Основы трёхмерного зрения

Сергей Носов,
Программист Itseez

Обзор

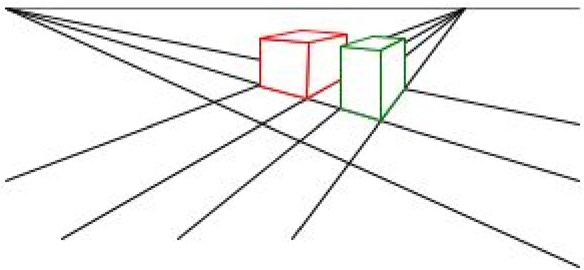
- Проективная геометрия
  - Однородные координаты на прямой, на плоскости
  - Аффинные преобразования
  - Проективные преобразования
- Стерео зрение
  - Краткий обзор подходов
  - Модель камеры и принципы стерео-зрения
  - Метод стерео соответствия Block Matching
- Практическое задание (на дом) и литература

Вопрос!

Пересекаются ли параллельные прямые?

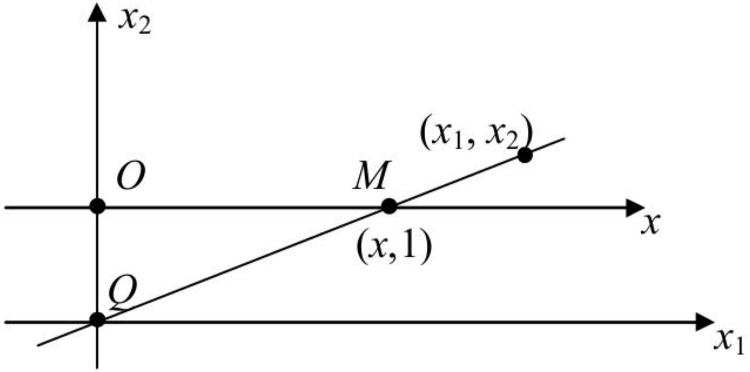
Ответ:

Да или нет, в зависимости от того, что называть параллельными прямыми и что называть пересечением.



Однородные координаты на прямой

- Поместим одномерную ось в двухмерное пространство
- Построим “биекцию” между точками исходной оси и прямыми, проходящими через центр координат двухмерного пространства



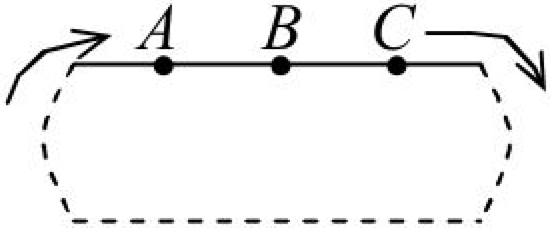
$$x = \frac{x_1}{x_2}$$

- У каждой точки есть много представлений (с точностью до множителя):

$$M = (x) = (x, 1) = (x_1, x_2) = \ldots$$

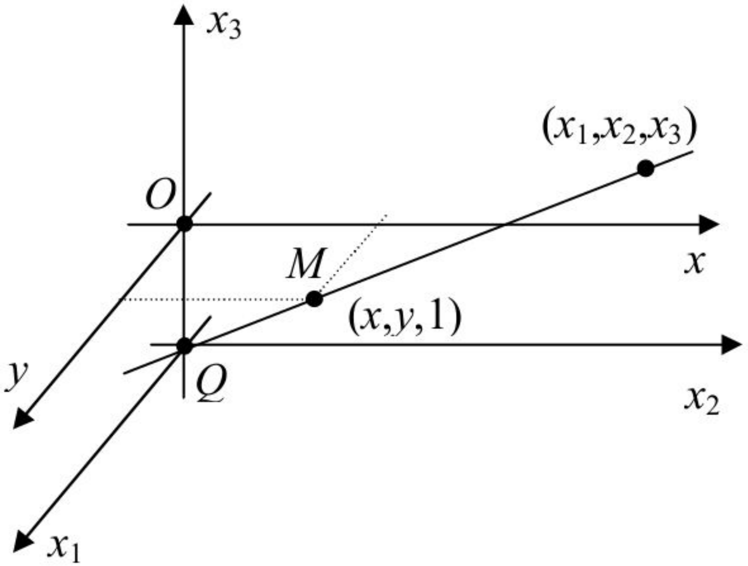
Несобственная точка

- Чтобы биекция стала полноценной - дополним исходную ось т.н. несобственной точкой.
- Её координаты: $(x_1, 0), x_1 \neq 0$
- Несобственная точка как бы замыкает исходную прямую в кольцо



Однородные координаты на плоскости

- Аналогично, поместим плоскую систему координат в трёхмерие



$$M = (x, y) = (x, y, 1) = (x_1, x_2, x_3) = \left(\frac{x_1}{x_3}, \frac{x_2}{x_3}, 1\right) = \ \dots$$

Несобственная прямая

- Аналогично, у прямых, с координатами вида $(x_1, x_2, 0)$ нет соответствующих точек на исходной плоскости.
- Не беда! Дополним ими плоскость и назовём несобственными. А их объединение - несобственной прямой.

Больше плюсов!

- В однородных координатах можно записать уравнение прямой:

$$ax+by+c=0 \to a\frac{x_1}{x_3}+b\frac{x_2}{x_3}+c = 0 \to ax_1+bx_2+cx_3 = 0$$

- Во вспомогательном пространстве это же уравнение описывает плоскость, проходящую через центр координат. А исходная прямая - это её пересечение с плоскостью $x_3 = 1$
- Из уравнения видно, что прямая пересекает несобственную прямую в точке $(-b, a, 0)$ (представление с точностью до множителя)
- А отсюда видно, что все параллельные прямые одного семейства пересекаются! К тому же в одной точке!
- Более того, верно общее утверждение - на проективной плоскости любые две прямые пересекаются в одной точке.

Меня терзают смутные сомнения…

Аффинные преобразования

- Довольно много “полезных” классов преобразований можно задать с помощью линейных функций:

$$\begin{cases} x' = a_{11}x + a_{12}y + p, \\ y' = a_{21}x + a_{22}y + q \end{cases}$$

(в матричном виде $\begin{bmatrix} x' \\ y' \end{bmatrix} = A \cdot \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} p \\ q \end{bmatrix}$, где $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$)

Аффинные преобразования (примеры)

- Сдвиг:

$$\begin{cases} x' = x + p \\ y' = y + q \end{cases}$$

- Растяжение вдоль осей:

$$\begin{cases} x' = 2x \\ y' = 3y \end{cases}$$

- Симметрия:

$$\begin{cases} x' = -x \\ y' = y \end{cases}$$

- Поворот:

$$\begin{cases} x' = x\cos\alpha - y\sin\alpha \\ y' = x\sin\alpha + y\cos\alpha \end{cases}$$

Обратные преобразования

- Возьмём преобразование:

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = A_0 \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} p_0 \\ q_0 \end{bmatrix}$$

- Обратное ему:

$$\begin{bmatrix} x \\ y \end{bmatrix} = A_0^{-1} \left(\begin{bmatrix} x' \\ y' \end{bmatrix} - \begin{bmatrix} p_0 \\ q_0 \end{bmatrix} \right)$$

Суперпозиция преобразований

- Применим ещё одно преобразование:

$$\begin{bmatrix} x'' \\ y'' \end{bmatrix} = A_1 \begin{bmatrix} x' \\ y' \end{bmatrix} + \begin{bmatrix} p_1 \\ q_1 \end{bmatrix} = A_1 A_0 \begin{bmatrix} x \\ y \end{bmatrix} + A_1 \begin{bmatrix} p_0 \\ q_0 \end{bmatrix} + \begin{bmatrix} p_1 \\ q_1 \end{bmatrix}$$

- Какое будет обратное ему?
- Как будет выглядеть суперпозиция трёх преобразований?
- А если третье - это обратное первому?

Преобразования в однородных координатах

- Главная “проблема”: в двухмерном пространстве описанные преобразования именно, что аффинны, а не линейны (из-за наличия свободного члена)
- Однако в трёхмерном пространстве им можно поставить в соответствие линейные преобразования:

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & p \\ a_{21} & a_{22} & q \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

Аффинные преобразования плоскости

Параллельный перенос $T(p,q)=\begin{bmatrix}1&0&p\\0&1&q\\0&0&1\end{bmatrix}$, (p,q) – вектор переноса.	Поворот вокруг начала координат $R(\varphi)=\begin{bmatrix}\cos\varphi&-\sin\varphi&0\\\sin\varphi&\cos\varphi&0\\0&0&1\end{bmatrix}$, φ – угол поворота.	Сжатие-растяжение вдоль осей координат $D(\lambda,\mu)=\begin{bmatrix}\lambda&0&0\\0&\mu&0\\0&0&1\end{bmatrix}$, λ,μ – коэффициенты.
Симметрия относительно оси Ox $S_x=\begin{bmatrix}1&0&0\\0&-1&0\\0&0&1\end{bmatrix}$.	Симметрия относительно оси Oy $S_y=\begin{bmatrix}-1&0&0\\0&1&0\\0&0&1\end{bmatrix}$.	Симметрия относительно начала координат O $S_O=\begin{bmatrix}-1&0&0\\0&-1&0\\0&0&1\end{bmatrix}$.

Суперпозиция в однородных координатах

- Поворот вокруг заданной точки (x_0, y_0) на заданный угол φ
  - Сдвиг начала координат в (x_0, y_0) : $T(-x_0, -y_0)$
  - Поворот относительно начала координат на угол φ : $R(\varphi)$
  - Сдвиг начала координат обратно: $T(x_0, y_0)$

$$T(x_0, y_0) \cdot R(\varphi) \cdot T(-x_0, -y_0)$$

$$\begin{bmatrix} 1 & 0 & x_0 \\ 0 & 1 & y_0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} \cos\varphi & -\sin\varphi & 0 \\ \sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & -x_0 \\ 0 & 1 & -y_0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \cos\varphi & -\sin\varphi & x_0(1-\cos\varphi) + y_0 \sin\varphi \\ \sin\varphi & \cos\varphi & -x_0 \sin\varphi + y_0(1-\cos\varphi) \\ 0 & 0 & 1 \end{bmatrix}$$

Меня <u>опять</u> терзают смутные сомнения…

Проективные (перспективные) преобразования

- Если умножить части на $x_3 = x'_3 \neq 0$, то получится аффинное преобразование в однородных координатах:

$$\begin{bmatrix} x'_1 \\ x'_2 \\ x'_3 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & p \\ a_{21} & a_{22} & q \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

- А если отказаться от требования $x_3 = x'_3 \neq 0$, то мы получим более общее, чем аффинное, а именно *проективное* преобразование:

$$\begin{bmatrix} x'_1 \\ x'_2 \\ x'_3 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

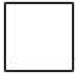
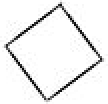
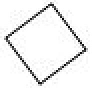
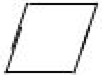
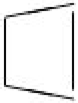
Выражение исходных (аффиных) координат

- Проективному преобразованию соответствует дробно-линейная функция в исходных координатах, хотя в проективном пространстве оно линейно!

$$\begin{cases} x_1' = a_{11}x_1 + a_{12}x_2 + a_{13}x_3, \\ x_2' = a_{21}x_1 + a_{22}x_2 + a_{23}x_3, \\ x_3' = a_{31}x_1 + a_{32}x_2 + a_{33}x_3, \end{cases}$$

$$\begin{cases} x' = \dfrac{a_{11}x + a_{12}y + a_{13}}{a_{31}x + a_{32}y + a_{33}}, \\ y' = \dfrac{a_{21}x + a_{22}y + a_{23}}{a_{31}x + a_{32}y + a_{33}}. \end{cases}$$

Иерархия классов двухмерных преобразований

Transformation	**Matrix**	**# DoF**	**Preserves**	**Icon**
translation	$\begin{bmatrix} \boldsymbol{I} \mid \boldsymbol{t} \end{bmatrix}_{2\times 3}$	2	orientation	
rigid (Euclidean)	$\begin{bmatrix} \boldsymbol{R} \mid \boldsymbol{t} \end{bmatrix}_{2\times 3}$	3	lengths	
similarity	$\begin{bmatrix} s\boldsymbol{R} \mid \boldsymbol{t} \end{bmatrix}_{2\times 3}$	4	angles	
affine	$\begin{bmatrix} \boldsymbol{A} \end{bmatrix}_{2\times 3}$	6	parallelism	
projective	$\begin{bmatrix} \tilde{\boldsymbol{H}} \end{bmatrix}_{3\times 3}$	8	straight lines	

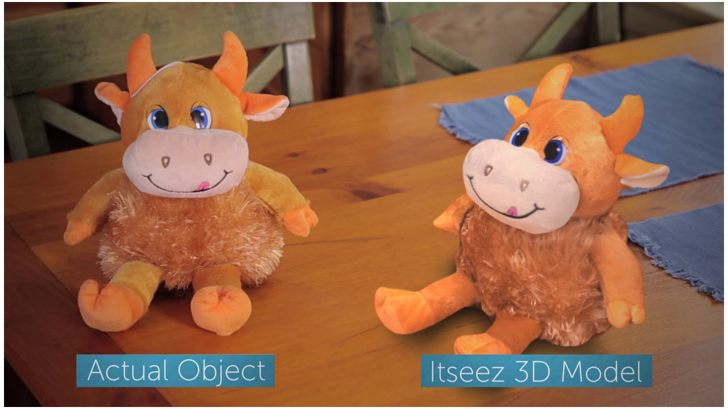
Пример проективного преобразования

Красота-то какая… Лепота!..

Основы стерео зрения

3D-сенсоры: Оценка расстояния до объектов

- Распознавание позы, жестов человека
  - Kinect
  - vCount
- Создание 3D-моделей
  - Itseez3D
  - Карты окружения
- Позиционирование
  - для автономных систем
  - для подсказок

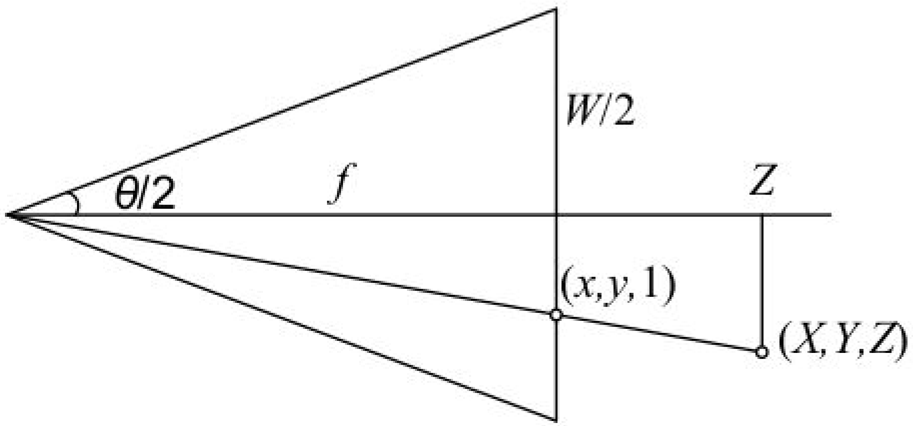


Разнообразие 3D-сенсоров

- Time-of-Flight
    - RADAR
    - Scanning LIDAR
    - Scannerless LIDAR
- Стерео зрение
    - Классическое
    - Structured light

Оптический обман зрения

Модель pinhole камеры



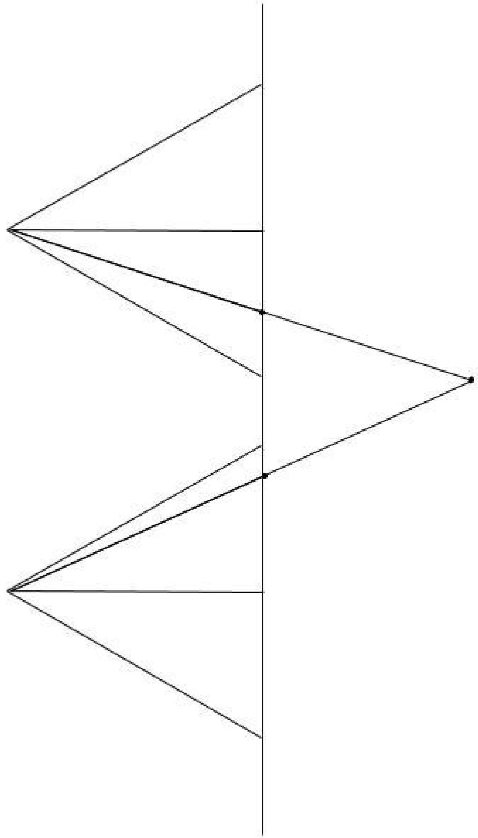
$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} \rightsquigarrow \begin{pmatrix} \frac{X}{Z} \\ \frac{Y}{Z} \\ 1 \end{pmatrix} \rightsquigarrow \begin{pmatrix} f_x \frac{X}{Z} + c_x \\ f_y \frac{Y}{Z} + c_y \\ 1 \end{pmatrix}$$

$$\begin{pmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} \frac{X}{Z} \\ \frac{Y}{Z} \\ 1 \end{pmatrix} = K \cdot \begin{pmatrix} \frac{X}{Z} \\ \frac{Y}{Z} \\ 1 \end{pmatrix} = \begin{pmatrix} u \\ v \\ 1 \end{pmatrix}$$

$$K^{-1} z \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = z \begin{pmatrix} \dot{x} \\ \dot{y} \\ 1 \end{pmatrix}$$

Стерео зрение

- Информация о соответствии точек на изображениях, снятых из различных положений (!) в пространстве позволяет разрешить неопределённость при вычислении расстояния
- Горизонтальное смещение (в пикселях) двух соответствующих точек на стерео-паре называется *диспаритетом* (disparity)



Вычисление расстояния

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \mathrm{K}^{-1} z \begin{pmatrix} u \\ v \\ 1 \end{pmatrix}$$

$$b = \frac{zd}{f_x}$$

$$\begin{pmatrix} x + b \\ y \\ z \end{pmatrix} = K^{-1} z \begin{pmatrix} u + d \\ v \\ 1 \end{pmatrix}$$

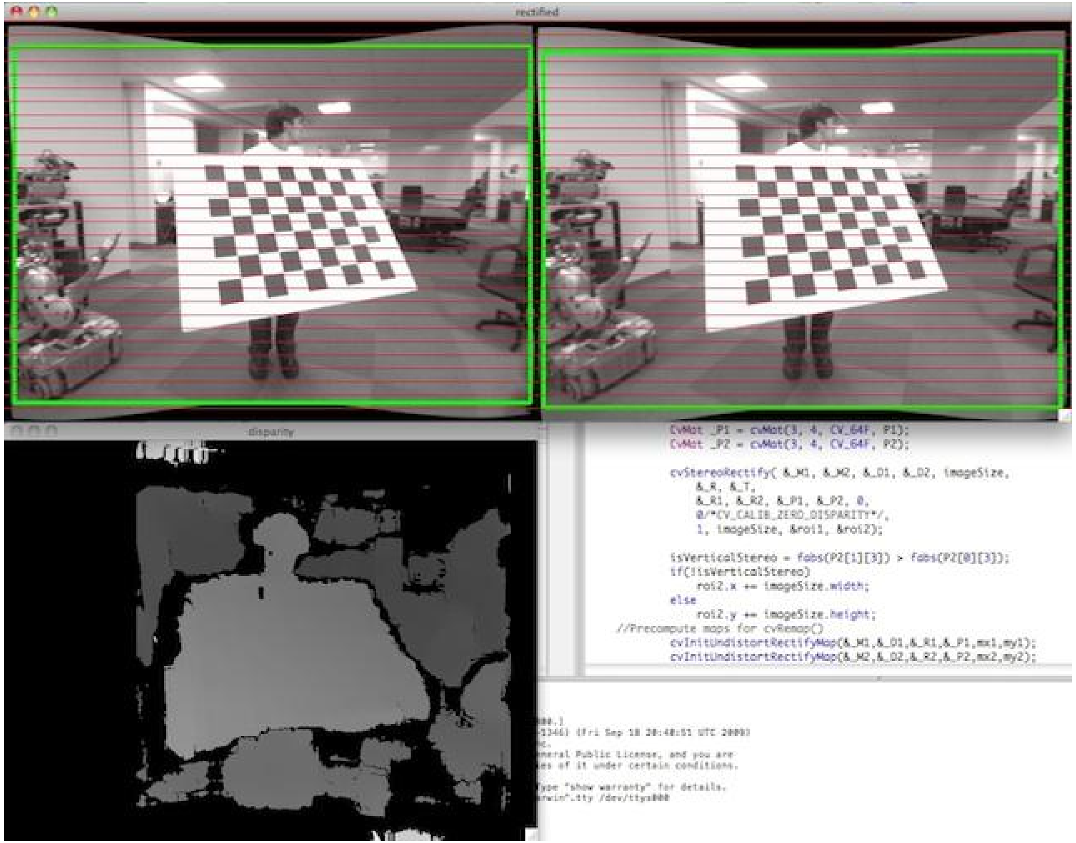
$$z = \frac{f_x b}{d}$$

$$\begin{pmatrix} b \\ 0 \\ 0 \end{pmatrix} = K^{-1} z \begin{pmatrix} d \\ 0 \\ 0 \end{pmatrix}$$

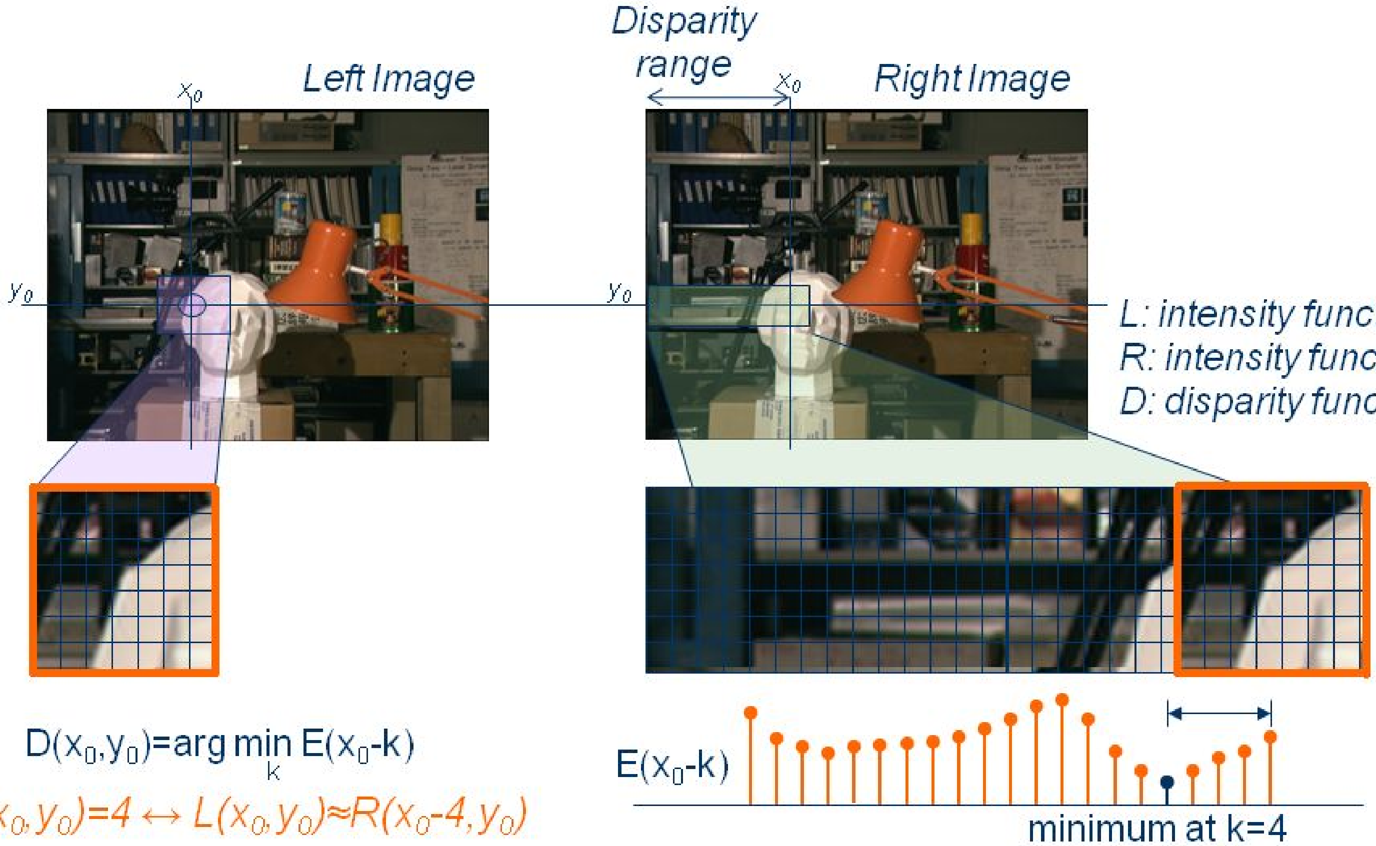
Калибрация стерео-камеры

- Оценка внутренних/intrinsic параметров:
  - f_x, f_y, c_x, c_y, distortion coefficients
- Оценка внешних/extrinsic (R, t) параметров
- Undistortion + Rectification
- Хороший результат: "средняя эпиполярная ошибка" меньше 0.25px

Пример калибрации



Block Matching



L: intensity function of left image
R: intensity function of right image
D: disparity function wrt left image

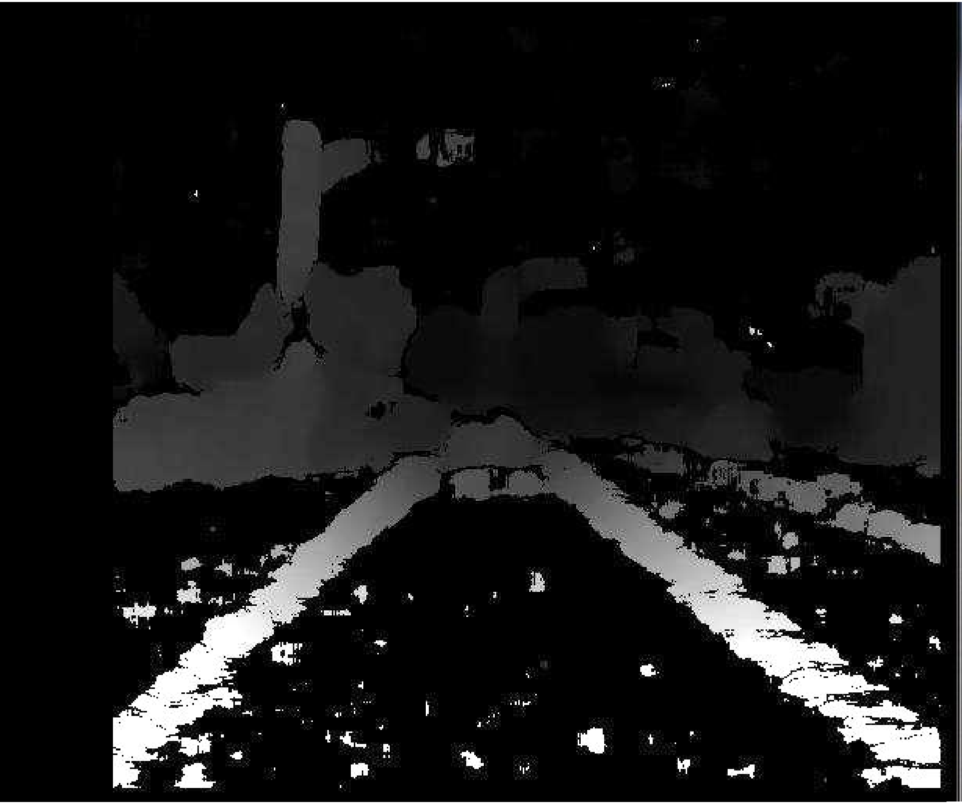
$$D(x_0,y_0)=\arg\min_k E(x_0-k)$$

$$D(x_0,y_0)=4 \leftrightarrow L(x_0,y_0)\approx R(x_0-4,y_0)$$

Краткая характеристика метода

- Выполняется за O(MND) операций
- Классический способ сравнения окон - SAD (Sum-of-Absolute-Differences), типовые размеры окон - от 11x11 до 31x31
- Могут использоваться бинарные дескрипторы, наиболее популярным выбором является Census Transform. Размеры окон в этом случае выбираются меньше, порядка 5x5, 7x7
- Может быть реализован очень эффективно на большинстве современных архитектур

Результат



Практическое задание (на дом)

- Дано *ректифицированное* видео со стерео камеры
- Вычислить диспаритет в каждой точке изображения (block matching)
- Используя известные параметры камеры:

$$K = \begin{pmatrix} 573.835 & 0 & 306 \\ 0 & 573.835 & 256 \\ 0 & 0 & 1 \end{pmatrix} \qquad b = 18\mathrm{cm}$$

Раскрасить точки, которые лежат в полосе шириной 40 см, центральная ось которой параллельна горизонтальной оси камеры и находится на 40 см выше её, в зависимости от расстояния:

  - Точки на расстояния больше 30 метров - зелёные
  - Точки на расстоянии от 0 до 30 метров - линейная смесь красного и зелёного.

Литература

- Шульц М.М. Аналитическая и вычислительная геометрия. Спецкурс. Учебное пособие. – Нижний Новгород: Нижегородский госуниверситет, 2010. – 125 с.
- Szeliski R. Computer Vision: Algorithms and Applications. Springer, 2010. http://szeliski.org/Book/
- http://docs.opencv.org/2.4/modules/calib3d/doc/camera_calibration_and_3d_reconstruction.html
- https://github.com/snosov1/jetson-screencasts-sources/tree/master/10-cv-stereo

Буду рад любым вопросам!