

Задача

Дано:

1. Модель камеры:

а) Разрешение в пикселях:

$$Width \times Height$$

б) Угол обзора по вертикали:

$$\alpha_{Height}$$

Известно, что изображение камеры — это проекция пространства на прямоугольник с соотношением сторонами $Width : Height$ через точку, расположенную за прямоугольником на его центральном перпендикуляре — точку фокуса.

2. n точек:

а) Для каждой точки реальные координаты в трехмерном пространстве:

$$X = \begin{pmatrix} x_r \\ y_r \\ z_r \end{pmatrix}$$

б) Для каждой точки координаты на изображение:

$$P_{pix} = \begin{pmatrix} u_{pix} \\ v_{pix} \end{pmatrix}$$

По данной информации требуется найти трехмерную матрицу трансформации камеры.

Решение

Прежде чем решать задачу введем некоторые договоренности

а) Оси координат расположены так, что:

$$OZ = OX \times OY$$

б) Трехмерная матрица трансформации имеет следующий вид:

$$T = \left(R \left| \begin{matrix} t_x \\ t_y \\ t_z \end{matrix} \right. \right)$$

, где R — трехмерная матрица поворота, а $t = \begin{pmatrix} t_x \\ t_y \\ t_z \end{pmatrix}$ — ввектор сдвига.

1. Перевод пиксельных координат в гомогенные.

Воспользуемся информацией о модели камеры и найдем гомогенные координаты точек.

Пусть начало координат находится в точки фокуса, а центр прямоугольника имеет координаты:

$$\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

Тогда на изображение ось OX будет направлена вправо, а ось OY вниз.

Рассмотрим треугольник образованный множеством точек изображения с координатой 0 по OX и точкой фокуса. Заметим, что имеет место следующее равенство:

$$\frac{v_{pix} - Height/2}{Height/2} = \frac{v}{MAX_y}$$

, где v — реальная координата по оси OY , а MAX_y — максимальная.

Здесь мы переводим v_{pix} из промежутка $[0, Height]$ в промежуток $[-Height/2, Height/2]$, а затем записываем равенство отношений пиксельных координат и реальных.

Отсюда получаем:

$$v = (v_{pix} - Height/2) \frac{MAX_y}{Height/2}$$

Теперь заметим, что:

$$MAX_y = \tan\left(\frac{\alpha_{Height}}{2}\right)$$

Так как мы принял, что изображение находится на расстоянии 1 от точки фокуса. А следовательно:

$$v = (v_{pix} - Height/2) \frac{\tan\left(\frac{\alpha_{Height}}{2}\right)}{Height/2}$$

Пусть F - фокусное расстояние в пикселях. В таком случае:

$$\frac{Height/2}{F} = \tan\left(\frac{\alpha_{Height}}{2}\right) \Leftrightarrow \frac{1}{F} = \frac{\tan\left(\frac{\alpha_{Height}}{2}\right)}{Height/2}$$

Получается, что вычислив $\frac{1}{F}$ один раз, мы можем с легкостью переводить координаты:

$$v = (v_{pix} - Height/2) \frac{1}{F}$$

По аналогии с этим можно получить и формулу для второй координаты:

$$u = (u_{pix} - Width/2) \frac{1}{F}$$

, где u реальная координата по оси OX .

Итого мы нашли гомогенные координаты:

$$P_H := \begin{pmatrix} u \\ v \\ 1 \end{pmatrix}$$

2. Функция минимизации.

Чего мы хотим? Мы хотим найти матрицу трансформации, соответствующую такому положению камеры, что трехмерные точки при проецировании на изображение совпадали бы с реальными позициями этих точек.

Рассмотрим идеальную ситуацию, когда проекции совпадают с реальными положениями. Это значит, что имеет место равенство:

$$\lambda P_H = TX_H$$

, где λ — некий коэффициент, T — матрица трансформации, а $X_H = \begin{pmatrix} X \\ 1 \end{pmatrix}$

А это условие можно записать и через векторное произведение:

$$P_H \times TX_H = 0$$

Введем кососимметричную матрицу P :

$$P := \begin{pmatrix} 0 & -1 & v \\ 1 & 0 & -u \\ -v & u & 0 \end{pmatrix}$$

Тогда предыдущее равенство переписывается так:

$$PTX_H = 0$$

Исходя из этого нам нужно подобрать такую матрицу трансформации чтобы сумма значений:

$$(PTX_H) \cdot (PTX_H) \rightarrow \min$$

стремила к нулю.

3. Трансформации при малых углах.

Матрицу поворота можно записать, как произведение матриц поворота вокруг осей:

$$R = R_z(\omega_z)R_y(\omega_y)R_x(\omega_x) = \begin{pmatrix} \cos \omega_z & -\sin \omega_z & 0 \\ \sin \omega_z & \cos \omega_z & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos \omega_y & 0 & \sin \omega_y \\ 0 & 1 & 0 \\ -\sin \omega_y & 0 & \cos \omega_y \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \omega_x & -\sin \omega_x \\ 0 & \sin \omega_x & \cos \omega_x \end{pmatrix}$$

Запишем разложение \sin и \cos в ряд до $\mathcal{O}(x^2)$:

$$\sin x = x + \mathcal{O}(x^2), \cos x = 1 + \mathcal{O}(x^2)$$

Тогда при достаточно малых углах (до ≈ 20 градусов) можно линеаризовать матрицу поворота:

$$R = \begin{pmatrix} 1 & -\omega_z & 0 \\ \omega_z & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & \omega_y \\ 0 & 1 & 0 \\ -\omega_y & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -\omega_x \\ 0 & \omega_x & 1 \end{pmatrix} = \begin{pmatrix} 1 & -\omega_z & \omega_y \\ \omega_z & 1 & -\omega_x \\ -\omega_y & \omega_x & 1 \end{pmatrix}$$

Получается, что матрица трансформации будет иметь вид:

$$T = \left(\begin{array}{ccc|c} 1 & -\omega_z & \omega_y & t_x \\ \omega_z & 1 & -\omega_x & t_y \\ -\omega_y & \omega_x & 1 & t_z \end{array} \right)$$

4. Нахождение малых углов.

Во 2ом пункте мы выяснили, что нужно минимизировать сумму значений:

$$(PTX_H) \cdot (PTX_H) \rightarrow \min$$

Давайте повнимательнее посмотрим на то, что здесь написано:

$$TX_H = \left(\begin{array}{ccc|c} 1 & -\omega_z & \omega_y & t_x \\ \omega_z & 1 & -\omega_x & t_y \\ -\omega_y & \omega_x & 1 & t_z \end{array} \right) \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix} = \begin{pmatrix} x - \omega_z y + \omega_y z + t_x \\ y + \omega_z x - \omega_x z + t_y \\ z - \omega_y x + \omega_x y + t_z \end{pmatrix} + \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

Пусть:

$$G = \begin{pmatrix} 1 & 0 & 0 & 0 & z & -y \\ 0 & 1 & 0 & -z & 0 & x \\ 0 & 0 & 1 & y & -x & 0 \end{pmatrix}, \mu = \begin{pmatrix} t_x \\ t_y \\ t_z \\ \omega_x \\ \omega_y \\ \omega_z \end{pmatrix}$$

Используя введенные обозначения можем переписать равенство:

$$TX_H = G\mu + X$$

Теперь перепишем функцию минимизации из 2го пункта:

$$(PG\mu + PX) \cdot (PG\mu + PX) \rightarrow \min$$

Для нахождения точки минимума возьмем производную и приравняем ее к 0 (числовые коэффициенты опущены):

$$G^T P^T (PG\mu + PX) = 0$$

Получаем простое уравнение:

$$G^T P^T PG\mu = -G^T P^T PX$$

Такое уравнение нетрудно решить. Для учета всех точек достаточно сложить уравнения:

$$(\sum G^T P^T PG)\mu = -\sum G^T P^T PX$$

5. Получение матрицы трансформации для малых углов.

В предыдущем пункте мы нашли углы. Для восстановления матрицы трансформации по этим углам воспользуемся следующей формулой:

$$R = I + \frac{\sin \theta}{\theta} \omega_{\times} + \frac{1 - \cos \theta}{\theta^2} \omega_{\times}^2$$

, где I — единичная матрица, $\theta = \sqrt{\omega_x^2 + \omega_y^2 + \omega_z^2}$, а $\omega_{\times} = \begin{pmatrix} 0 & -\omega_z & \omega_y \\ \omega_z & 0 & -\omega_x \\ -\omega_y & \omega_x & 0 \end{pmatrix}$

6. Вычисление ответа.

На данный момент мы имеем алгоритм нахождения матрицы трансформации. Но он не дает точного результата, так как мы линеаризуем матрицы поворота. Для нахождения точного результата будем действовать итеративно. Пусть изначально матрица трансформации равна:

$$T = \left(\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{array} \right)$$

То есть она никак не трансформирует точки. Далее придерживаемся следующего алгоритма:

а) Переведем точки в систему координат камеры, соответствующей текущей трансформации:

$$X' = TX$$

б) Найдем новую матрицу трансформации T' , используя в качестве исходных координат точек, координаты X' .

с) Пересчитаем текущую матрицу трансформации следующим образом:

$$T = T'T$$

Для перемножения матриц, к каждой снизу дописывается строка $(0 \ 0 \ 0 \ 1)$.

д) Перейдем в пункт а).

Таким образом, шаг за шагом, мы будем приближаться к верному ответу. Сходится это довольно быстро, так что 20 итераций хватает с лихвой.

Мы ставили задачу сопоставить точки на изображение и проекции трехмерных точек. Эту задачу можно сформулировать и по-другому:

Точка на изображение задает луч, исходящий из точки фокуса камеры. Нужно найти такое положение камеры, в котором лучи проходили бы через свои точки в трехмерном пространстве.

Однако, важно понимать следующее: наш метод устроен так, что он работает не с лучами, а с прямыми. А значит на кадр может спроецироваться точка, которая в действительности не видна, которая находится на дополнении луча до прямой.

В связи с этим может сложиться ситуация, при которой ни одна из точек не видна. Поэтому после нахождения трансформации необходимо проверить, верно ли направлена камера, и если нет, перевернуть ее и вновь посчитать трансформацию, используя трансформацию перевернутой камеры в качестве стартовой.

Можно также отметить, что из-за неидеальности камер такая перевернутая трансформация даст большую ошибку, чем трансформация верно установленной камеры. Исходя из этого можно сделать несколько попыток работы алгоритма описанного выше, но использовать не единичную матрицу, а случайную матрицу поворота. После чего просто выбрать трансформацию дающую наименьшую ошибку (что такое ошибка можно вспомнить во втором пункте).

Вышесказанное стоит применять, когда матрица трансформации ищется впервые. В случае же, когда мы работаем с потоком виде, стоит использовать матрицу трансформации предыдущего кадра, так как она близка к искомой. Это может значительно ускорить работу и избавить от необходимости лишних проверок.