Компьютерное зрение '2014

Who? Александр Вахитов

When? December 9, 2014

План лекции

Определение

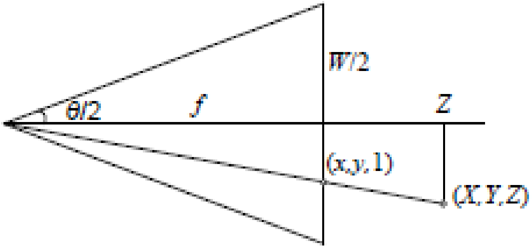
Свойства

Однородные координаты

$$x \in \mathbb{R}^N \quad (N = 2, 3)$$

$$x \to \begin{pmatrix} x \\ 1 \end{pmatrix}$$

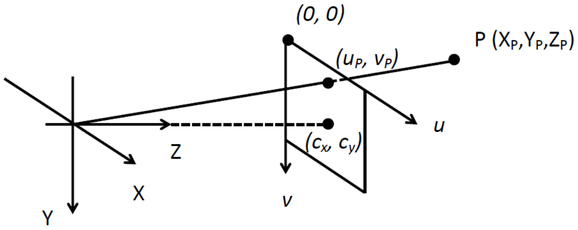
Модель проективной камеры



Понятия

Угол обзора θ, фокусное расстояние f, размер матрицы W

Проективная камера в 3D



$$u = f\frac{X}{Z} + c_x; \quad v = f\frac{Y}{Z} + c_y$$

Проектирование в однородных координатах

$$K = \begin{pmatrix} f & 0 & c_x \\ 0 & f & c_y \\ 0 & 0 & 1 \end{pmatrix}$$

$$\lambda \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \begin{pmatrix} & & 0 \\ & K & 0 \\ & & 0 \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}$$

Евклидово преобразование в однородных координатах

$$\begin{pmatrix} X' \\ Y' \\ Z' \\ 1 \end{pmatrix} = \begin{pmatrix} & R & & t \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}$$

Проектирование на камеру точек из глобальной системы координат

$$\lambda \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = K(R\ t) \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}$$

Основные задачи геометрии в компьютерном зрении

- определение положения калиброванной камеры по проекциям точек
- определение положения точки по ее проекциям на камеры (триангуляция)
- калибровка камеры по проекциям точек

Положение калиброванной камеры по проекциям точек

Дано Матрица камеры K, проекции $(u_i, v_i)_{i=1}^N$, точки $(X_i, Y_i, Z_i)_{i=1}^N$

Найти R, t :

$$\sum_i \|\pi_u(X_i, Y_i, Z_i, K, R, t) - u_i\|^2 +$$
$$+ \sum_i \|\pi_v(X_i, Y_i, Z_i, K, R, t) - v_i\|^2 \to \min_{R,t}$$

Если все точки принадлежат одному объекту, то задача эквивалентна поиску ракурса съемки объекта

Положение калиброванной камеры по проекциям точек: метод

- Домножение на K^{-1}
- Поиск $P = [R\ t] \in M(3,4)$
- Уточнение - поиск $\phi, t : R = R(\phi)$, $\phi \in \mathbb{R}^3$

Положение камеры относительно точек объекта

Пусть известны координаты N точек трехмерного объекта $\{(X_i, Y_i, Z_i)\}_{i=1}^N$ и их проекции на плоскость изображения $\{(x_i, y_i)\}_{i=1}^N$. Задача нахождения матрицы камеры P по этим данным называется задачей Perspective-n-Point (PnP).

$$\lambda_i \begin{pmatrix} x_i \\ y_i \\ 1 \end{pmatrix} = P \begin{pmatrix} X_i \\ Y_i \\ Z_i \\ 1 \end{pmatrix}, \quad i = 1 \dots N.$$

Неизвестны: λ_i, P.

Положение камеры относительно объекта (2)

$$\lambda_i \begin{pmatrix} x_i \\ y_i \\ 1 \end{pmatrix} = P \begin{pmatrix} X_i \\ Y_i \\ Z_i \\ 1 \end{pmatrix}, \quad i = 1 \dots N.$$

Неизвестны: λ_i, P.

- Сколько уравнений?

Положение камеры относительно объекта (2)

$$\lambda_i \begin{pmatrix} x_i \\ y_i \\ 1 \end{pmatrix} = P \begin{pmatrix} X_i \\ Y_i \\ Z_i \\ 1 \end{pmatrix}, \quad i = 1 \dots N.$$

Неизвестны: λ_i, P.

- Сколько уравнений?
- Сколько точек необходимо для однозначного решения?

Положение камеры относительно объекта (2)

$$\lambda_i \begin{pmatrix} x_i \\ y_i \\ 1 \end{pmatrix} = P \begin{pmatrix} X_i \\ Y_i \\ Z_i \\ 1 \end{pmatrix}, \quad i = 1 \dots N.$$

Неизвестны: λ_i, P.

- Сколько уравнений?
- Сколько точек необходимо для однозначного решения?

Положение камеры относительно объекта (2)

$$\lambda_i \begin{pmatrix} x_i \\ y_i \\ 1 \end{pmatrix} = P \begin{pmatrix} X_i \\ Y_i \\ Z_i \\ 1 \end{pmatrix}, \quad i = 1 \dots N.$$

Неизвестны: λ_i, P.

- Сколько уравнений?
- Сколько точек необходимо для однозначного решения?
- Какие возможны неоднозначности?

Положение камеры относительно объекта (2)

$$\lambda_i \begin{pmatrix} x_i \\ y_i \\ 1 \end{pmatrix} = P \begin{pmatrix} X_i \\ Y_i \\ Z_i \\ 1 \end{pmatrix}, \quad i = 1 \dots N.$$

Неизвестны: λ_i, P.

- Сколько уравнений?
- Сколько точек необходимо для однозначного решения?
- Какие возможны неоднозначности?

Положение камеры относительно объекта (2)

$$\lambda_i \begin{pmatrix} x_i \\ y_i \\ 1 \end{pmatrix} = P \begin{pmatrix} X_i \\ Y_i \\ Z_i \\ 1 \end{pmatrix}, \quad i = 1 \ldots N.$$

Неизвестны: λ_i, P.

- Сколько уравнений?
- Сколько точек необходимо для однозначного решения?
- Какие возможны неоднозначности?
- Если $P = K[R\, t]$, то сколько точек должно хватать?

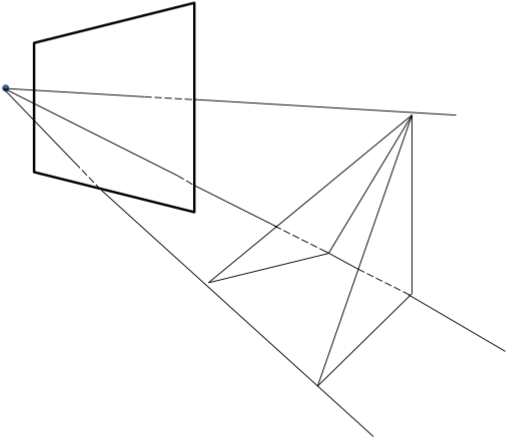
Положение камеры относительно объекта (2)

$$\lambda_i \begin{pmatrix} x_i \\ y_i \\ 1 \end{pmatrix} = P \begin{pmatrix} X_i \\ Y_i \\ Z_i \\ 1 \end{pmatrix}, \quad i = 1 \dots N.$$

Неизвестны: λ_i, P.

- Сколько уравнений?
- Сколько точек необходимо для однозначного решения?
- Какие возможны неоднозначности?
- Если $P = K[R\,t]$, то сколько точек должно хватать?

Неоднозначность



Две позиции треугольника соответствуют одному изображению

Параметризация вращений

Известный факт: любое вращение есть поворот относительно одной оси на известный угол. Пусть ось $n \in S^2$, угол $\theta \in \mathbb{R}$, тогда $\phi = n\theta$. Для вычисления R по ϕ:

- $\theta = \|\phi\|$, $n = \phi / \|\phi\|$.
- формула Родрига (Rodrigues) (по оси и углу дает матрицу поворота) - cv2.rogrigues в OpenCV

Метод Левенберга Марквардта

Метод Ньютона для нелинейных наименьших квадратов:

$$\sum (x_i - \mu_i(w))^2 \to \min$$

Есть библиотека levmar
`http://users.ics.forth.gr/~lourakis/levmar/`

Гомография

Определение

Гомография H - Обратимое отображение проективного пространства $\mathbb{P}^n$ на само себя, такое, что в однородных координатах оно представимо как домножение на матрицу.

Пример

$$x = (1, 1)^T, \quad H = \begin{pmatrix} 5 & -5 \\ 0 & 5 \end{pmatrix} \quad y = Hx = ?$$

Гомография (2)

Уравнение прямой l в однородных координатах на проективной плоскости $\mathbb{P}^3$:

$$l^T x = 0$$

Свойство

Гомография H переводит прямые в прямые.

$$x_H = Hx, \quad l_H = ?$$

Гомография (2)

Уравнение прямой l в однородных координатах на проективной плоскости $\mathbb{P}^3$:

$$l^T x = 0$$

Свойство

Гомография H переводит прямые в прямые.

$$x_H = Hx, \quad l_H = ?$$

$$l_H = H^{-T} l \implies l_H^T x_H = 0.$$

Гомография (2)

Уравнение прямой l в однородных координатах на проективной плоскости $\mathbb{P}^3$:

$$l^T x = 0$$

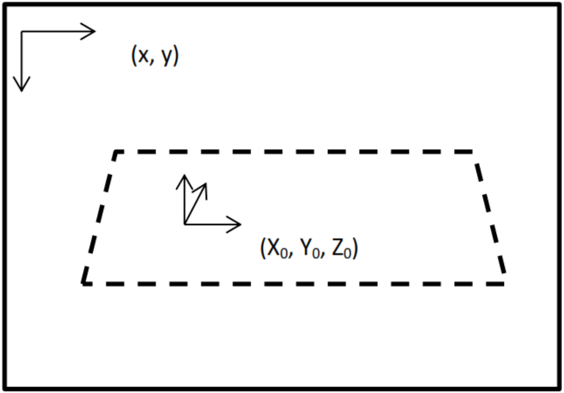
Свойство

Гомография H переводит прямые в прямые.

$$x_H = Hx, \quad l_H = ?$$

$$l_H = H^{-T} l \implies l_H^T x_H = 0.$$

Гомография между реальной и изображенной плоскостями



$$\mathbf{x} \equiv K[\mathbf{r}_1\ \mathbf{r}_2\ \mathbf{r}_3\ \mathbf{t}]\begin{pmatrix} X \\ Y \\ 0 \\ 1 \end{pmatrix} = K[\mathbf{r}_1\ \mathbf{r}_2\ \mathbf{t}]\begin{pmatrix} X \\ Y \\ 1 \end{pmatrix}$$

Выбросы

Выброс (outlier) - наблюдение, не подчиняющееся принятой математической модели
Например:

$$y_i = x + v_i, \quad v \in N(0,1), \; i = 1 \dots K.$$

Пусть $x = 2$

$$\{y_i\} = \{3.53263030828475$$
$$1.23033408624632$$
$$2.37137881276006$$
$$1.77441559772875$$
$$3.11735613881447$$
$$72.5746540499043$$
$$0.910935704947764$$
$$2.03255746416497$$
$$2.55252702111222$$

Фильтрация выбросов методом RANSAC

Задача найти $\hat{x}$ по наблюдениям y_i

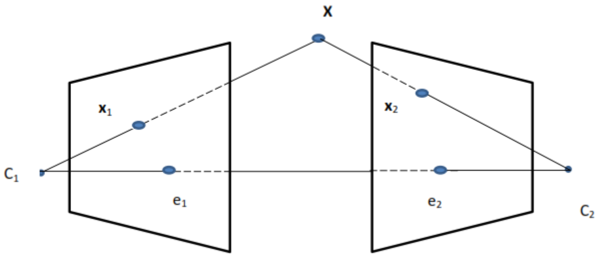
Алгоритм В цикле по $j = 1 \dots N$

- случайно выбрать M наблюдений $\{y_{i_j}\}$
- построить гипотезу о значении неизвестного вектора $\hat{x}_j$ по $\{y_i\}$
- определить множество S_j наблюдений, удовлетворяющих гипотезе $\hat{x}_j$
- $V_j = |S_j|$

Выбрать

$$\hat{x} = \hat{x}_j : \quad j = argmax\ V_j$$

Двухкамерное сопоставление



Система координат задана камерой C_1.

$$\lambda_1 \mathbf{x_2} = K[R\ t]\begin{pmatrix}\mathbf{X}\\ \mathbf{1}\end{pmatrix} \implies e_2 = Kt, \ \ \mathbf{x_1} = \lambda_2 \mathbf{X}$$

Проведем прямую через $\mathbf{x_2}, \mathbf{e_2}$:

$$l(\mathbf{x_2}, \mathbf{e_2}) = \mathbf{x_2} \times \mathbf{e_2} \equiv (K(R\lambda_3 \mathbf{X} + t)) \times Kt = \lambda_3 KR\mathbf{X} \times Kt =$$

$$= F\mathbf{X} = \lambda_2 F\mathbf{x_1} \implies \mathbf{x_2^T} F\mathbf{x_1} = 0$$

Двухкамерное сопоставление (2)

Определение Эпиполярная линия (прямая) - прямая, проходящая через проекцию центра одной камеры на другую.

Определение Фундаментальная матрица - матрица, домножение которой на однородные координаты проекции точки на первую камеру дает однородные координаты эпиполярной прямой на второй камере.

$$l(\mathbf{x_2}, \mathbf{e_2}) = F\mathbf{x_1} \implies \mathbf{x_2}^T F\mathbf{x_1} = 0$$

Записав систему уравнений, можно вычислить фундаментальную матрицу
Пусть $\bar{\mathbf{x_1}} = K^{-1}\mathbf{x_1}$, $\bar{\mathbf{x_2}} = K^{-1}\mathbf{x_2}$. Аналогично,

$$l(\bar{\mathbf{x_2}}, \bar{\mathbf{e_2}}) = R\mathbf{X} \times t \equiv [t]_x R\bar{\mathbf{x_1}},$$

Отсюда $E = [t]_x R$ где $[t]_x$ - матрица, домножение на которую дает векторное произведение на t.

Существенная матрица

Определение

Существенная матрица - фундаментальная матрица, заданная для камер, у которых матрица камеры - единичная ($K = I$).

$$SVD(E) = Udiag(1, 1, 0)V^T$$

Пусть

$$W = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad Z = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Возможны две факторизации $E = SR$:

$$S = UZU^T,\ R = UWV^T \text{ или } R = UW^TV^T$$

Какая из факторизаций верна, решаем проверкой того, что получившиеся точки лежат перед камерой

Вычислительные методы для фундаментальной и существенной матрицы

Даны: пары проекций точек на первую и вторую камеру $\{(\bar{\mathbf{x}}_\mathbf{i}, \bar{\mathbf{y}}_\mathbf{i})\}_{i=1}^N$
Найти: F (или E)
сколько точек надо?

Вычислительные методы для фундаментальной и существенной матрицы

Даны: пары проекций точек на первую и вторую камеру $\{(\bar{\mathbf{x}}_{\mathbf{i}}, \bar{\mathbf{y}}_{\mathbf{i}})\}_{i=1}^{N}$
Найти: F (или E)
сколько точек надо?

- *8-точечный метод* Составляем $Af = 0$, решаем
- *7-точечный метод* По определению, $rank(F) = 2 \implies det(F) = 0$. Следовательно, можно взять 7 точек, тогда $f = \alpha v_7 + (1-\alpha)v_8$, отсюда находим $\alpha : rank(F) = 0$.

Также, можно использовать RANSAC и ММП. При применении RANSAC, 7-точечный метод стабильнее.

Вычислительные методы для фундаментальной и существенной матрицы

Даны: пары проекций точек на первую и вторую камеру $\{(\bar{\mathbf{x}}_\mathbf{i}, \bar{\mathbf{y}}_\mathbf{i})\}_{i=1}^{N}$
Найти: F (или E)
сколько точек надо?

- *8-точечный метод* Составляем $Af = 0$, решаем
- *7-точечный метод* По определению, $rank(F) = 2 \implies det(F) = 0$. Следовательно, можно взять 7 точек, тогда $f = \alpha v_7 + (1-\alpha)v_8$, отсюда находим $\alpha : rank(F) = 0$.

Также, можно использовать RANSAC и ММП. При применении RANSAC, 7-точечный метод стабильнее.

Гомография бесконечно далекой плоскости

$$\mathbf{x} \equiv K[\mathbf{r_1}\ \mathbf{r_2}\ \mathbf{r_3}\ \mathbf{t}] \begin{pmatrix} X \\ Y \\ Z \\ 0 \end{pmatrix} = K[\mathbf{r_1}\ \mathbf{r_2}\ \mathbf{r_3}] \begin{pmatrix} X \\ Y \\ Z \end{pmatrix}$$

Не зависит от t (эффект параллакса наблюдается только для достаточно близких точек).

Задача вычисления гомографии

Модель:

1. $\lambda \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = P \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}$ проектирование

2. $\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} v_x \\ v_y \end{pmatrix}$ зашумление

- нахождение гомографии без шумов по парам точка-проекция - *DLT*
- устойчивость к шумам - *MLE*
- возможность не учитывать ложные соответствия “точка-проекция” (т.н. выбросы, outliers) - *RANSAC*

1. Нахождение гомографии методом Direct Linear Transformation

$$\lambda_i \mathbf{x_i}' = H\mathbf{x} \implies \mathbf{x_i}' \times H\mathbf{x_i} = 0.$$

Два линейно независимых уравнения (согласно определению векторного произведения)
А сколько уравнений нужно?
Получаем систему вида

$$A\mathbf{h} = \mathbf{0}, \quad h - \text{векторизованная матрица } H$$

$$\text{svd}: \; A = USV^T \implies \|Av_n\| = \min_{v:\,\|v\|=1} \|Av\|,$$

$$V = [v_1 \dots v_n]$$

SVD - Singular Value Decomposition

$$A = USV^T$$

Любая прямоугольная матрица A представима в виде произведения ортогональной U, диагональной прямоугольной S и другой ортогональной V

$$S = diag(s_1, \ldots, s_n),\ s_1 \geq s_2 \geq \ldots \geq s_n \geq 0.$$

Когда $U = V$?

2. Устойчивость к шумам

Метод максимального правдоподобия (MLE, ММП)
Обозначим x как наблюдение, w как параметр модели. Тогда ММП заключается в максимизации

$$P(x|w) \to \max$$

Если $P(x|w) \in \mathcal{N}(\mu, \Sigma)$, то

$$P(x|w) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\sum \frac{(x_i - \mu_i(w))^2}{2\sigma_i^2}}$$

Пусть $\forall i \ \sigma_i = \sigma$. Тогда

$$P(x|w) \to \max \Leftrightarrow ln\ P(x|w) \to \max \Leftrightarrow$$

$$\sum (x_i - \mu_i(w))^2 \to \min$$

2. Устойчивость к шумам (2)

Метод Левенберга Марквардта
Предназначен для локальной минимизации функций в виде суммы квадратов разностей наблюдений и значений функций

$$\sum (x_i - \mu_i(w))^2 \to \min$$

Есть библиотека levmar
`http://users.ics.forth.gr/~lourakis/levmar/`

3. Устойчивость к выбросам: RANSAC (Random Sample Consensus)

Пусть выборка пар “точка-проекция” содержит также ошибочные, неверно сопоставленные пары. Стоит задача выявить модель и множество пар, с ней согласованных

В цикле

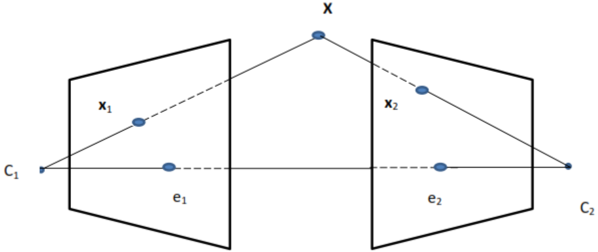
- Выбрать случайно k пар (достаточно для гипотезы о значении матрицы гомографии)
- Посчитать гомографию по k выбранным парам
- Проверить все остальные пары: подходят ли они под построенную гипотезу?

Выполняется фиксированное число итераций либо до момента, пока заданный процент имеющихся пар не подойдет под модель
Проблемы: случайный характер (раз от разу программа ведет себя по-разному), сложности с проверкой, подходят ли пары под гипотезу

Метод определения положения камеры по точкам объекта и их проекциям

Аналогично подсчету гомографии.
Сколько пар точка-проекция нужно для DLT в этом случае?
Сколько итераций RANSAC нужно в этой задаче, если 60% пар точка-проекция - ошибочные?

Двухкамерное сопоставление



Система координат сонаправлена с камерой C_1.

$$\mathbf{x_2} \equiv K[R\ t]\begin{pmatrix}\mathbf{X}\\ 1\end{pmatrix} \implies e_2 = Kt,\quad \mathbf{x_1} \equiv \mathbf{X}$$

$$l(\mathbf{x_2}, \mathbf{e_2}) = \mathbf{x_2} \times \mathbf{e_2} \equiv (KR\mathbf{X} + Kt) \times Kt = KR\mathbf{X} \times Kt =$$

$$= F\mathbf{X} \equiv F\mathbf{x_1} - \text{прямая}$$

Двухкамерное сопоставление (2)

Определение Эпиполярная линия (прямая) - прямая, проходящая через проекцию центра одной камеры на другую.

Определение Фундаментальная матрица - матрица, домножение которой на однородные координаты проекции точки на первую камеру дает однородные координаты эпиполярной прямой на второй камере.

$$l(\mathbf{x_2}, \mathbf{e_2}) \equiv F\mathbf{x_1} \implies \mathbf{x_2}^T F\mathbf{x_1} = 0$$

Записав систему уравнений на проекции точек, можно вычислить фундаментальную матрицу
Пусть $\bar{\mathbf{x_1}} = K^{-1}\mathbf{x_1}$, $\bar{\mathbf{x_2}} = K^{-1}\mathbf{x_2}$. Аналогично,

$$l(\bar{\mathbf{x_2}}, \bar{\mathbf{e_2}}) = R\mathbf{X} \times t \equiv [t]_x R\bar{\mathbf{x_1}},$$

Отсюда $E = [t]_x R$ где $[t]_x$ - матрица, домножение на которую дает векторное произведение на t.

Существенная матрица

Определение

Существенная матрица - фундаментальная матрица, заданная для камер, у которых матрица камеры - единичная ($K = I$).

$$SVD(E) = Udiag(1,1,0)V^T$$

Пусть

$$W = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad Z = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Возможны две факторизации $E = SR$:

$$S = UZU^T,\ R = UWV^T \ \text{ или } \ R = UW^TV^T$$

Какая из факторизаций верна, решаем проверкой того, что получившиеся точки лежат перед камерой

Вычислительные методы для фундаментальной и существенной матрицы

Даны: пары проекций точек на первую и вторую камеру $\{(\bar{\mathbf{x}}_\mathbf{i}, \bar{\mathbf{y}}_\mathbf{i})\}_{i=1}^N$
Найти: F (или E)
сколько точек надо?

Вычислительные методы для фундаментальной и существенной матрицы

Даны: пары проекций точек на первую и вторую камеру $\{(\bar{\mathbf{x}}_\mathbf{i}, \bar{\mathbf{y}}_\mathbf{i})\}_{i=1}^N$
Найти: F (или E)
сколько точек надо?

- *8-точечный метод* Составляем $Af = 0$, решаем
- *7-точечный метод* По определению, $rank(F) = 2 \implies det(F) = 0$. Следовательно, можно взять 7 точек, тогда $f = \alpha v_7 + (1-\alpha) v_8$, отсюда находим $\alpha : rank(F) = 0$.

Также, можно использовать RANSAC и ММП. При применении RANSAC, 7-точечный метод стабильнее.

Вычислительные методы для фундаментальной и существенной матрицы

Даны: пары проекций точек на первую и вторую камеру $\{(\bar{\mathbf{x}}_\mathbf{i}, \bar{\mathbf{y}}_\mathbf{i})\}_{i=1}^{N}$
Найти: F (или E)
сколько точек надо?

- *8-точечный метод* Составляем $Af = 0$, решаем
- *7-точечный метод* По определению, $rank(F) = 2 \implies det(F) = 0$. Следовательно, можно взять 7 точек, тогда $f = \alpha v_7 + (1 - \alpha) v_8$, отсюда находим $\alpha : rank(F) = 0$.

Также, можно использовать RANSAC и ММП. При применении RANSAC, 7-точечный метод стабильнее.

Задачи

1. Известна формула для фундаментальной матрицы, переводящей координаты проекции на первую камеру в координаты прямой на второй камере. Как через нее выразить фундаментальную матрицу, переводящую координаты проекции на вторую камеру - в координаты прямой на первой камере?
2.