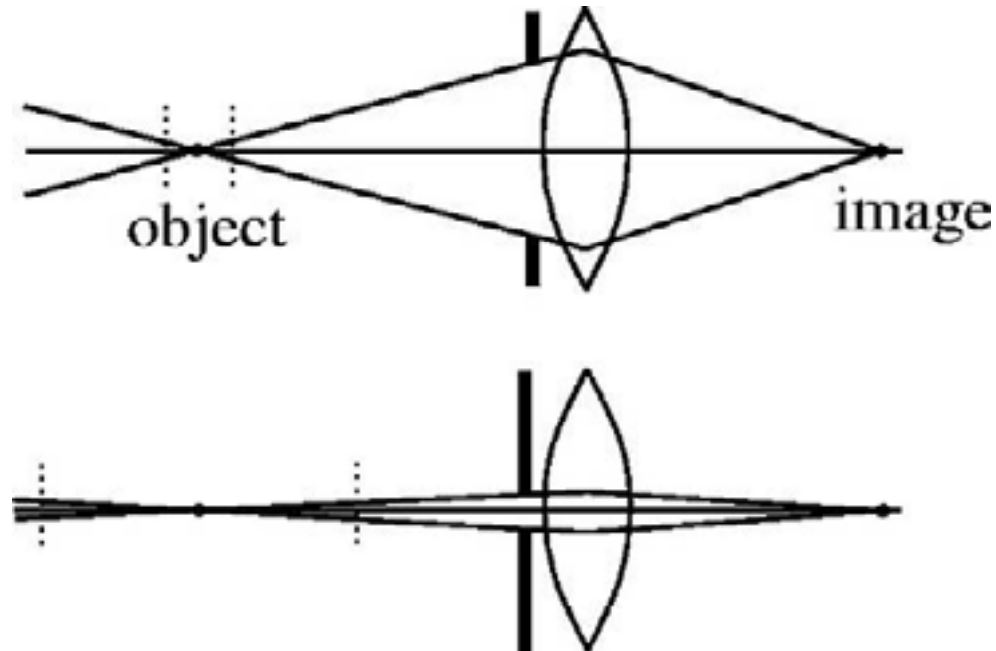
Управление глубиной резкости



Изменяя диафрагму можно изменять размер «пятен рассеяния»



Управление глубиной резкости



- Диафрагма управляет глубиной резкости
 - Уменьшение диафрагмы увеличивает интервал, на котором объект находится приблизительно в фокусе
 - Маленькая апертура также уменьшает количество света – приходится увеличивать выдержку (время экспозиции)



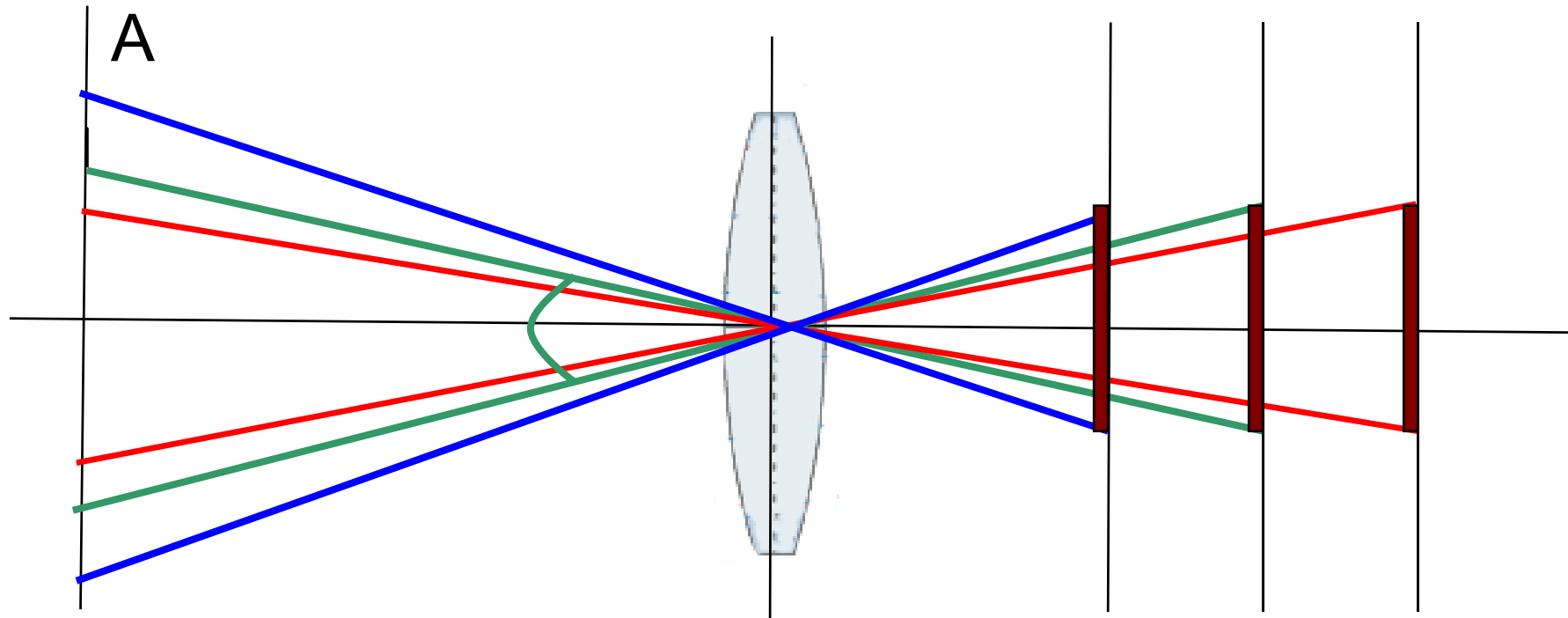
Красивый эффект

Microsoft
Research





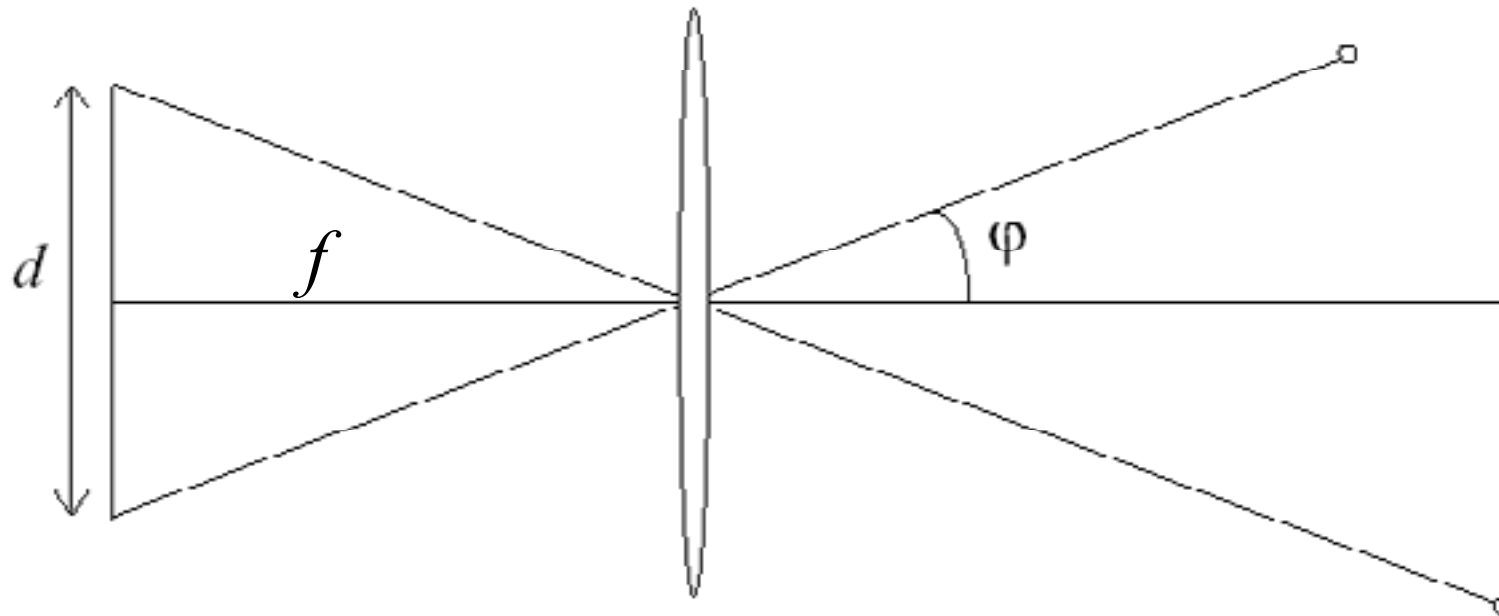
Угол обзора (поле зрения)



- Размер пленки и фокусное расстояние определяют угол обзора (field-of-view) камеры.
- Размеры пленки фиксирован
- Изменение фокусного расстояния управляет полем зрения



Угол обзора (Field of View)

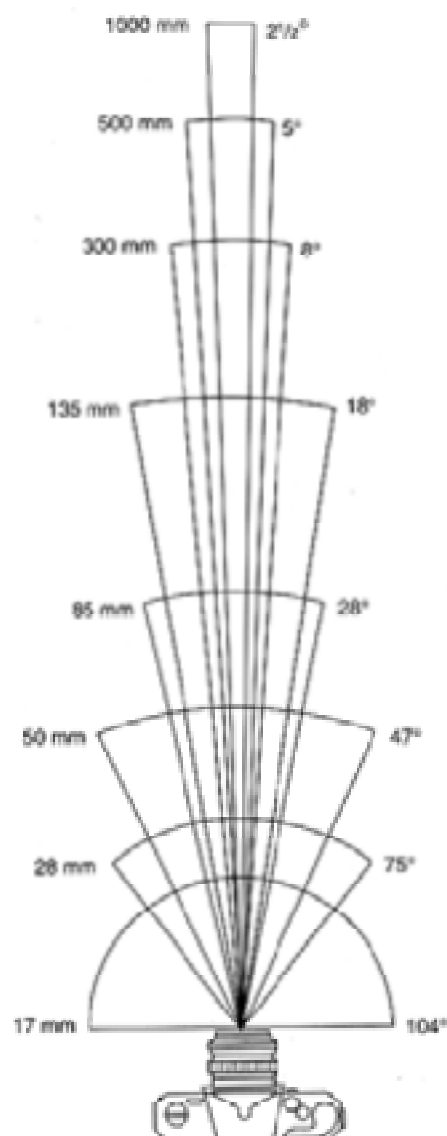


$$\varphi = \tan^{-1}\left(\frac{d}{2f}\right)$$

- Больше фокусное расстояние – меньше угол обзора
- Меньше фокусное расстояние – больше угол обзора



Трансфокация (Zoom)



17mm



28mm



50mm

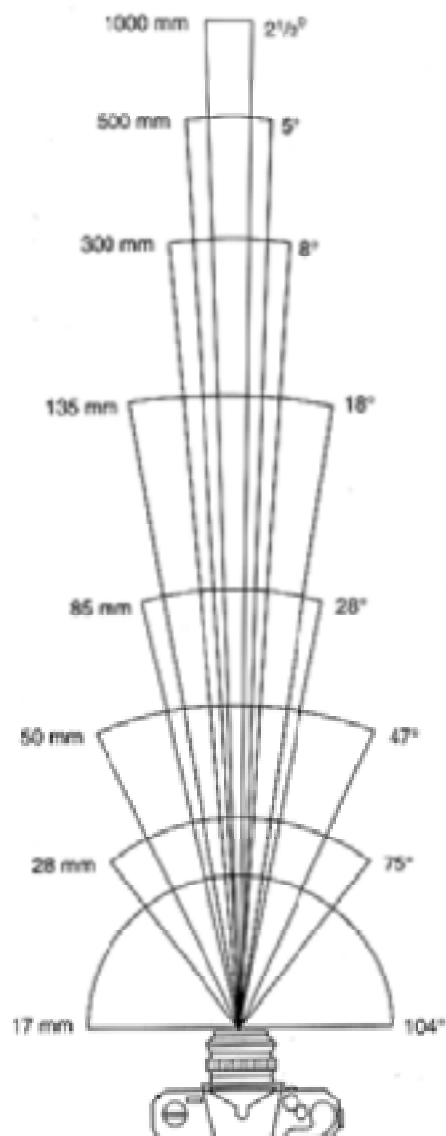


85mm

From London and Upton



Трансфокация (Zoom)



135mm



300mm



17mm



104°

From London and Upton



Поле зрения / Фокусное расстояние



Большое поле зрения
Камера близко от машины



Маленькое поле зрения
Камера далеко от машины

Чем больше фокусное расстояние – чем «параллельнее» лучи и тем меньше перспективных искажений



Забавные эффекты



<http://www.hash.com/users/jsherwood/tutes/focal/Zoomin.mov>



Figure 5.1

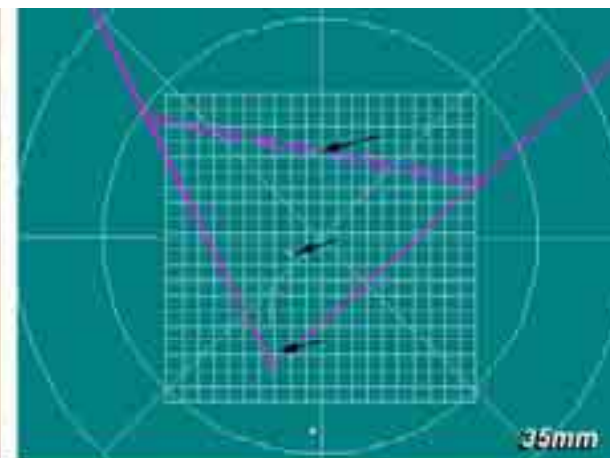
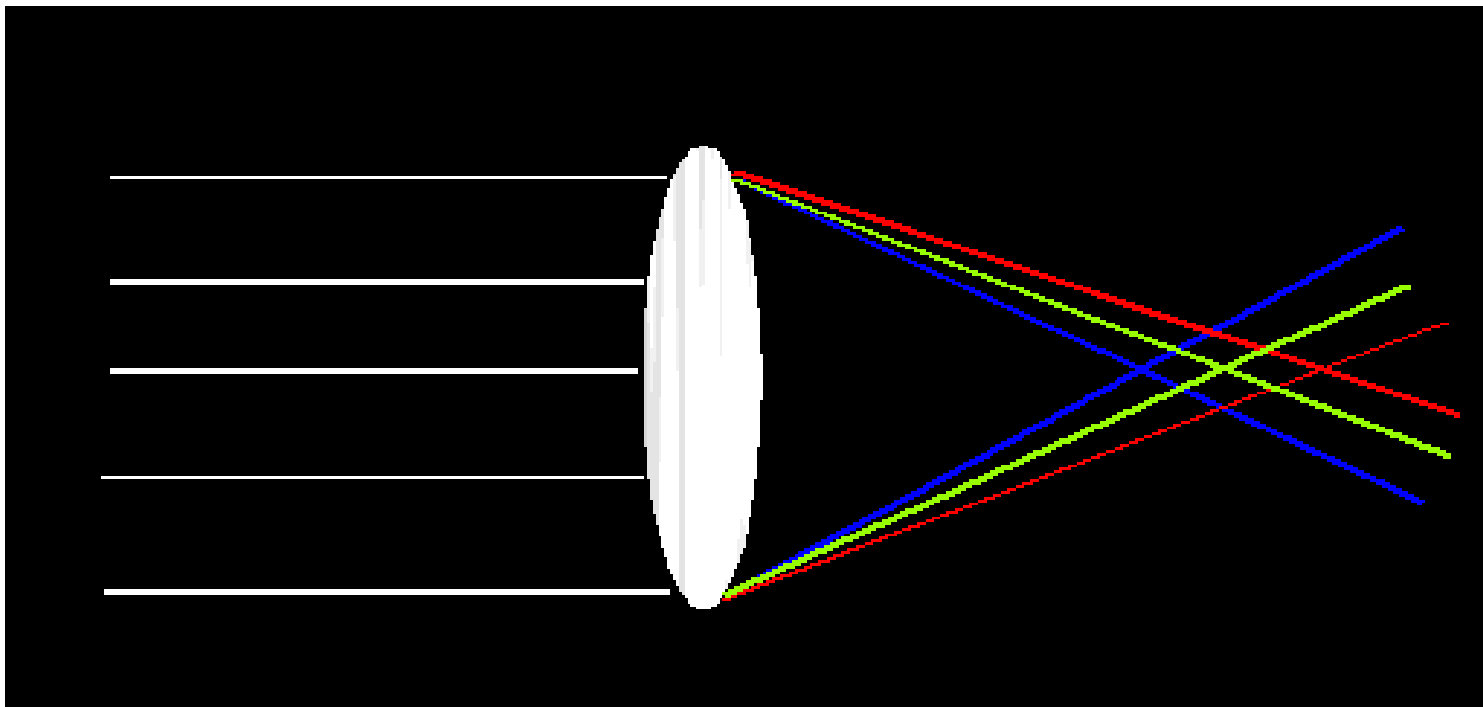


Figure 5.2



Хроматическая аберрация

- Угол преломления света зависит от длины волны
- Лучи разного цвета преломляются по-разному
- Лучи разного цвета от одной и той же точки расходятся по краям изображения





Хроматическая аберрация



Центр изображения



Край изображения



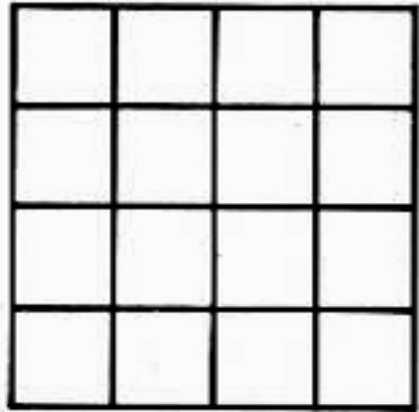
Радиальная дисторсия



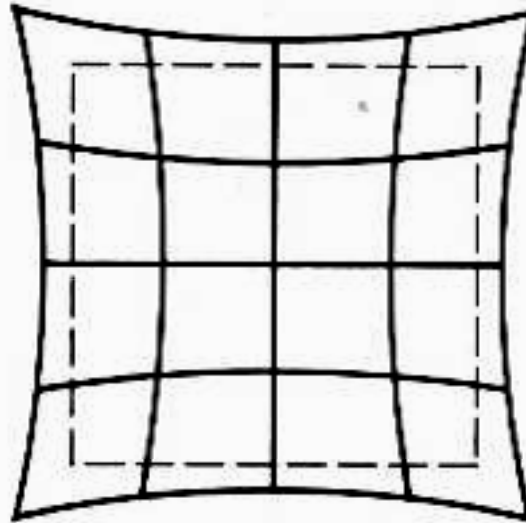
Прямые линии по краям изображения превращаются в кривые



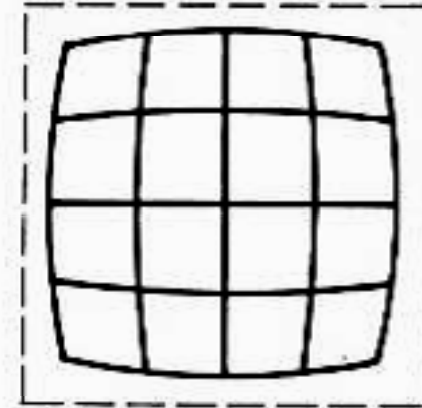
Радиальная дисторсия



Нет дисторсии



«Подушка» (положительная)



Бочкообразная

- Идеально тонких линз не бывает!
- Нарушается допущение $\sin(x) \sim x$
- Искажения наиболее заметны по краям изображения

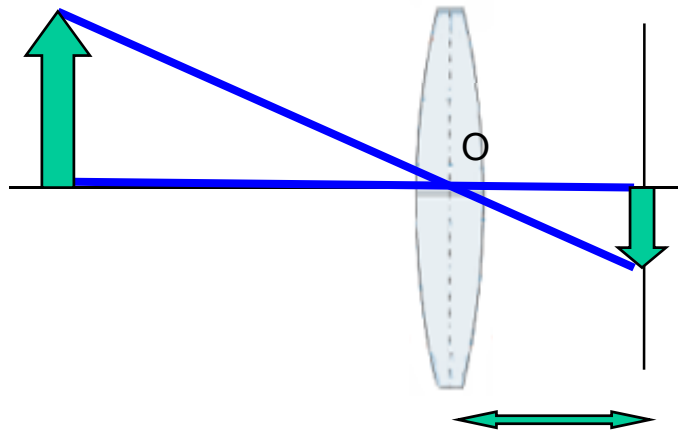


План лекции

- Устройство фотоаппарата
- **Модель перспективной проекции**
- Гомография
- Геометрия двух камер



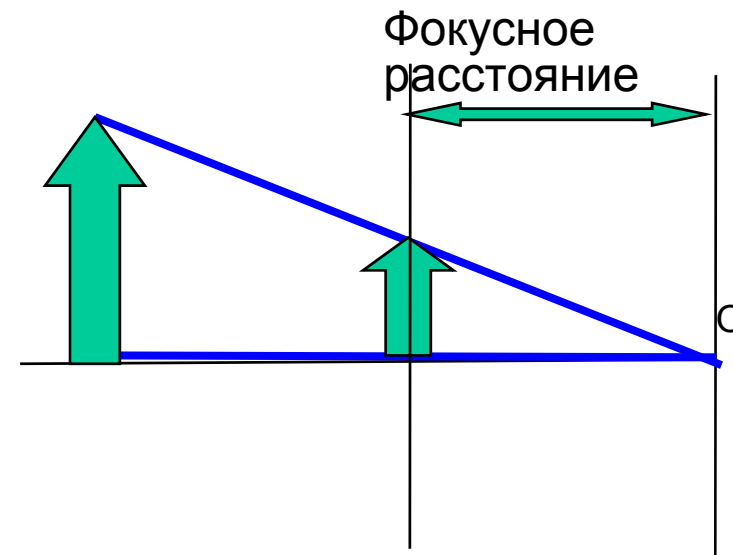
Модель



Фокусное расстояние

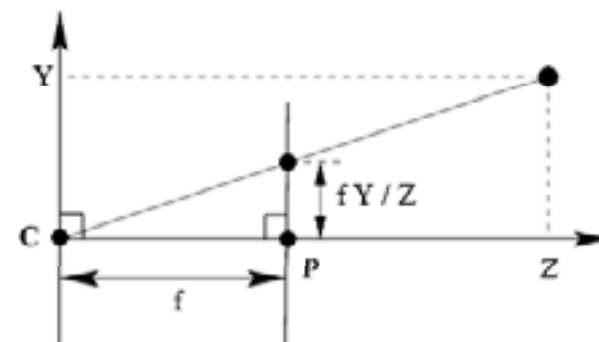
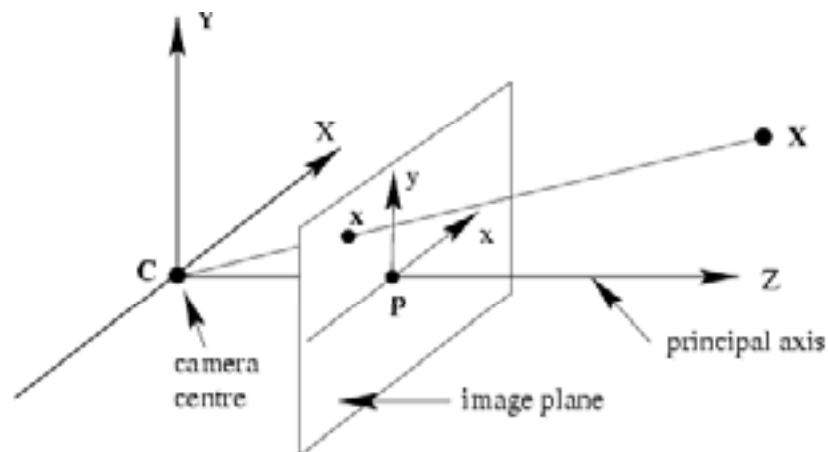
- Камера-обскура
- Изображение позади фокуса
- Изображение перевернутое

- Модель перспективной проекции
- Перенесем объект на противоположенную сторону
- То же самое фокусное расстояние!
- Изображение нормальное, не перевернутое





Модель камеры-обскуры



$C = (0,0,0)$ – центр
камеры (проекции)

$X=(X,Y,Z)$ – точка в 3х
мерном пространстве

$x=(x,y)$ – проекция X
на изображении

C , x , X лежат на одной прямой

Тогда простейшее уравнение
перспективной проекции:

$$x = f \frac{X}{Z} \quad y = f \frac{Y}{Z}$$

Перспективная проекция –
нелинейное преобразование!



В линейном виде

- Хотелось бы записать всё в виде векторов и линейных преобразований
- Для этого придётся воспользоваться представлением в однородных координатах



Однородные координаты

Рассмотрим прямые на плоскости:

$$ax + by + c = 0$$

$$(ka)x + (kb)y + (kc) = 0$$

Вектор $(a, b, c)^T$

Вектор $(ka, kb, kc)^T$

- Два вектора, отличающихся масштабом, можно считать эквивалентными
- *Однородным* (homogenous) вектором назовём класс эквивалентности, по указанному отношению эквивалентности
- $(a, b, c)^T$ - экземпляр однородного вектора
- Множество классов эквивалентности векторов в $\mathbb{R}^3 - (0,0,0)^T$ образуют проективное пространство P^2



Однородные координаты

- Точка $(x, y)^T$ лежит на линии $l = (a, b, c)^T$ если $ax + by + c = 0$

$$(x, y, 1)(a, b, c)^T = 0$$

- Отметим, что $(kx, ky, k)l = 0$, только при $(x, y, 1)l = 0$

- Тогда пусть $(kx, ky, k)^T$ - представление точки $(x, y)^T$ в однородных координатах пространства P^2



Однородные координаты

Перевод из обычных в однородных:

$$(x, y) \Rightarrow \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

Однородные координаты
точки изображения

$$(x, y, z) \Rightarrow \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

Однородные координаты
точки сцены

Перевод из однородных в обычные:

$$\begin{bmatrix} x \\ y \\ w \end{bmatrix} \Rightarrow (x/w, y/w)$$

$$\begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} \Rightarrow (x/w, y/w, z/w)$$



Базовая геометрия

- Точка x лежит на линии l тогда и только тогда $x^T l = 0$
- Пересечение двух двух линий можно описать как векторное произведение двух линий $x = l_i \times l_j$

- Пример: линии $x = 1, y=1$

$$\begin{aligned} -1x + 1 = 0 & \quad (-1, 0, 1)^T & -1y + 1 = 0 & \quad (0, -1, 1)^T \\ x = l_1 \times l_2 = \begin{vmatrix} i & j & k \\ -1 & 0 & 1 \\ 0 & -1 & 0 \end{vmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \end{aligned}$$

- Линию можно задать как векторное произведение двух точек

$$line_{ij} = x_i \times x_j$$

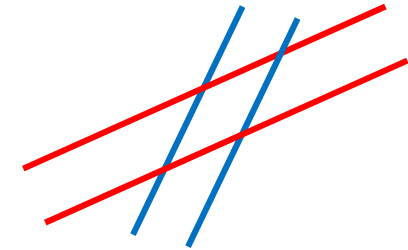


Идеальные точки

- Рассмотрим параллельные линии:

$$ax + by + c = 0$$

$$ax + by + c' = 0$$



- Точка пересечения будет: $(b, -a, 0)^T$
- Точки вида $(x, y, 0)^T$ - идеальные точки в «бесконечности»
- Пример: $x=1$, $x=2$

$$x = l_1 \times l_2 = \begin{vmatrix} i & j & k \\ -1 & 0 & 1 \\ -1 & 0 & 2 \end{vmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

Точка в бесконечности, в направлении у-оси



Проективное преобразование

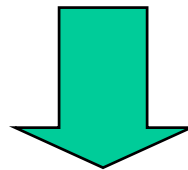
- *Проективность (projectivity)* – это такое обратимое преобразование h из P^2 , при котором точки x_1, x_2, x_3 лежат на одной прямой тогда и только тогда, когда $h(x_1), h(x_2), h(x_3)$ – лежат на одной прямой
- Отображение h в P^2 – проективность тогда и только тогда, когда существует невырожденная матрица H размером 3×3 , когда для любой точки x верно $h(x) = Hx$

$$\begin{bmatrix} x'_1 \\ x'_2 \\ x'_3 \end{bmatrix} \cong \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{12} & h_{33} \end{bmatrix} * \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$



Аффинные преобразования

$$X \rightarrow aX + T_x$$



$$\begin{bmatrix} aX + T_x \\ bY + T_y \\ cZ + T_z \\ 1 \end{bmatrix} \approx \begin{bmatrix} a & 0 & 0 & T_x \\ 0 & b & 0 & T_y \\ 0 & 0 & c & T_z \\ 0 & 0 & 0 & 1 \end{bmatrix} * \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$



P в матричной форме

Запишем теперь перспективную проекцию в матричном виде:

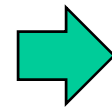
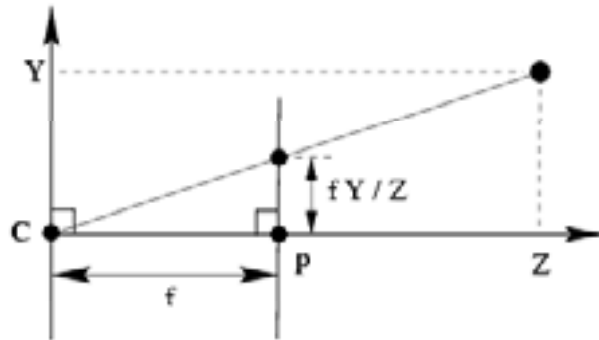
$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \cong \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} * \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} X/Z \\ Y/Z \\ 1 \end{bmatrix} \quad \begin{aligned} x &= \frac{X}{Z} \\ y &= \frac{Y}{Z} \end{aligned}$$

$$P = [I \mid 0]$$

$$\lambda x = PX$$



Внутренняя калибровка (f)



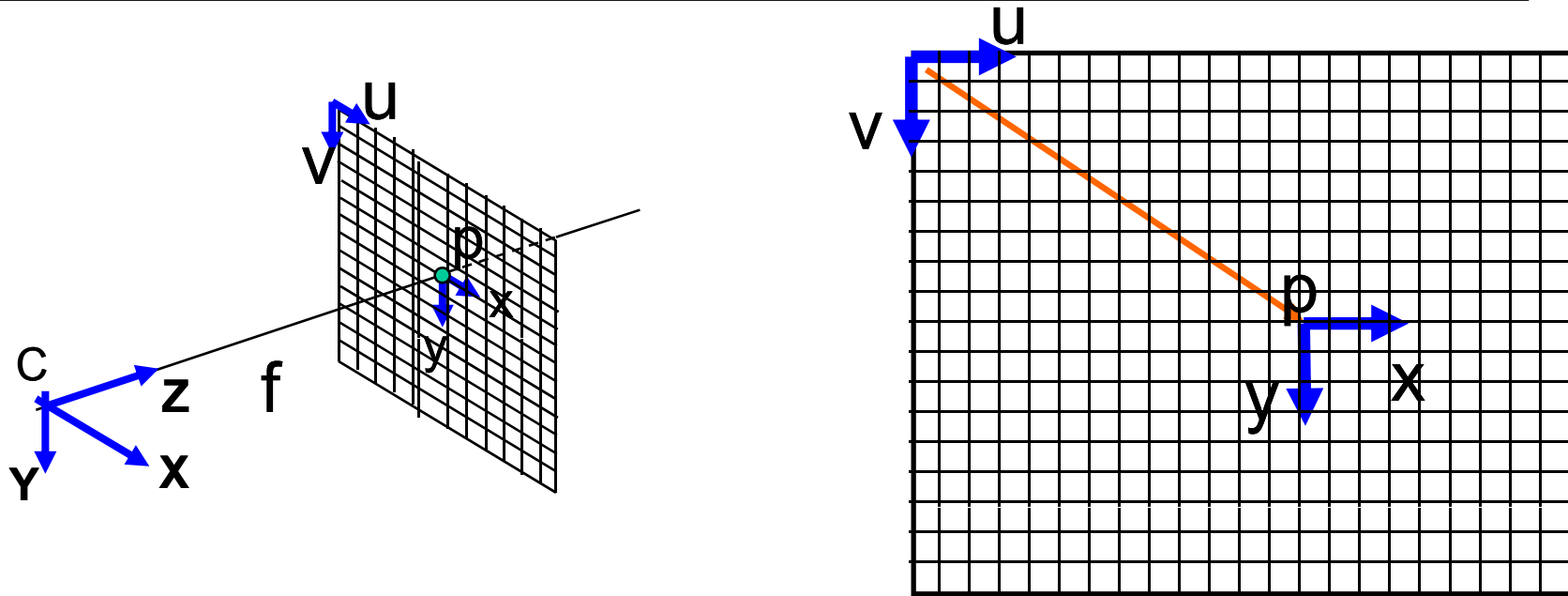
$$x = f \frac{X}{Z} \quad y = f \frac{Y}{Z}$$



$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \cong \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & f & 0 \end{bmatrix} * \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \Rightarrow P = \underbrace{\begin{bmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{bmatrix}}_K \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}}_{[I|0]}$$



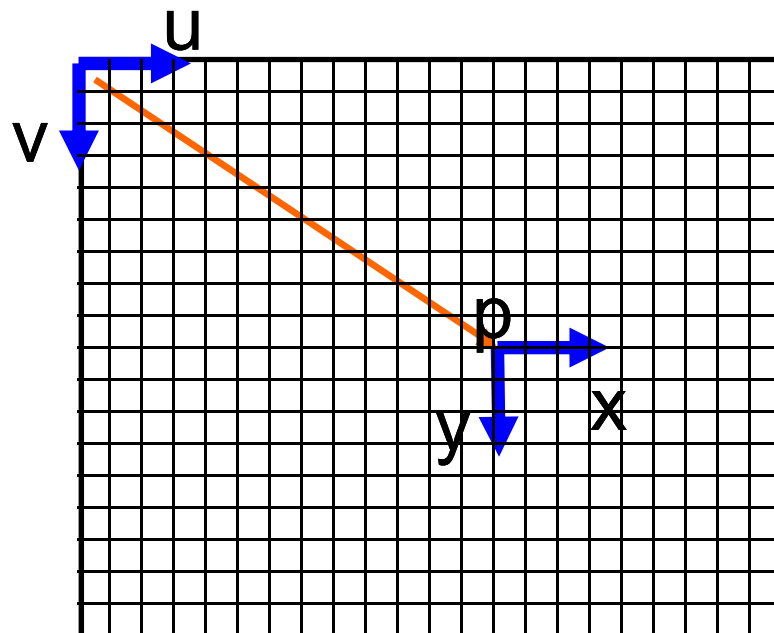
Внутренняя калибровка



- p – основание перпендикуляра из C на картинную плоскость (принципиальная точка)
- $p = (0, 0, f)$ в мировых координатах, центр системы координат в картинной плоскости
- В картинной области расположена матрица (изображение)
- Нужно преобразовать из мировых координат в пиксельные



Внутренняя калибровка



- Для перевода в координаты изображения в пикселях нужно сделать масштабирование и сдвиг

$$(x, y) \rightarrow (sx + c_x, sy + c_y) \cong (x / pix + c_x, y / pix + c_y)$$

Где pix – размер пикселя, (c_x, c_y) – принципиальная точка в координатах изображения



Внутренняя калибровка

$$(x, y) \rightarrow (sx + c_x, sy + c_y) \cong (x / pix + c_x, y / pix + c_y)$$

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \cong \begin{bmatrix} 1 / pix & 0 & c_x \\ 0 & 1 / pix & c_y \\ 0 & 0 & 1 \end{bmatrix} * \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} x / pix + c_x \\ y / pix + c_y \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 / pix & 0 & c_x \\ 0 & 1 / pix & c_y \\ 0 & 0 & 1 \end{bmatrix} * \begin{bmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{bmatrix} \cong \begin{bmatrix} f / pix & 0 & c_x \\ 0 & f / pix & c_y \\ 0 & 0 & 1 \end{bmatrix}$$

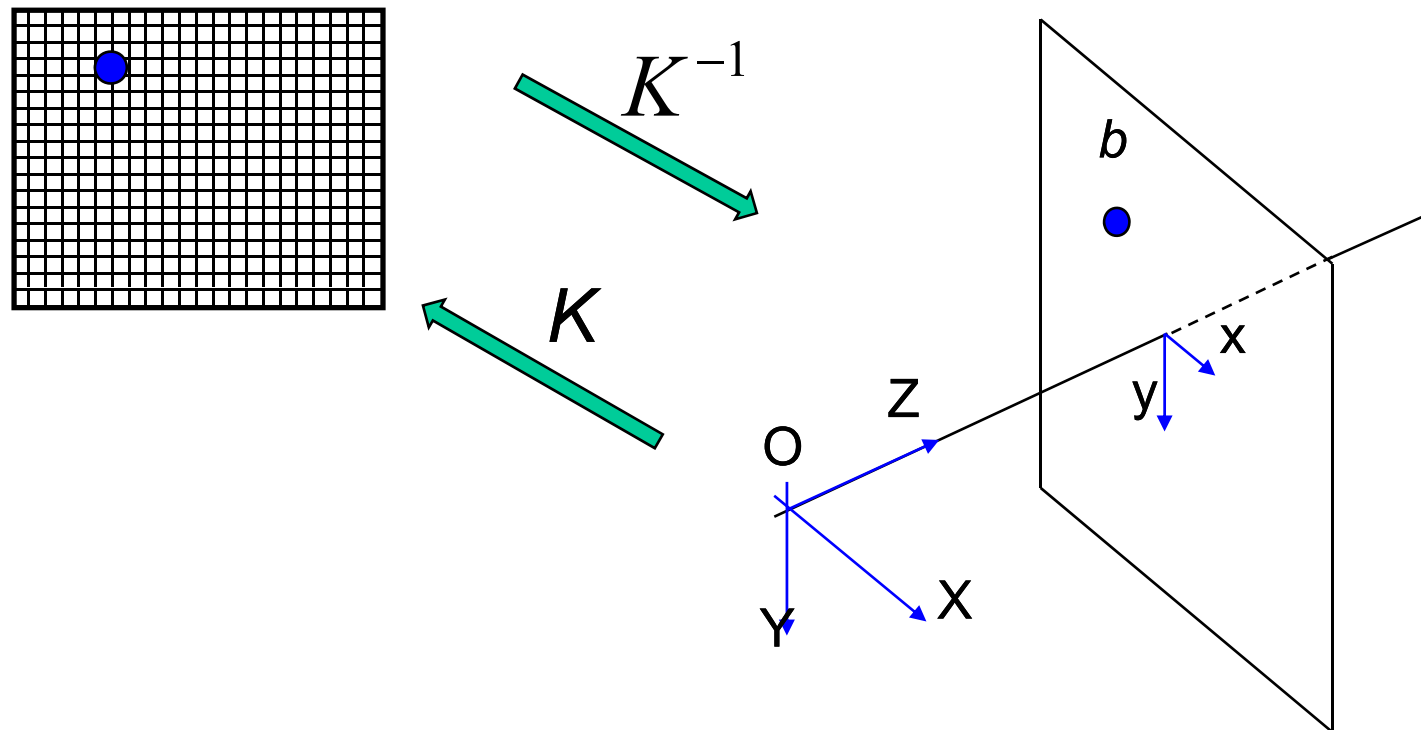


Внутренняя калибровка

Или матрица калибровки камеры (camera calibration matrix), internal calibration

$$K = \begin{bmatrix} a_x & s & c_x \\ 0 & a_y & c_y \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} f / pix & s & c_x \\ 0 & f / pix & c_y \\ 0 & 0 & 1 \end{bmatrix}$$

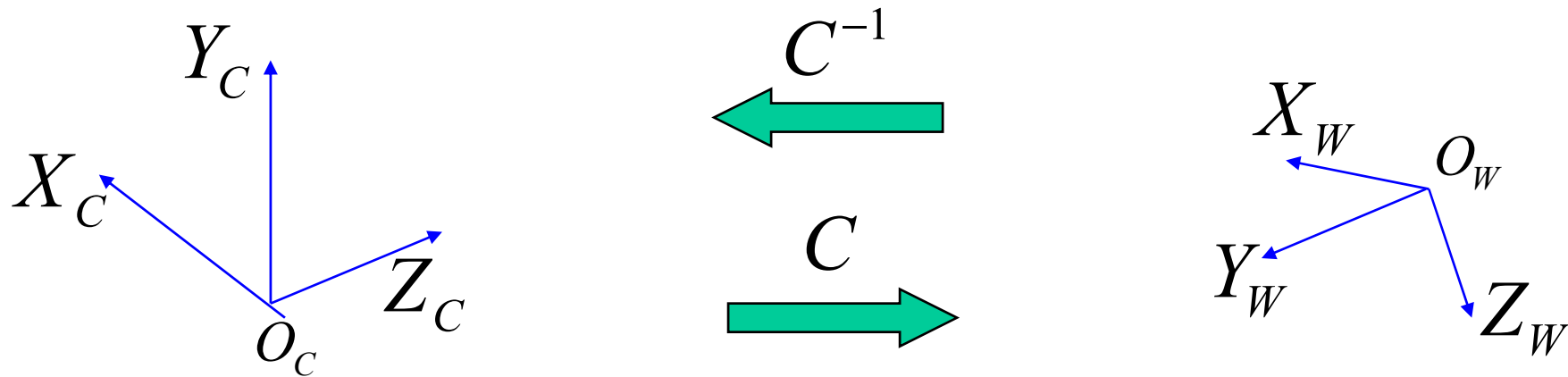
- a – масштабирующие коэффициенты
- s – наклон, угол между осями x и y изображения
- c – положение принципиальной точки изображения (в пикселях)
- f , pix – фокусное расстояние, размер пиксела



- Внутренняя калибровка K отображает точки с картинной плоскости на изображение
- Обратная матрица K^{-1} отображает точки с изображения на картинную плоскость, с $f=1$
 - Координаты в пространстве с центром в C будут $(x, y, 1, 1)$
 - Отображение точек на изображении в лучи



Из мировой в координаты камеры



$$C = \begin{bmatrix} R & T \\ [0,0,0] & 1 \end{bmatrix}$$

- Положение и ориентация камеры в мировых координатах задается евклидовым преобразованием C
- Обратное преобразование
 - Из мировых координат в координаты камеры
 - Обратное к C преобразование! – $\text{Inverse}(C)$



Внешняя калибровка

- Матрица преобразования из мировой системы координат в систему координат камеры называется матрицей *внешней калибровки*

$$C^{-1} = \left(\begin{bmatrix} R & T \\ [0,0,0] & 1 \end{bmatrix} \right)^{-1} = \begin{bmatrix} R^T & -R^T T \\ [0,0,0] & 1 \end{bmatrix}$$

- Внешняя калибровка определяется положением и ориентацией камеры в пространстве



Конечная проективная камера

$$P = K \cdot [I \mid 0] \cdot C^{-1}$$

↑
Матрица
проекции
камеры

↑
Внутренняя
калибровка

↑
Центральная
проекция

↑
Внешняя
калибровка

- Матрица P размером 3×4 , которую можно факторизовать вышеуказанным образом, где K – верхнетреугольная, называется *конечной проективной камерой (finite projective camera)*
- У неё 11 степеней свободы (5 внутренних параметров, 3 параметра поворота и 3 параметра положения центра проекции)



Обобщенная проективная камера

- Произвольная однородная матрица ранга 3 называется *обобщенной проективной камерой* (*general projective camera*)

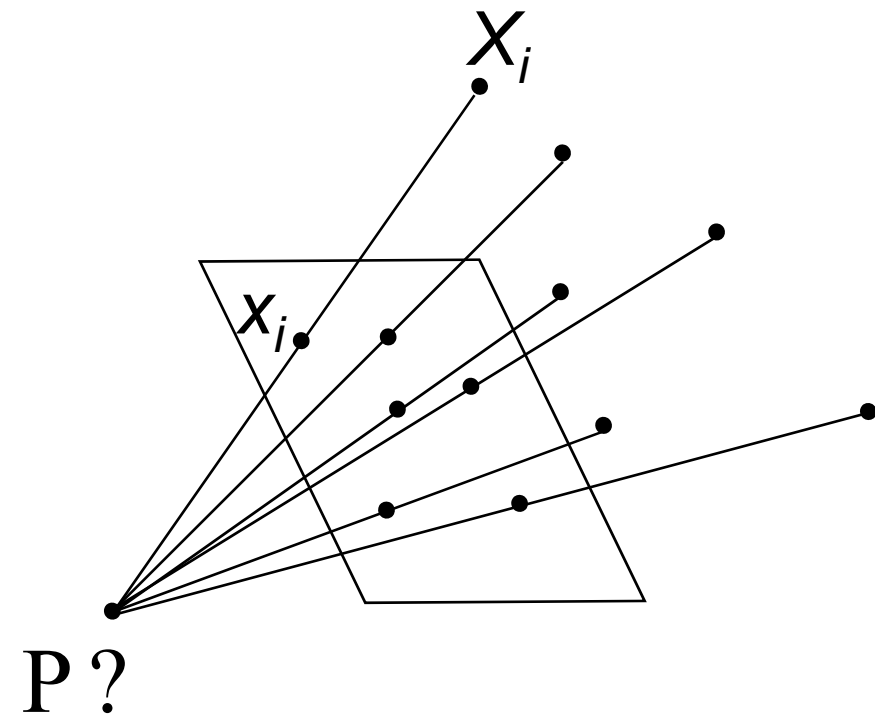
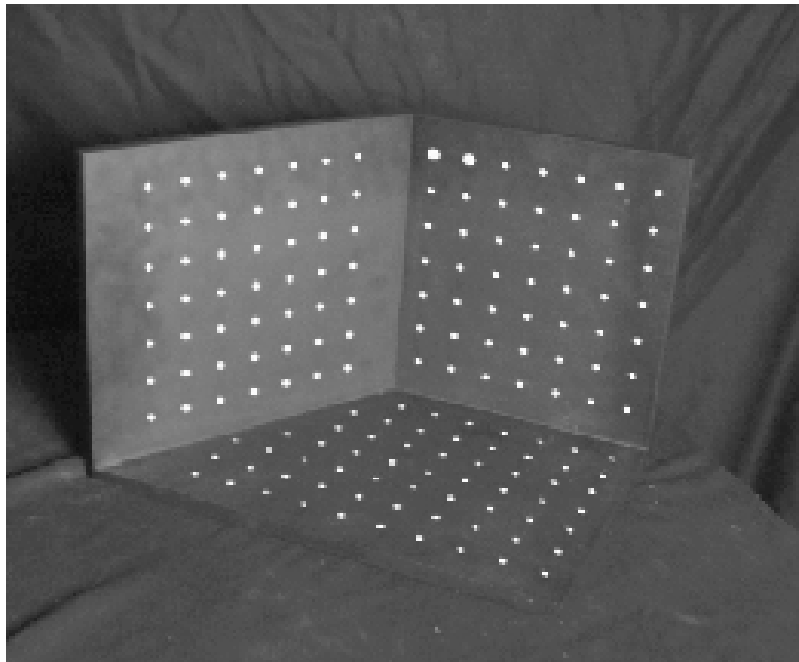
$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \cong \begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \end{bmatrix} * \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad \lambda x = PX$$

- У P в общем случае также 11 степеней свободы (неизвестных параметров)
- Если ранг < 3 , тогда всё проецируется в линию или точку, а не плоскость, т.е. изображения не получается



Калибровка камеры

- Даны n точек с известными 3D координатами X_i и известными проекциями x_i , оценить параметры камеры





DLT-метод

- Direct Linear Transformation (прямое линейное преобразование)

$$\lambda \mathbf{x}_i = \mathbf{P} \mathbf{X}_i \qquad \lambda \begin{bmatrix} x_i \\ y_i \\ w_i \end{bmatrix} = \begin{bmatrix} \mathbf{P}_1^T \\ \mathbf{P}_2^T \\ \mathbf{P}_3^T \end{bmatrix} \mathbf{X}_i \qquad \mathbf{x}_i \times \mathbf{P} \mathbf{X}_i = 0$$

$$\begin{bmatrix} 0 & -w_i \mathbf{X}_i^T & y_i \mathbf{X}_i^T \\ w_i \mathbf{X}_i^T & 0 & -x_i \mathbf{X}_i^T \\ -y_i \mathbf{X}_i^T & x_i \mathbf{X}_i^T & 0 \end{bmatrix} \begin{pmatrix} \mathbf{P}_1 \\ \mathbf{P}_2 \\ \mathbf{P}_3 \end{pmatrix} = 0$$

- Два линейно независимых уравнения (берут обычно первые два)



Калибровка камеры

$$\begin{bmatrix} 0^T & X_1^T & -y_1 X_1^T \\ X_1^T & 0^T & -x_1 X_1^T \\ \dots & \dots & \dots \\ 0^T & X_n^T & -y_n X_n^T \\ X_n^T & 0^T & -x_n X_n^T \end{bmatrix} \begin{pmatrix} P_1 \\ P_2 \\ P_3 \end{pmatrix} = 0 \quad \begin{matrix} Ap = 0 \\ \|p\| = 1 \end{matrix}$$

- У P 11 степеней свободы (12 параметров, но масштаб произвольный)
- Каждая пара (x,X) даёт нам уравнения
- 6 соответствий необходимо для решения
- Однородные наименьшие квадраты



- Singular Value Decomposition: $A = UDV^T$, где
 - U, V - ортогональные
 - D – диагональная (м.б. упорядочены столбцы по убыванию элементов)
 - D – состоит из сингулярных чисел
- Одно из самых важных разложений
- Связь с собственными значениями:

$$A = UDV^T$$

$$A^T A = VDU^T UDV^T$$

$$A^T A = VDDV^T = VD^2V^T$$

- Сингулярные значения – корни квадратные собственных значений матрицы $A^T A$



Наименьшие квадраты

- Имеем уравнение $Ap = 0$ при $\|p\| = 1$
- Решим с помощью SVD-разложения: $A = UDV^T$
- Нужно минимизировать $\|UDV^T p\|$
- Однако: $\|UDV^T p\| = \|DV^T p\|$ $\|V^T p\| = \|p\|$
- Тогда нам нужно минимизировать $\|DV^T p\|$ при $\|V^T p\| = 1$
- Обозначим $y = V^T p$
- Тогда нам нужно минимизировать $\|Dy\|$ при $\|y\| = 1$
- $y = (0, \dots, 0, 1)^T$, тогда $x = Vy$ – последний столбец V



Калибровка камеры

$$\begin{bmatrix} 0^T & X_1^T & -y_1 X_1^T \\ X_1^T & 0^T & -x_1 X_1^T \\ \dots & \dots & \dots \\ 0^T & X_n^T & -y_n X_n^T \\ X_n^T & 0^T & -x_n X_n^T \end{bmatrix} \begin{pmatrix} P_1 \\ P_2 \\ P_3 \end{pmatrix} = 0 \quad \begin{aligned} Ap &= 0 \\ \|p\| &= 1 \end{aligned}$$

- Для копланарных точек, удовлетворяющих уравнению $\Pi^T X = 0$, получим вырожденные решения $(\Pi, 0, 0)$, $(0, \Pi, 0)$, или $(0, 0, \Pi)$



Калибровка камеры

- Когда мы получили матрицу проекции, нам необходимо извлечь из неё внутренние и внешние параметры калибровки
- Представим матрицу проецирования в следующем виде:

$$P = KR^T [I \mid -T] = [M \mid -MT]$$

- Для разложения M в KR^T воспользуемся RQ-факторизацией
 - Q – ортогональная матрица
 - K – верхнетреугольная
- Потом найдем T



“Gold Standard”

- Метод «Золотого стандарта» – метод, оптимизирующий параметры по наиболее корректной, оптимальной метрике
- В случае калибровки камеры, это расстояния от проекций 3D точек до их измеренных значений (ошибка репроекции):

$$\min_P \sum_i d^2(x_i, P X_i)$$

- Требуется использование нелинейных методов оптимизации (градиентный спуск и т.д.)



Схема калибровки камеры

- DLT-метод для получения матрицы проецирования
- RQ-факторизация для извлечения матрицы внутренней калибровки K и внешней калибровки R и T
- Уточнение параметров калибровки с помощью нелинейной оптимизации ошибки репроекции:

$$\min_P \sum_i d^2(x_i, P X_i)$$

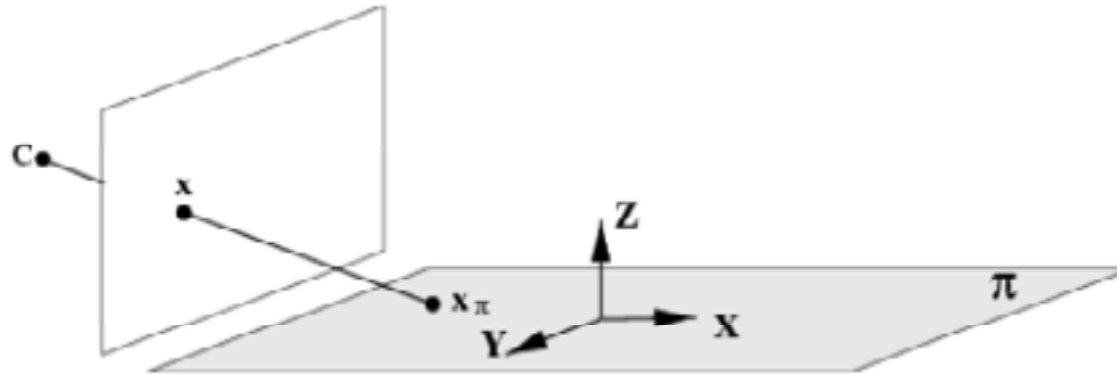


План лекции

- Устройство фотоаппарата
- Модель перспективной проекции
- **Гомография**
- Геометрия 2х камер



Гомография



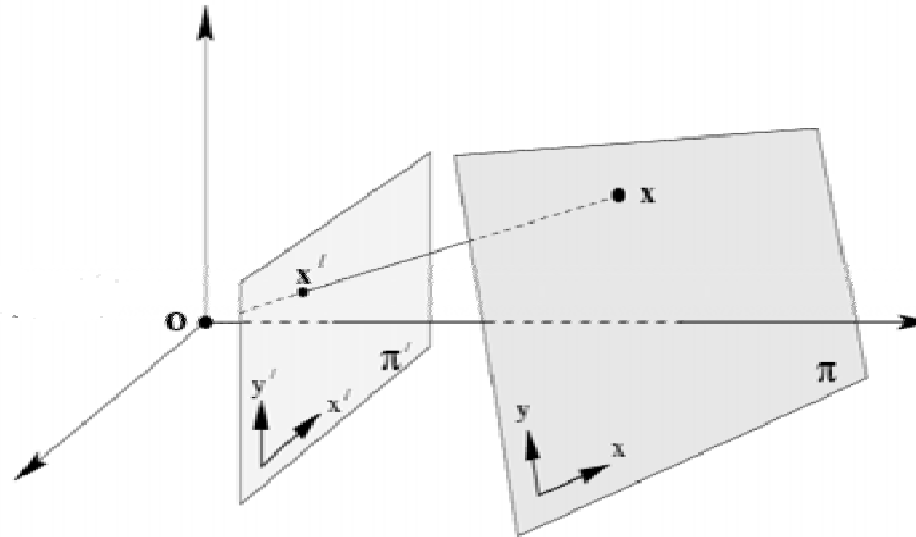
- Перспективное преобразование плоскости (гомография)
- Пусть точка X лежит на плоскости
- Выберем специальную мировую систему координат

$$\begin{bmatrix} u \\ v \\ w \end{bmatrix} \cong \begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \end{bmatrix} * \begin{bmatrix} X \\ Y \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} p_{11} & p_{12} & p_{14} \\ p_{21} & p_{22} & p_{24} \\ p_{31} & p_{32} & p_{34} \end{bmatrix} * \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix}$$



Гомография

- Для произвольной плоскости аналогично (домножим на матрицу евклидова преобразования)
- 8 степеней свободы



$$\begin{bmatrix} u \\ v \\ w \end{bmatrix} \cong \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} * \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$



Вычисление гомографии (DLT)

- (x, x') – известные соответствия на 2х изображениях
- Линейный (DLT) метод:
- Составим систему линейных уравнений $\lambda x' = Hx$
 - Каждое соответствие даст 2 уравнения на параметры H
 - 4 соответствия дадут 8 уравнений на 8 неизвестных
 - Никакие 3 не должны лежать на одной прямой!
 - Можем взять и больше соответствий...
- Сведем к однородной задаче наименьших квадратов
 - Уже знаем, как решать!



Вычисление гомографии (Gold)

- Критерий получше (transfer error)

$$\sum_i d^2(x'_i, Hx_i) + d^2(H^{-1}x'_i, x_i)$$

- Вычислить достаточно просто

- Оптимальный критерий (distance error)

$$\sum_i d^2(x'_i, \hat{x}'_i) + d^2(x_i, \hat{x}_i)$$

где $\hat{x}'_i = H\hat{x}_i$ - идеальное соответствие

- Нужно вычислять «идеальные соответствия», это сложно



Использование гомографии

- 4 соответствия (x, x') , из которых никакие 3 не лежат на одной прямой, определяют гомографию H
- Значит, любые 4 точки в общей позиции можно преобразовать в любые 4 другие точки в общей позиции
- Это можно использовать



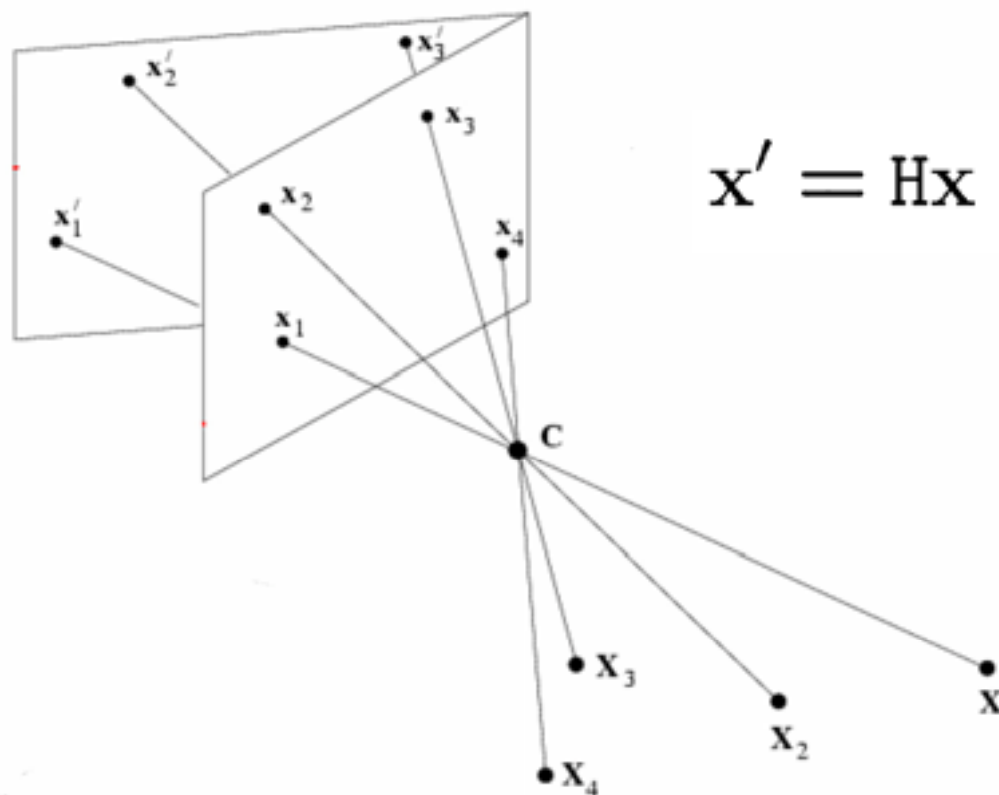
Ректификация изображения



- Выбираем 4 точки, которые являются вершинами прямоугольника с известными размерами
- Вычисляем гомографию (4е точки -> прямоугольник)
- Разворачиваем изображение
- Можем измерять расстояния на плоскости!



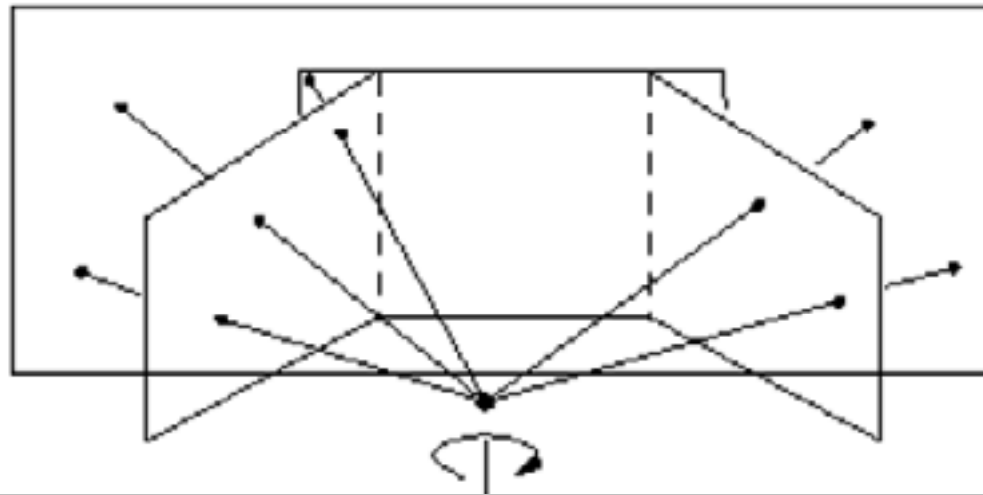
Стационарная камера



Если центр камеры не изменился, то одно изображение в другое преобразуется гомографией



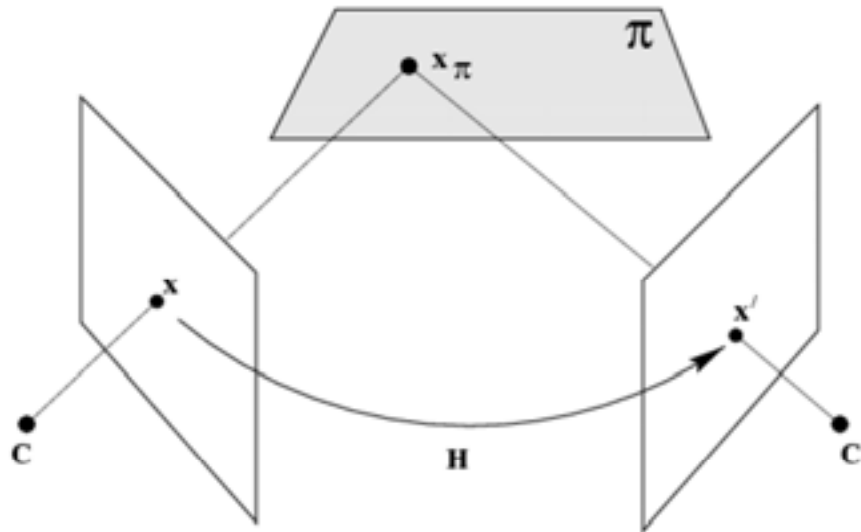
Построение мозаики (панорамы)



Проецируем все изображения на общую плоскость



Плоская сцена



$$x = H_{1\pi} x_\pi \quad x' = H_{2\pi} x_\pi$$

$$x' = H_{2\pi} x_\pi$$

$$= H_{2\pi} H_{1\pi}^{-1} x = Hx$$

- Если наблюдаемая сцена – плоская, тогда преобразование точек между изображениями можно тоже описать гомографией
- Это результат используется в дальнейшем



Геометрия 2х камер

Microsoft
Research





План лекции

- Устройство фотоаппарата
- Модель перспективной проекции
- Гомография
- Геометрия двух камер



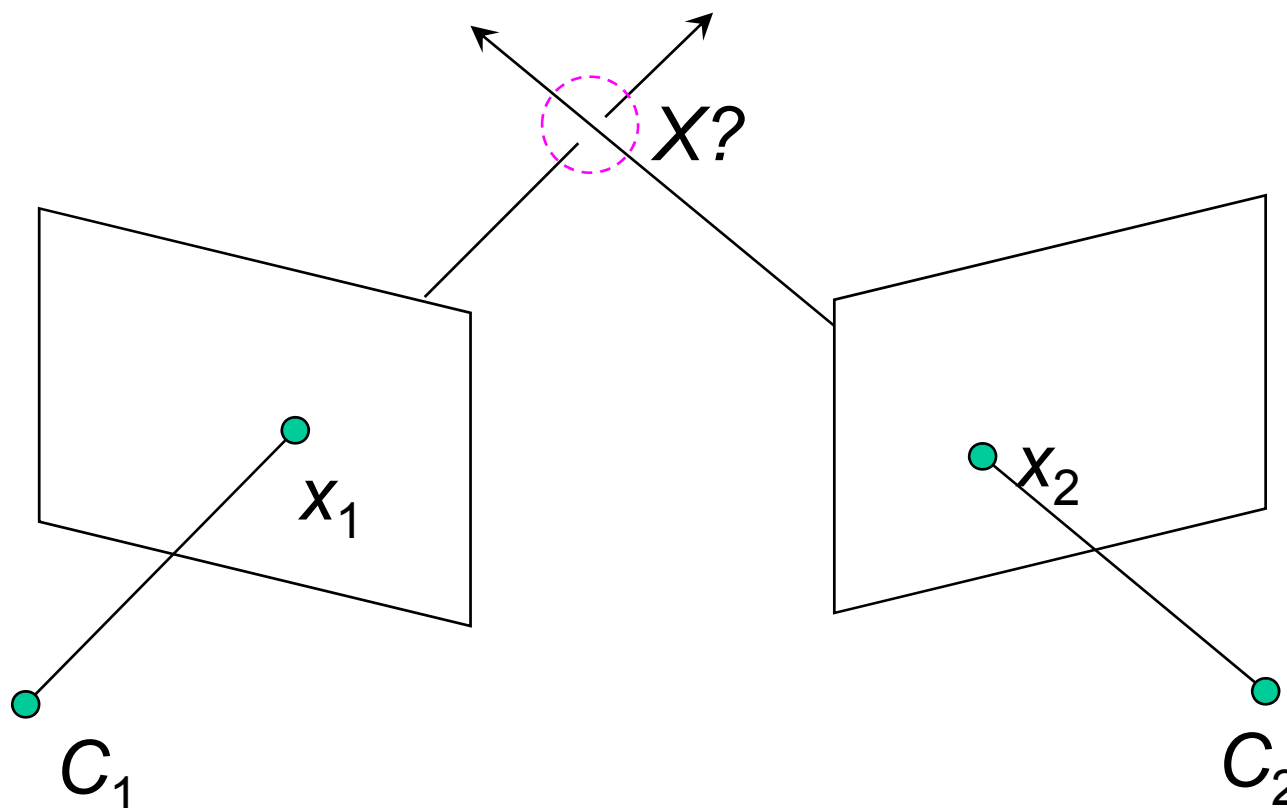
Задачи многовидовой геометрии

- **Геометрия сцены (структура)** : Пусть даны соответствующие точки в 2D на 2х и более изображениях, определить координаты соответствующих точек в 3D
- **Соответствие (стерео-сопоставление)** : Дана точка в одном изображении, как мы можем ограничить положение соответствующей точки на втором изображении?
- **Калибровка камер (движение)**: Дан набор соответствующих точек в 2х и более изображениях, определить матрицы калибровки камер для этих изображений (видов)



Триангуляция (геометрия сцены)

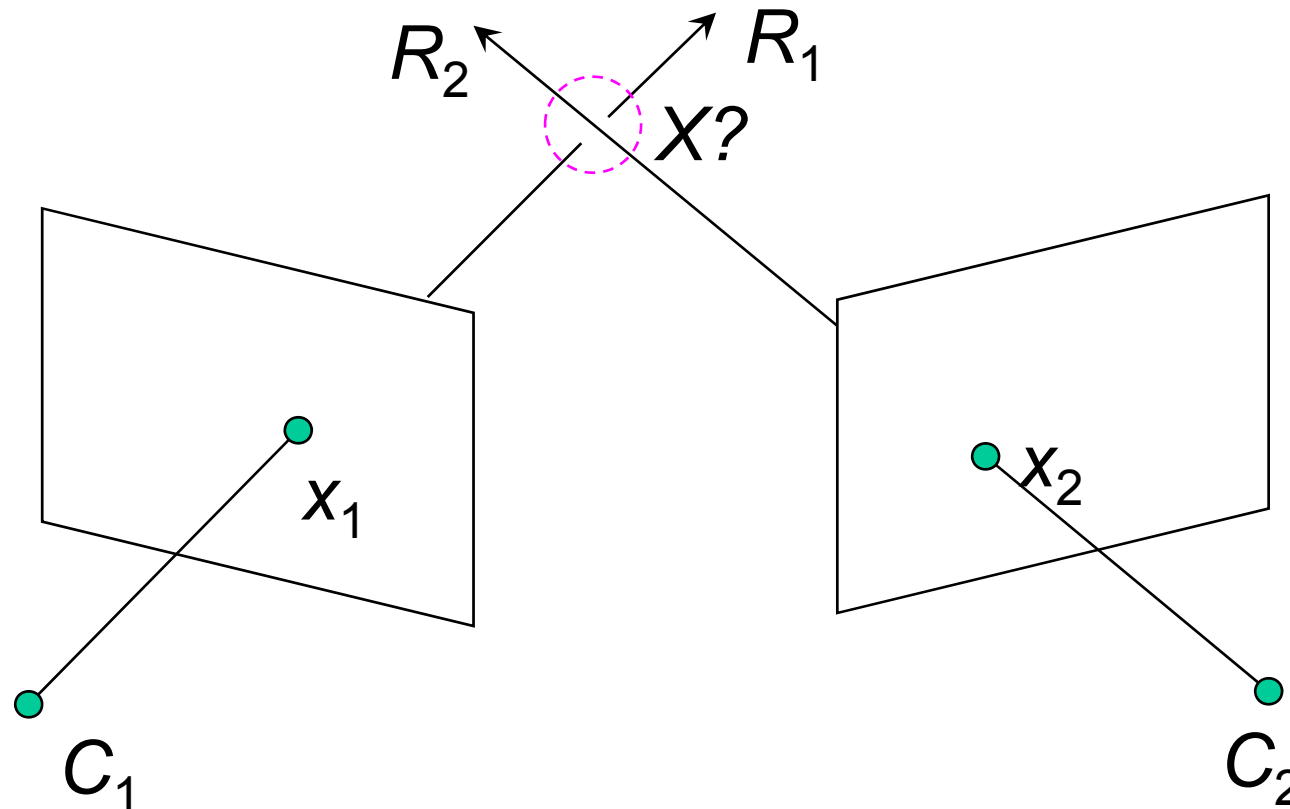
- Даны проекции x_1 , x_2 точки $\mathbf{X}=(X,Y,Z)$ на 2 или более изображения (с известными матрицами калибровки), найти координаты точки





Триангуляция

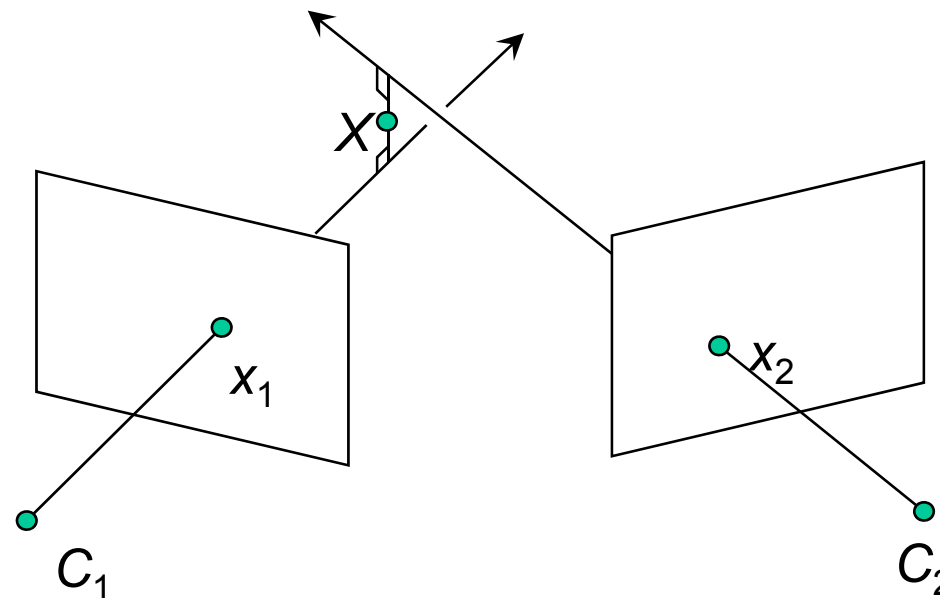
- Можно найти пересечение лучей, соответствующих точкам x_1 и x_2 , однако из-за ошибок измерения и оценки они не пересекаются





Геометрический подход

- Найдем кратчайший отрезок, соединяющий два луча и пусть X будет средней точкой на этом отрезке
- Алгоритм средней точки (midpoint)





Линейный подход

$$\lambda_1 x_1 = P_1 X \quad x_1 \times P_1 X = 0 \quad [x_{1 \times}] P_1 X = 0$$

$$\lambda_2 x_2 = P_2 X \quad x_2 \times P_2 X = 0 \quad [x_{2 \times}] P_2 X = 0$$

Проекция на каждую камеру даст
2 уравнения на X

$$x_1(p_1^{3T} X) - (p_1^{1T} X) = 0$$

$$y_2(p_1^{3T} X) - (p_1^{2T} X) = 0$$

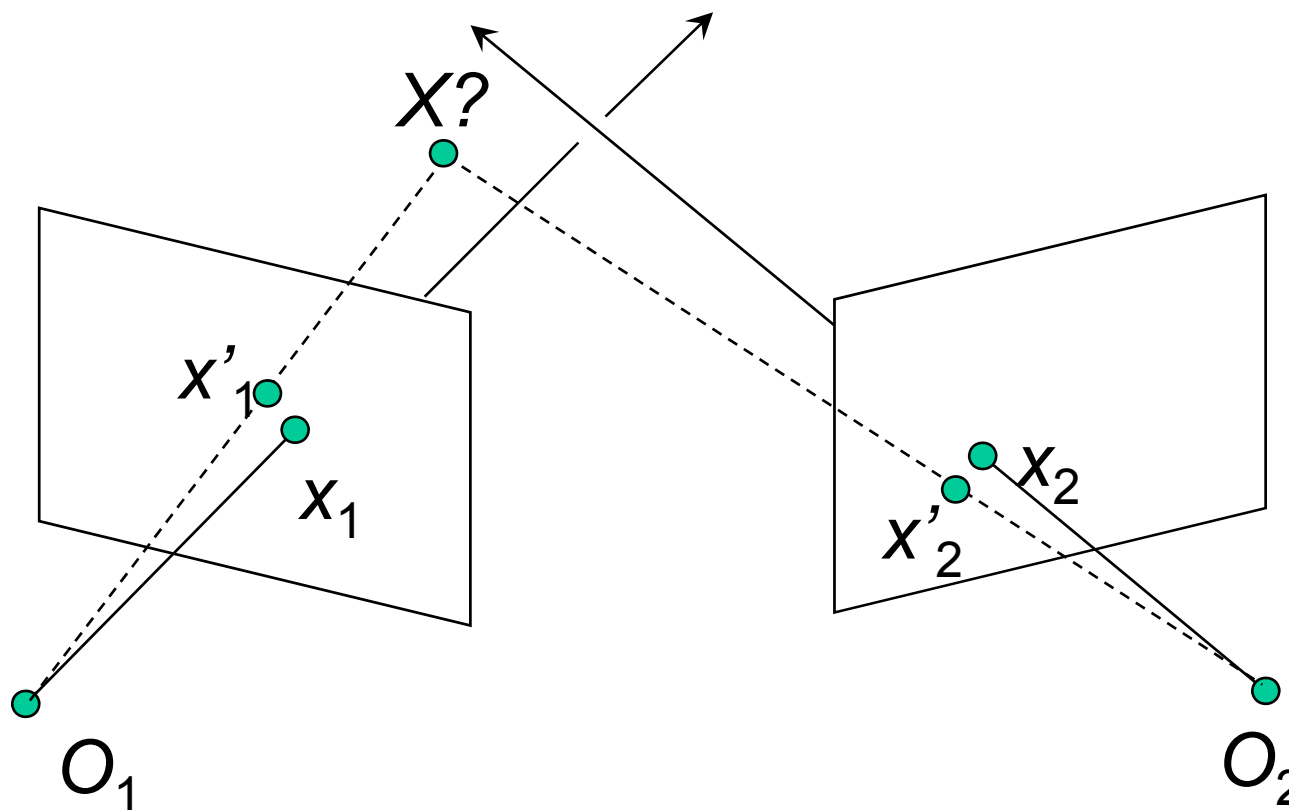
Проекция на обе камеры
приведет к знакомой задаче на
однородные наименьшие
квадраты

$$A = \begin{bmatrix} x_1 p_1^{3T} - p_1^{1T} \\ y_1 p_1^{3T} - p_1^{2T} \\ x_2 p_2^{3T} - p_2^{1T} \\ y_2 p_2^{3T} - p_2^{2T} \end{bmatrix} x = 0 \quad \begin{aligned} AX &= 0 \\ \|X\| &= 1 \end{aligned}$$



Нелинейный подход

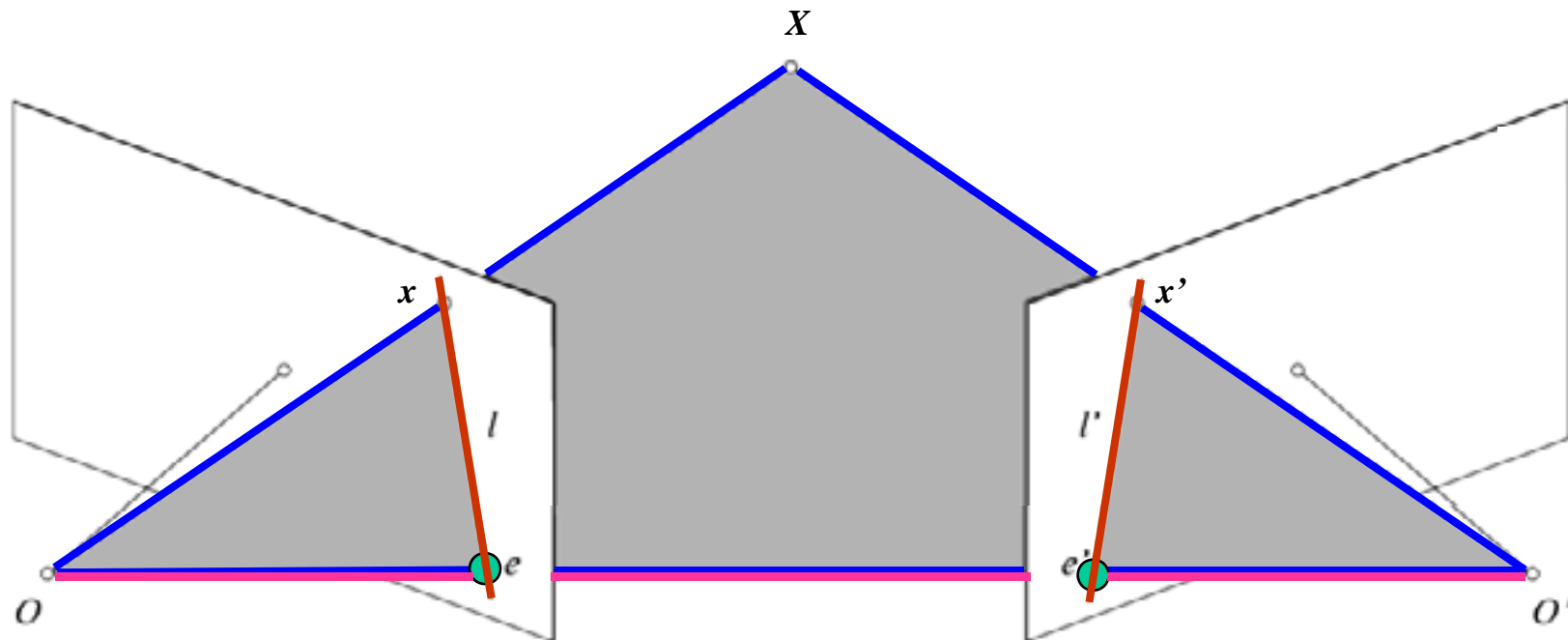
- Найти X , минимизирующий ошибку $d^2(x_1, P_1 X) + d^2(x_2, P_2 X)$



- Требуется применение нелинейной оптимизации
- Результат DLT используем как начальное приближение



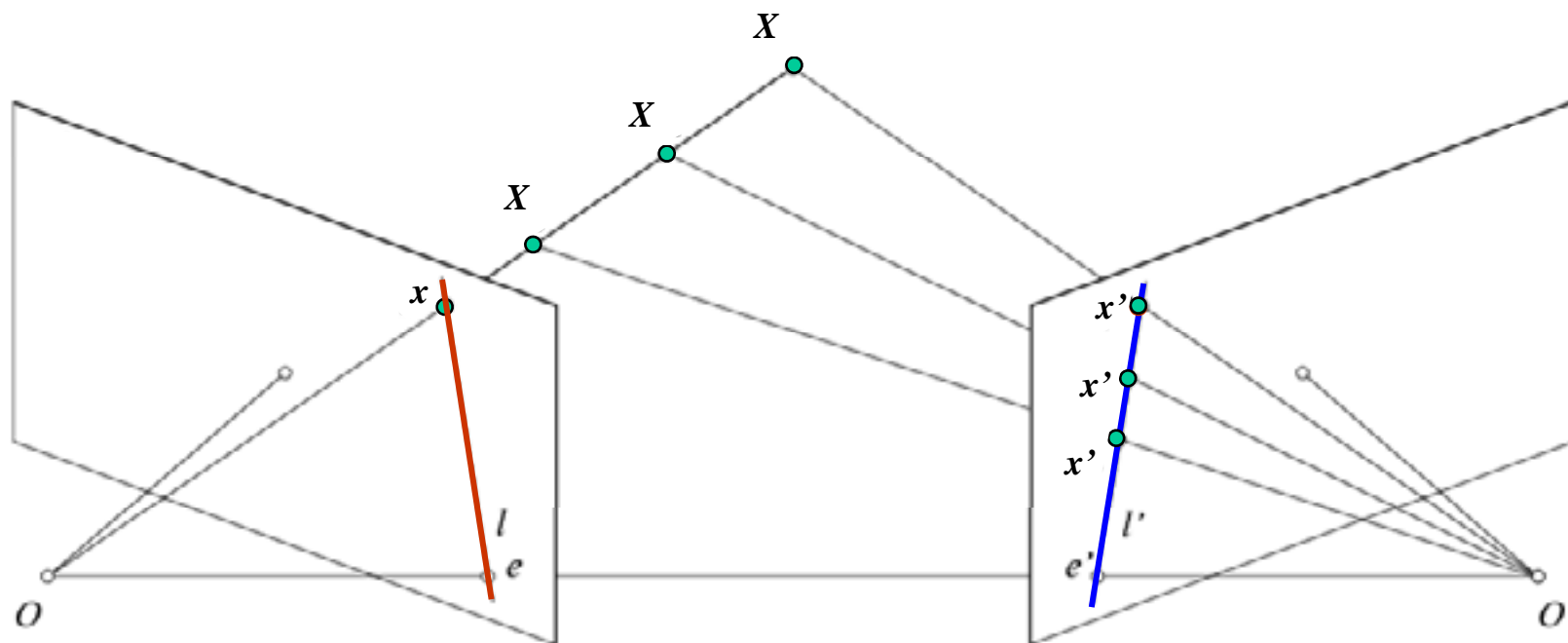
Эпилярная геометрия



- **Базовая линия (Baseline)** – линия, соединяющая центры камер
- **Эпилярная плоскость (Epipolar Plane)** – плоскость, проходящая через базовую линию (пучок плоскостей)
- **Эпиполи/Эпиляры (Epiholes)**
= пересечение базовой линии с картинными плоскостями
= проекции центра проекций второй камеры
- **Эпилярные линии (Epipolar Lines)** – пересечение эпилярной плоскости с картинной плоскостью (дают соответствующие пары)



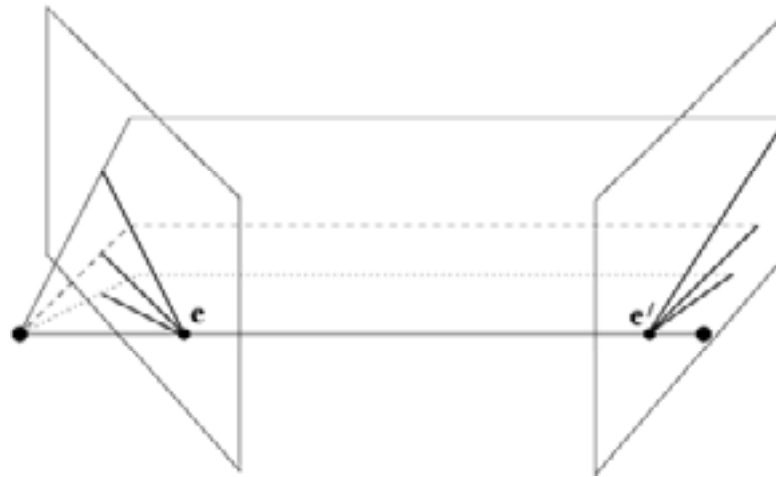
Эпиполярное ограничение



- Пусть даны изображения I и I' с центрами проекции O и O' , и проекция x точки X (неизвестной) на изображение I .
- Что мы можем сказать про точку x' – проекцию X на I' ?
 - x' должна лежать на эпиполярной линии l'
- Эпиполярное ограничение:
 - Проекции x и x' точки сцены X на изображения I и I' должны лежать на соответствующих эпиполярных линиях l и l'

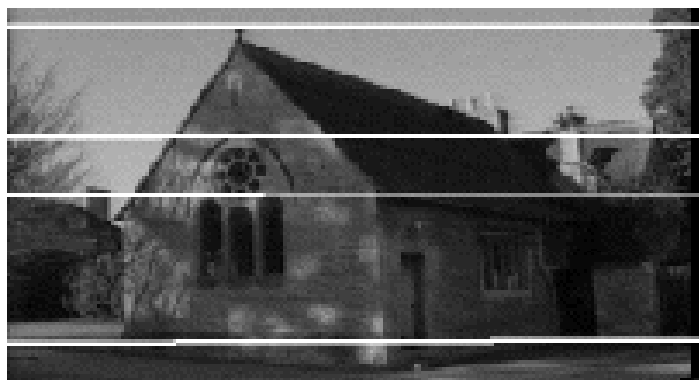


Пример: Сближающиеся камеры





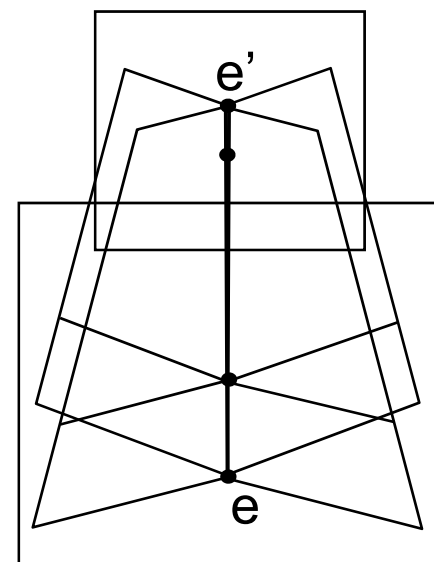
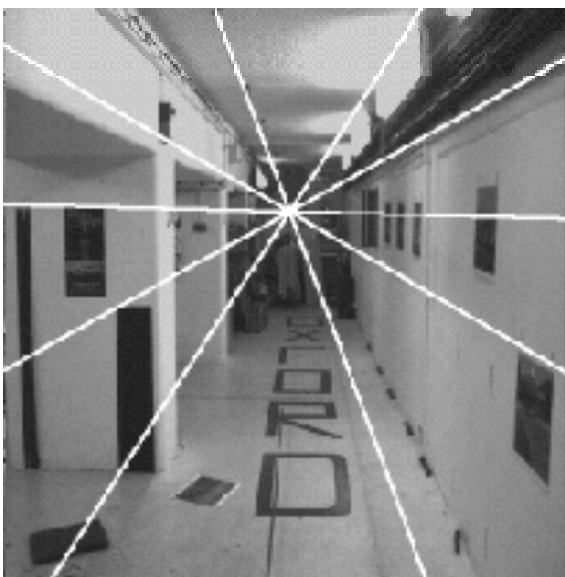
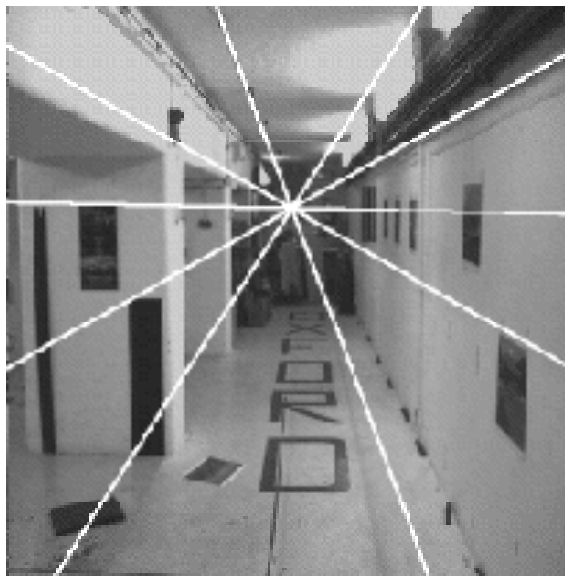
Пример: стереопара



Движение, параллельное картинной плоскости



Пример: движение вперед

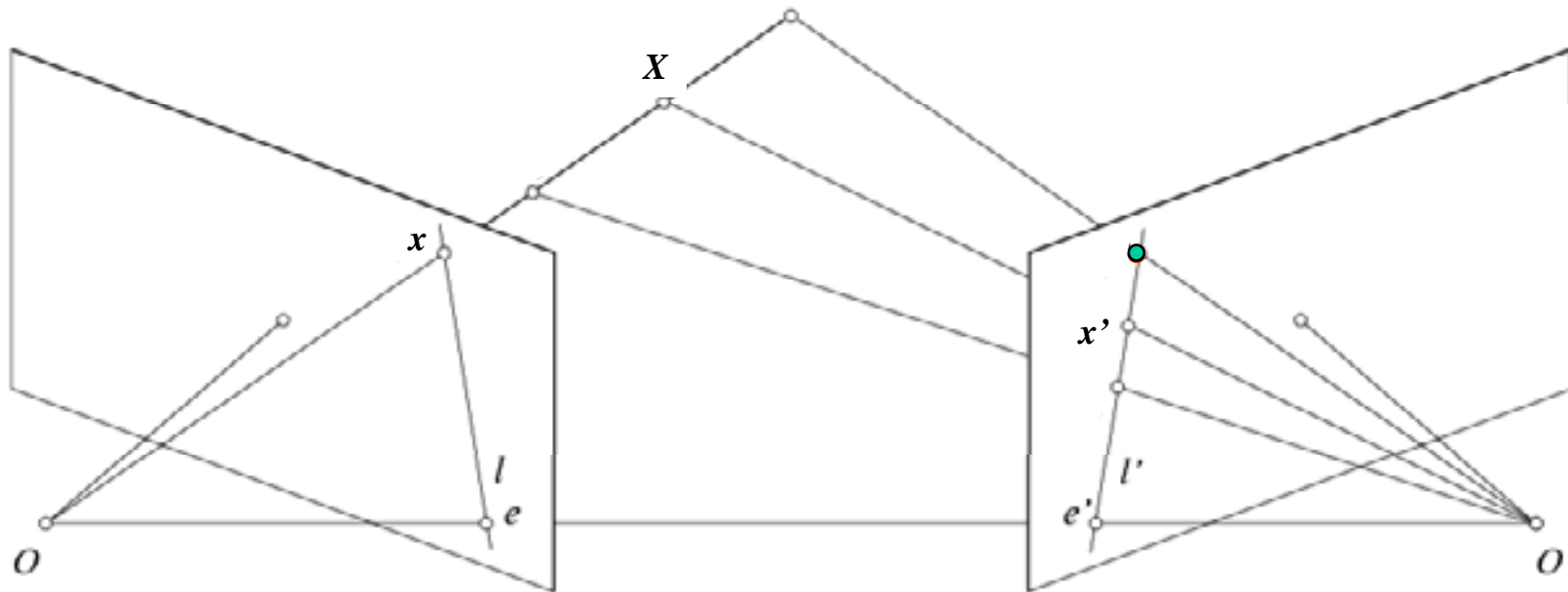


У эпиполей одинаковые координаты в обоих изображениях.

Точки движутся вдоль линий, исходящих из e – «фокуса расширения»



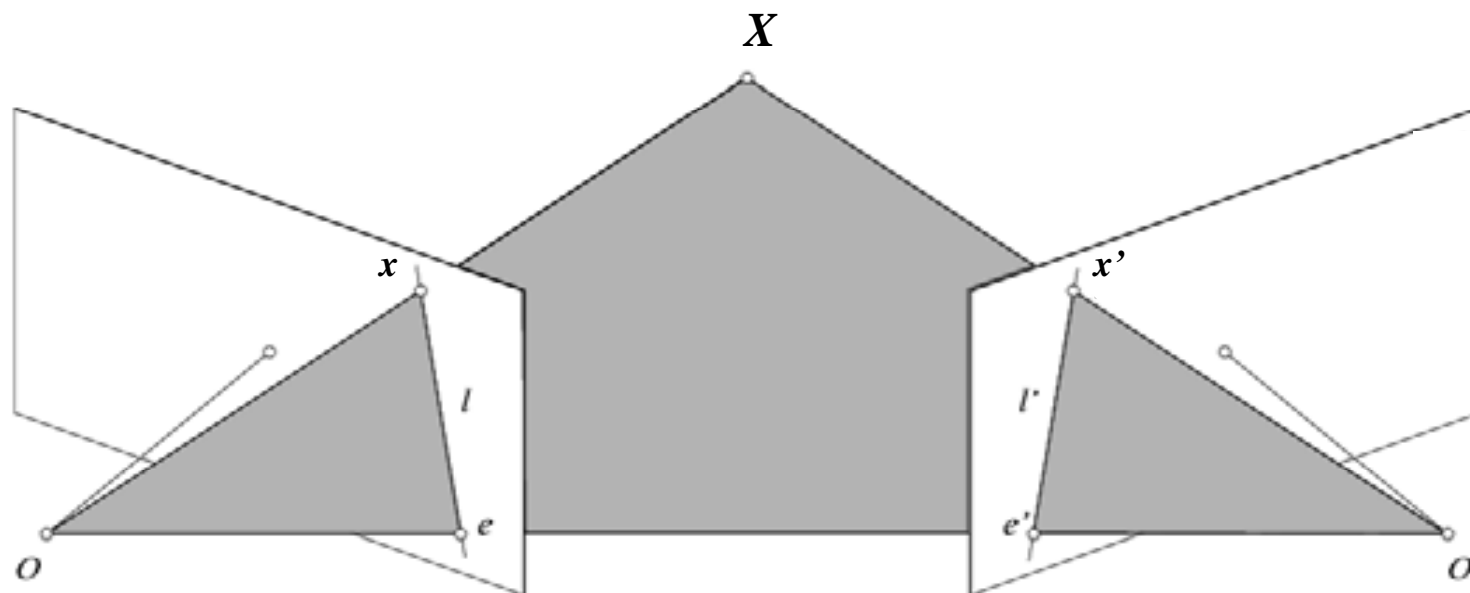
Эпиполярное ограничение



- Интересный вывод:
 - Линия l' не зависит от X , если x и калибровка камер известна
 - Линия l не зависит от X , если x' и калибровка камер известна
 - Между x , x' и калибровкой камер существует зависимость
 - Попробуем её вывести



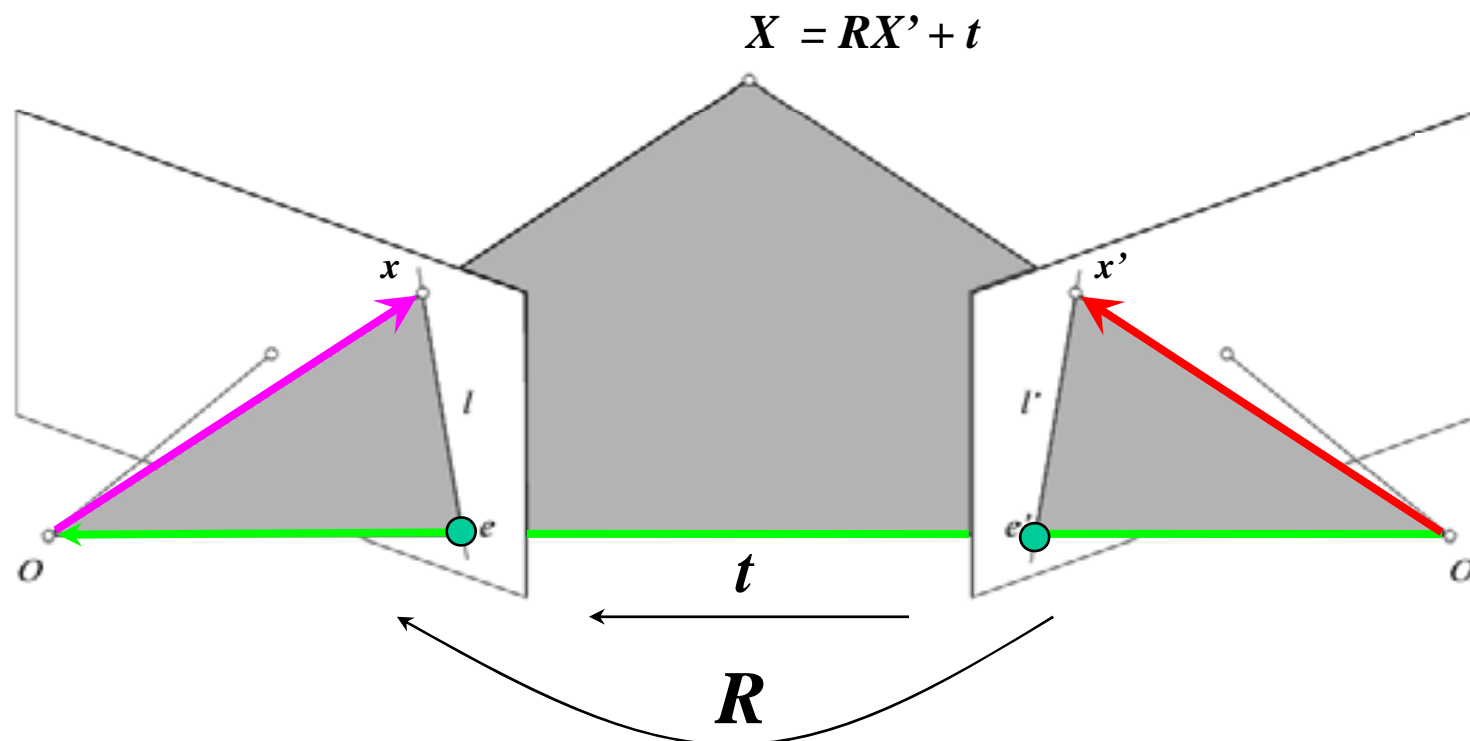
Калиброванный случай



- Совместим глобальную систему координат с координатной системой первой камеры
- Допустим, известны внутренние калибровки камер K , K' и внешняя калибровка второй камеры R и T
- Умножим матрицу проекции каждой камеры на обратную матрицу калибровки K^{-1} , K'^{-1}
 - Также отобразим точки x , x' с изображения на картинную плоскость



Калиброванный случай



Матрица камеры $[I|0]$

$$X = (u, v, w, 1)^T$$

$$x = (u, v, w)^T$$

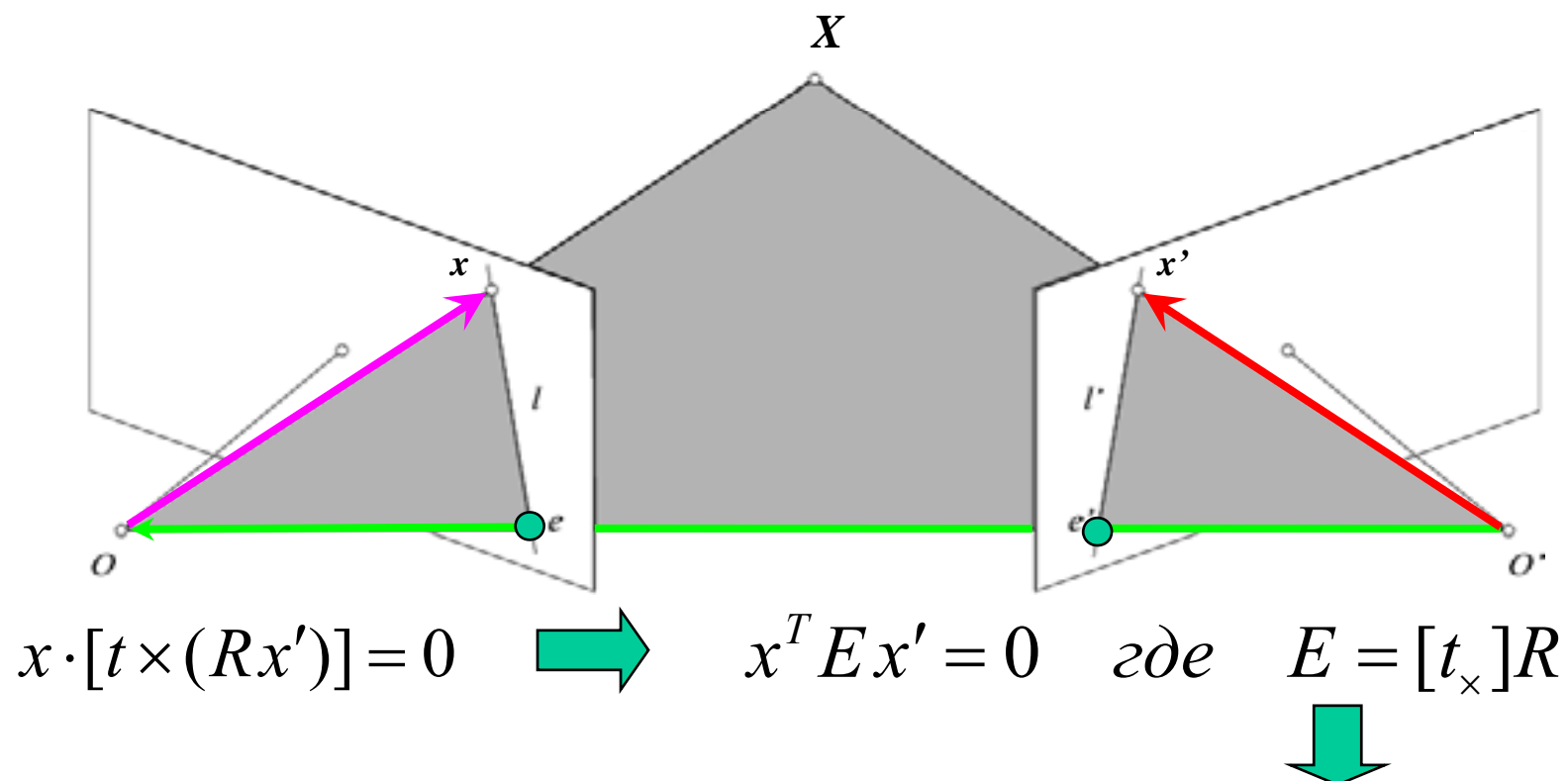
Матрица камеры: $[R^T | -R^T t]$

Вектор x' во второй системе координат имеет координаты Rx' в первой

Вектора x , t , и Rx' копланарны



Калиброванный случай

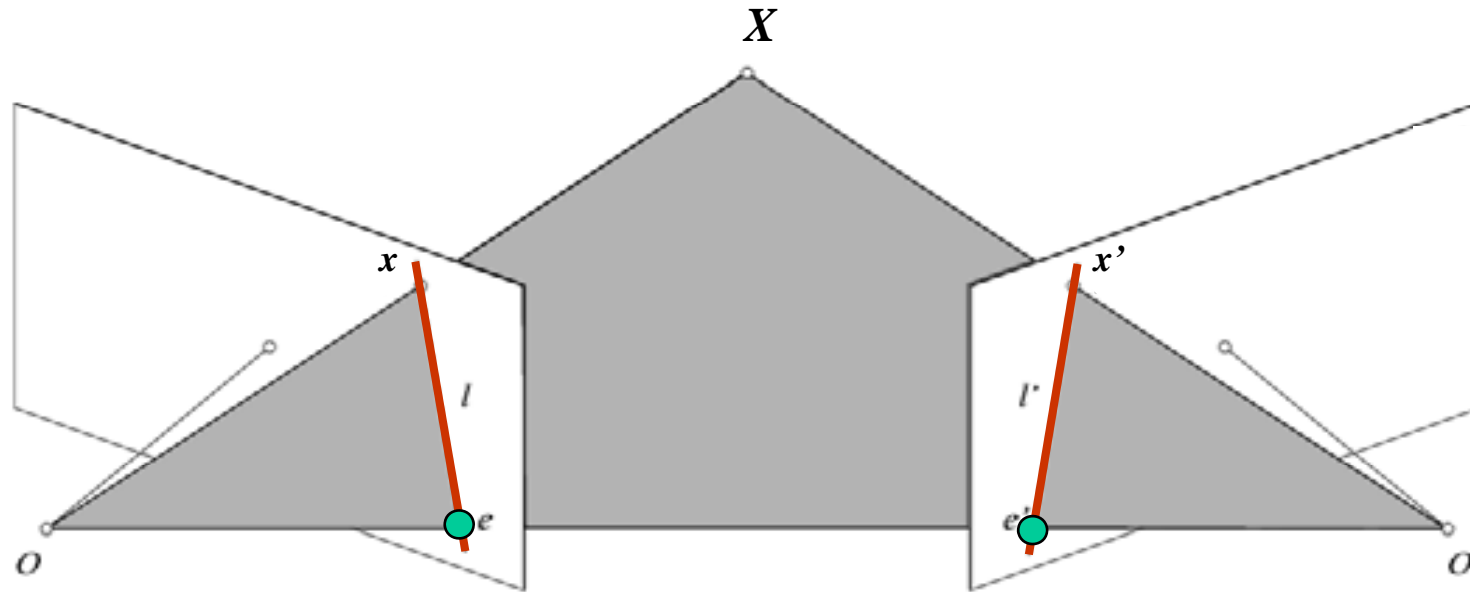


**Существенная матрица
(Essential Matrix)**
(Longuet-Higgins, 1981)

Вектора x , t , и Rx' копланарны



Свойства существенной матрицы

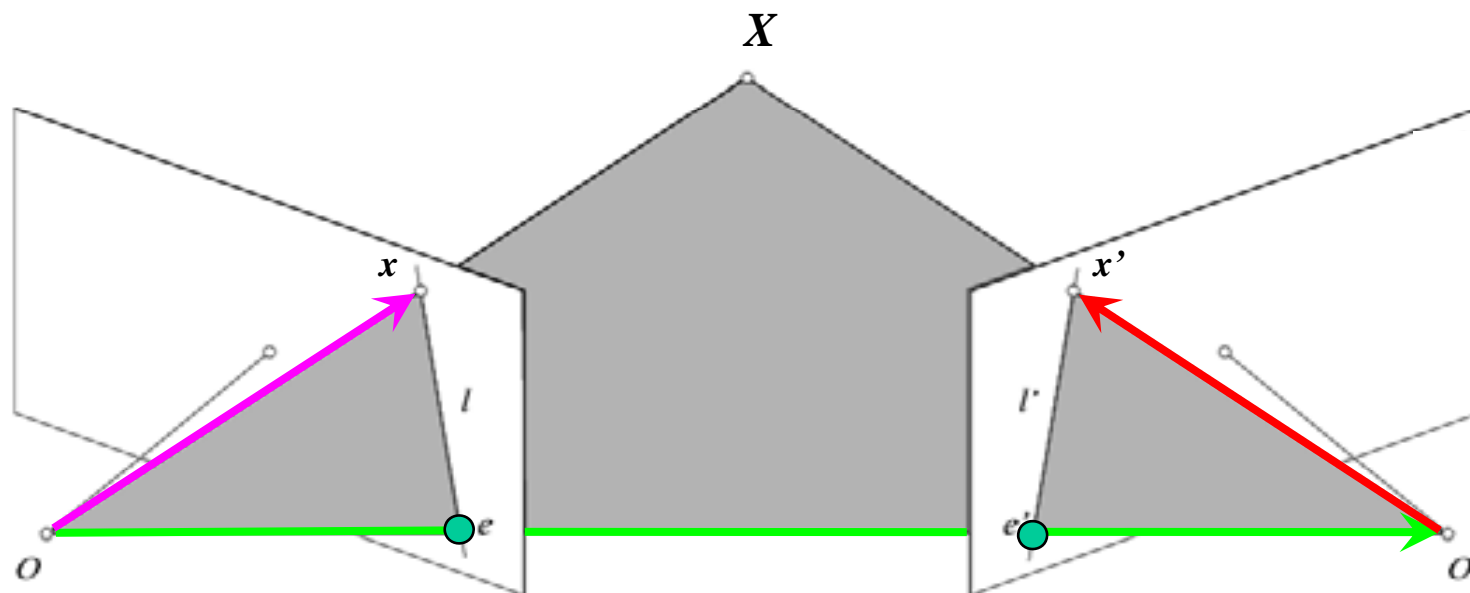


$$x \cdot [t \times (Rx')] = 0 \quad \Rightarrow \quad x^T E x' = 0 \quad \text{где} \quad E = [t_x]R$$

- $E x'$ - эпиполярная линия, соответствующая x' ($l = E x'$)
- $E^T x$ - эпиполярная линия, соответствующая x ($l' = E^T x$)
- $E e' = 0$ и $E^T e = 0$
- E вырождена (ранг 2)
- E имеет 5 степеней свободы



Некалиброванный случай



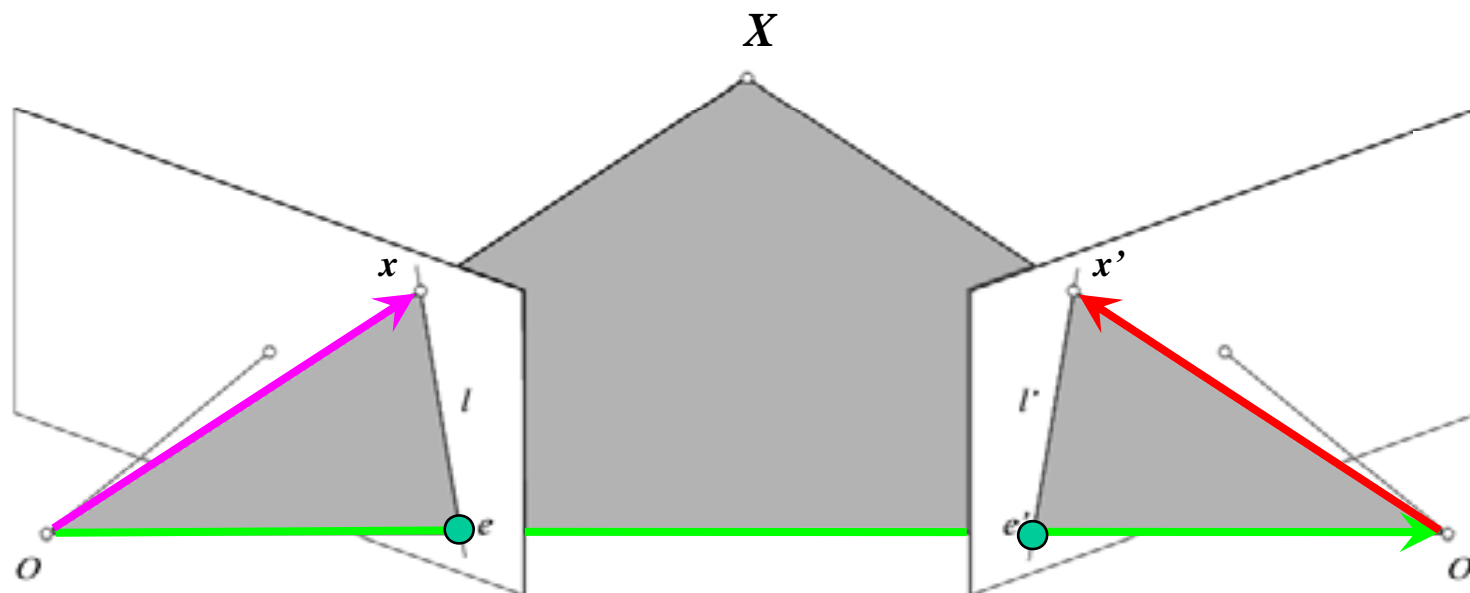
- Пусть матрицы калибровки K и K' для обеих камер неизвестны
- Запишем эпиполярное ограничение через *неизвестные* нормализованные координаты:

$$\hat{x}^T E \hat{x}' = 0$$

$$x = K \hat{x}, \quad x' = K' \hat{x}'$$



Некалиброванный случай



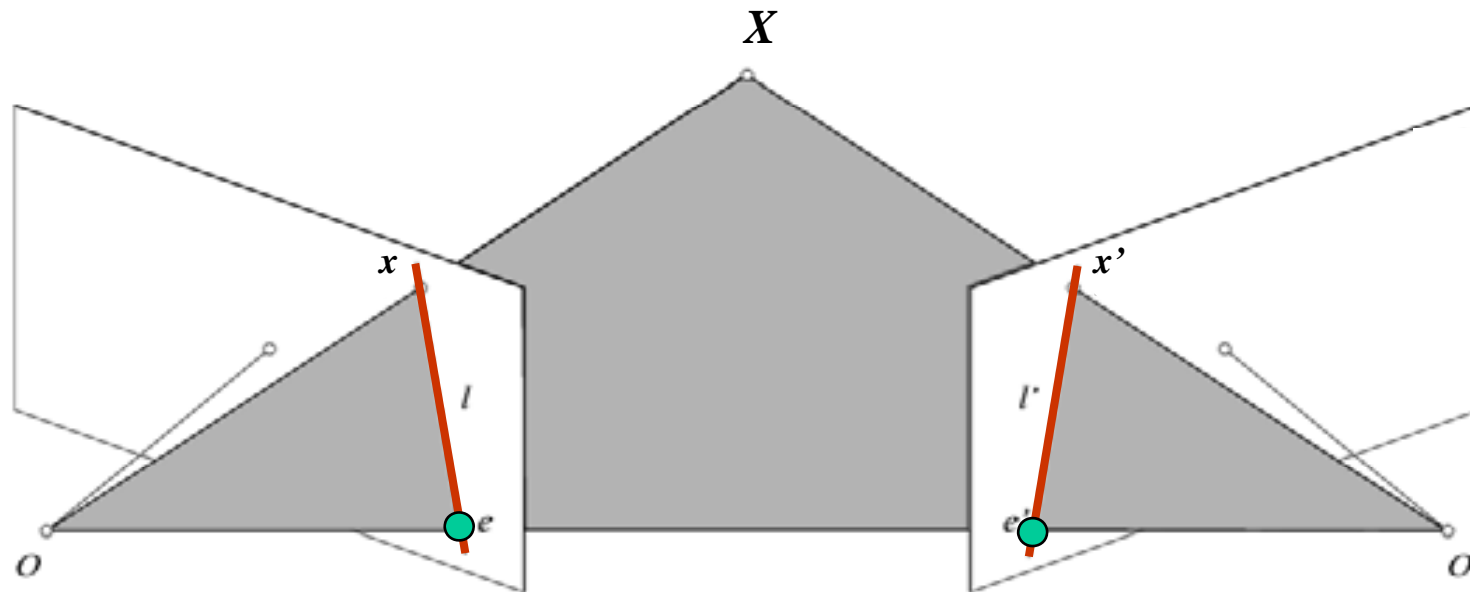
$$\hat{x}^T E \hat{x}' = 0 \quad \longrightarrow \quad x^T F x' = 0 \quad \text{где} \quad F = K^{-T} E K'^{-1}$$

$$x = K \hat{x}$$

$$x' = K' \hat{x}'$$

**Фундаментальная матрица
(Fundamental Matrix)**

(Faugeras and Luong, 1992)



$$\hat{x}^T E \hat{x}' = 0 \quad \longrightarrow \quad x^T F x' = 0 \quad \text{где} \quad F = K^{-T} E K'^{-1}$$

- $F x$ - эпиллярная линия, соответствующая x' ($l = F x'$)
- $F^T x$ - эпиллярная линия, соответствующая x ($l' = F^T x$)
- $F e' = 0$ и $F^T e = 0$
- F вырождена (ранг 2)
- F имеет семь степеней свободы



8-и точечный алгоритм

- Фундаментальная матрица накладывает ограничения на пары соответствующих точек
- Попробуем её вычислить напрямую из соответствий, без калибровки камеры

- Даны соответствия (x, x') : $x = (u, v, 1)^T$, $x' = (u', v', 1)^T$

$$(u, v, 1) \begin{pmatrix} F_{11} & F_{12} & F_{13} \\ F_{21} & F_{22} & F_{23} \\ F_{31} & F_{32} & F_{33} \end{pmatrix} \begin{pmatrix} u' \\ v' \\ 1 \end{pmatrix} = 0 \quad \Rightarrow \quad (uu', uv', u, vu', vv', v, u', v', 1) \begin{pmatrix} F_{11} \\ F_{12} \\ F_{13} \\ F_{21} \\ F_{22} \\ F_{23} \\ F_{31} \\ F_{32} \\ F_{33} \end{pmatrix} = 0$$

- 1 соответствие даёт 1 уравнение на F
- Если точек много, можем свести к обычной задаче однородных наименьших квадратов



8-и точечный алгоритм

$$(uu', uv', u, vu', vv', v, u', v', 1) \begin{pmatrix} F_{11} \\ F_{12} \\ F_{13} \\ F_{21} \\ F_{22} \\ F_{23} \\ F_{31} \\ F_{32} \\ F_{33} \end{pmatrix} = 0 \begin{pmatrix} u_1 u'_1 & u_1 v'_1 & u_1 & v_1 u'_1 & v_1 v'_1 & v_1 & u'_1 & v'_1 \\ u_2 u'_2 & u_2 v'_2 & u_2 & v_2 u'_2 & v_2 v'_2 & v_2 & u'_2 & v'_2 \\ u_3 u'_3 & u_3 v'_3 & u_3 & v_3 u'_3 & v_3 v'_3 & v_3 & u'_3 & v'_3 \\ u_4 u'_4 & u_4 v'_4 & u_4 & v_4 u'_4 & v_4 v'_4 & v_4 & u'_4 & v'_4 \\ u_5 u'_5 & u_5 v'_5 & u_5 & v_5 u'_5 & v_5 v'_5 & v_5 & u'_5 & v'_5 \\ u_6 u'_6 & u_6 v'_6 & u_6 & v_6 u'_6 & v_6 v'_6 & v_6 & u'_6 & v'_6 \\ u_7 u'_7 & u_7 v'_7 & u_7 & v_7 u'_7 & v_7 v'_7 & v_7 & u'_7 & v'_7 \\ u_8 u'_8 & u_8 v'_8 & u_8 & v_8 u'_8 & v_8 v'_8 & v_8 & u'_8 & v'_8 \end{pmatrix} \begin{pmatrix} F_{11} \\ F_{12} \\ F_{13} \\ F_{21} \\ F_{22} \\ F_{23} \\ F_{31} \\ F_{32} \end{pmatrix} = - \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

Минимизируем:

$$\sum_{i=1}^N (x_i^T F x'_i)^2$$

При условии

$$|F|^2 = 1$$



Приведение к рангу 2

- Фундаментальная матрица F имеет ранг 2
- В общем случае, результат DLT-оценки F имеет ранг 3 (из-за влияния шума)
- Необходимо привести к рангу 2
- Найдем F' , минимизирующую норму Фробениуса:

$$\|F - F'\|, \text{ где } \text{rang}(F') = 2$$

- Воспользуемся SVD-разложением:
- $F = UDV^T$, или $F = U \text{diag}(r, s, t) V^T$, где $r \geq s \geq t$
- $F' = U \text{diag}(r, s, 0) V^T$



Недостатки 8и точного метода

$$\begin{pmatrix} u_1 u'_1 & u_1 v'_1 & u_1 & v_1 u'_1 & v_1 v'_1 & v_1 & u'_1 & v'_1 \\ u_2 u'_2 & u_2 v'_2 & u_2 & v_2 u'_2 & v_2 v'_2 & v_2 & u'_2 & v'_2 \\ u_3 u'_3 & u_3 v'_3 & u_3 & v_3 u'_3 & v_3 v'_3 & v_3 & u'_3 & v'_3 \\ u_4 u'_4 & u_4 v'_4 & u_4 & v_4 u'_4 & v_4 v'_4 & v_4 & u'_4 & v'_4 \\ u_5 u'_5 & u_5 v'_5 & u_5 & v_5 u'_5 & v_5 v'_5 & v_5 & u'_5 & v'_5 \\ u_6 u'_6 & u_6 v'_6 & u_6 & v_6 u'_6 & v_6 v'_6 & v_6 & u'_6 & v'_6 \\ u_7 u'_7 & u_7 v'_7 & u_7 & v_7 u'_7 & v_7 v'_7 & v_7 & u'_7 & v'_7 \\ u_8 u'_8 & u_8 v'_8 & u_8 & v_8 u'_8 & v_8 v'_8 & v_8 & u'_8 & v'_8 \end{pmatrix} \begin{pmatrix} F_{11} \\ F_{12} \\ F_{13} \\ F_{21} \\ F_{22} \\ F_{23} \\ F_{31} \\ F_{32} \end{pmatrix} = - \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$



Недостатки 8и точечного метода

250906.36	183269.57	921.81	200931.10	146766.13	738.21	272.19	198.81
2692.28	131633.03	176.27	6196.73	302975.59	405.71	15.27	746.79
416374.23	871684.30	935.47	408110.89	854384.92	916.90	445.10	931.81
191183.60	171759.40	410.27	416435.62	374125.90	893.65	465.99	418.65
48988.86	30401.76	57.89	298604.57	185309.58	352.87	846.22	525.15
164786.04	546559.67	813.17	1998.37	6628.15	9.86	202.65	672.14
116407.01	2727.75	138.89	169941.27	3982.21	202.77	838.12	19.64
135384.58	75411.13	198.72	411350.03	229127.78	603.79	681.28	379.48

$$\begin{pmatrix} F_{11} \\ F_{12} \\ F_{13} \\ F_{21} \\ F_{22} \\ F_{23} \\ F_{31} \\ F_{32} \end{pmatrix} = - \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

- Плохая численная обусловленность
- Можно поправить масштабированием данных



Нормализованный 8-точечный

(Hartley, 1995)

- Переносим центр данных с изображения в начало координат, и масштабируем так, чтобы среднее квадратичное расстояние от точек до начала координат было 2 пикселя
- Применяем 8и точечный метод для вычисления F из нормализованных точек
- Приводим ранг матрицы F к 2
- Преобразуем матрицу к нормальным единицам: если T и T' нормализующие преобразования, тогда фундаментальная матрица в исходных координатах равна $T^T F T'$



“Gold standard”

- Ошибка $\sum_{i=1}^N (x_i^T F x'_i)^2$ - алгебраическая

- Не имеет физического смысла

- «Оптимальная» ошибка:

$$\sum_{i=1}^N [d^2(x'_i, \hat{x}'_i) + d^2(x_i, \hat{x}_i)] , \text{ где } (\hat{x}'_i)^T F \hat{x}_i = 0$$

- Т.е. сумма квадратов расстояний от измеренных точек, до «идеальных»
- Требуется нелинейная оптимизация



Сравнение алгоритмов



	8-point	Normalized 8-point	Nonlinear least squares
Av. Dist. 1	2.33 pixels	0.92 pixel	0.86 pixel
Av. Dist. 2	2.18 pixels	0.85 pixel	0.80 pixel



Если данные зашумлены?

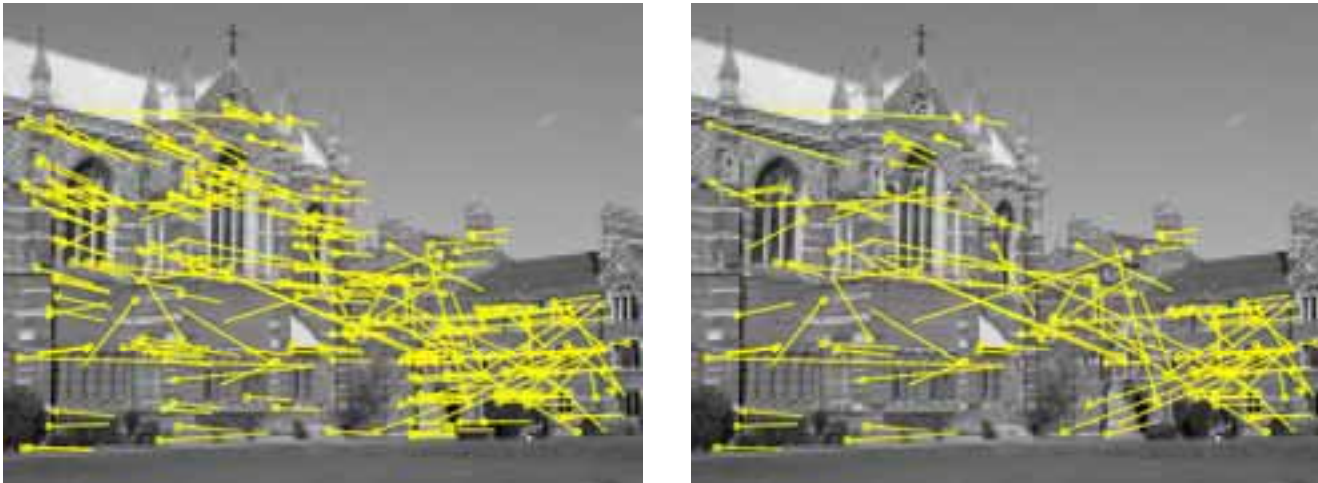
- В соответствиях могут быть ошибки!
 - Ложные соответствия
- Нужно использовать робастные методы
- Обычно используют метод RANSAC с ошибкой:

$$\sum_{i=1}^N \left[d^2(x'_i, Fx_i) + d^2(x_i, F^T \hat{x}_i) \right]$$

- Затем уточняют по всем найденным корректным
соответствиям



Зачем нужны?



- При сопоставлении локальных особенностей возникает много ошибок
- Можем робастно оценить фундаментальную матрицу по множеству соответствий
- Используем фундаментальную матрицу для фильтрации ложных соответствий
 - Если ошибка переноса/репроекции больше порог - выброс

$$\sum_{i=1}^N [d^2(x'_i, Fx_i) + d^2(x_i, F^T \hat{x}_i)]$$



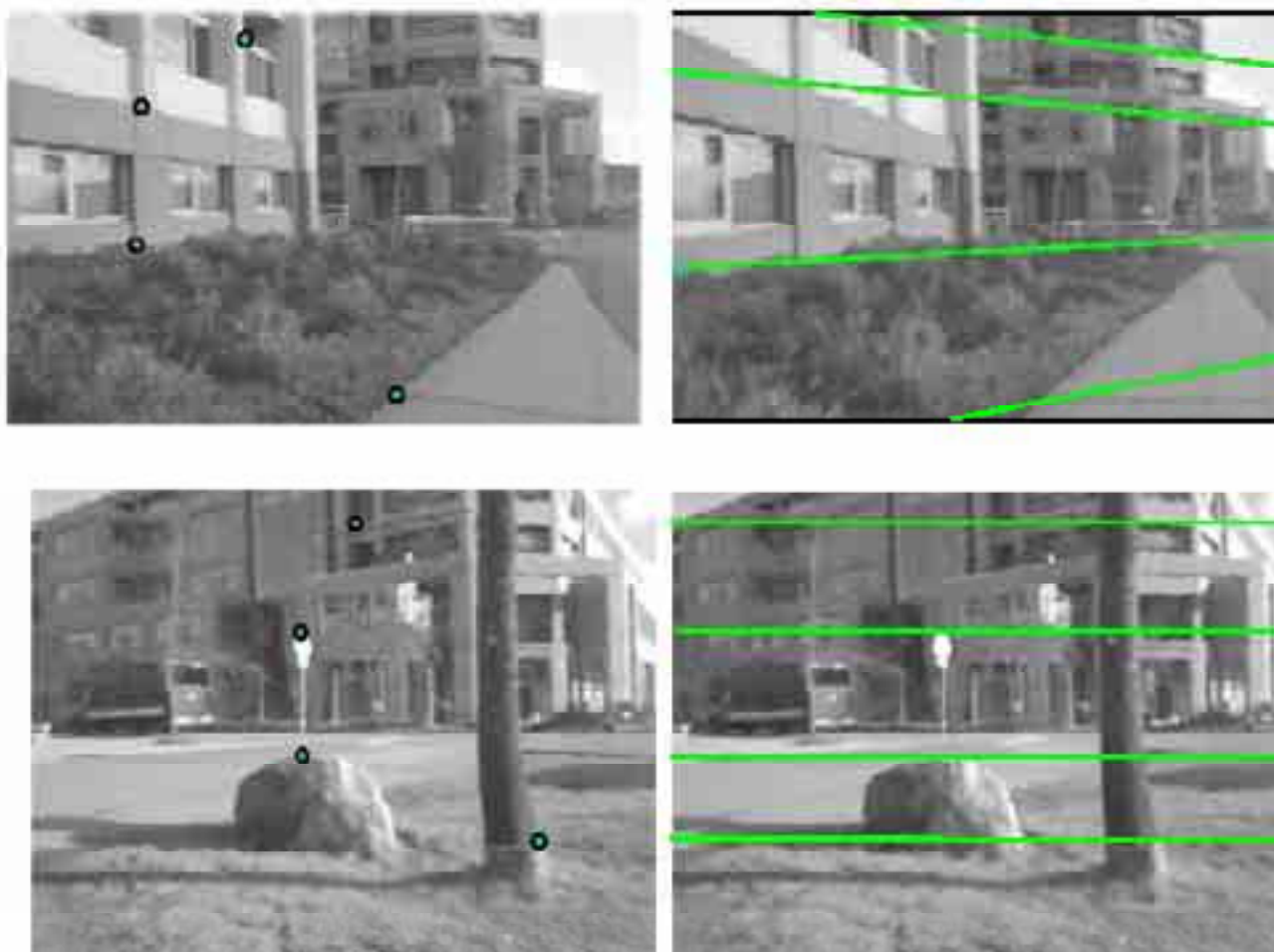
Калибровка камер

- Вычисление фундаментальной матрицы еще называют «слабая калибровка»
- Если известны матрицы калибровки двух камер, тогда мы можем вычислить существенную матрицу: $E = K^T F K'$
- Существенная матрица дает относительный сдвиг и поворот камеры (внешнюю калибровку)

$$E = [t_{\times}]R$$



Упрощение сопоставления



Зная существенную матрицу, можно для любой точки на изображении вычислить соответствующую эпиллярную линию на другом изображении, и искать соответствие только на линии



На следующих лекциях

- Структура из движения
- Стерео-реконструкция
- Реконструкция по 1 камере