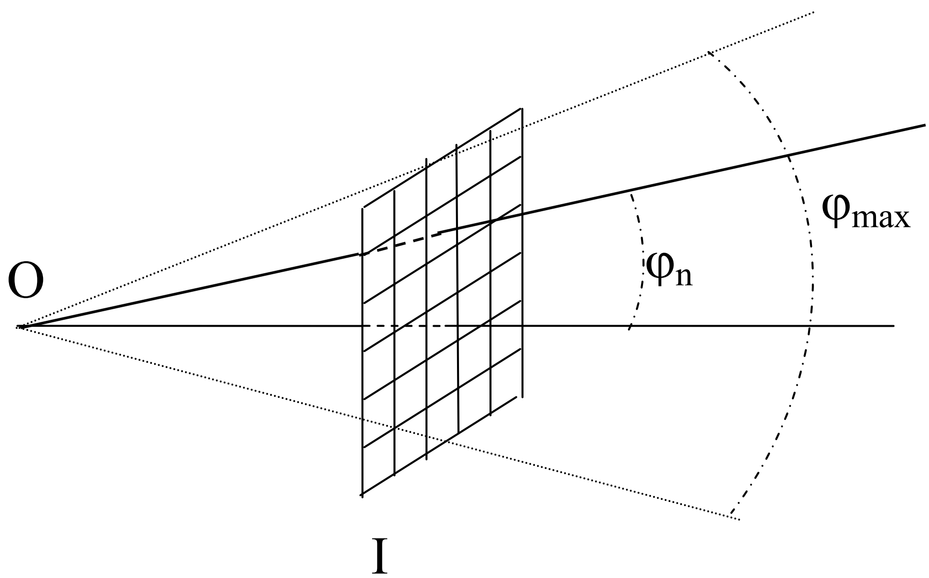
К вопросу о точности восстановления трехмерной сцены при стерео измерениях.

Вначале отметим два обстоятельства. Во-первых, задача восстановления трехмерных сцен, определяется исключительно отношениями размеров или углами, от абсолютных размеров ничего не зависит. Понять это легко, обращаясь к старым фильмам. В них часто показывали весьма масштабные сцены: сражающиеся эскадры, грандиозные постройки, витязь, сражающийся с огромным драконом. Все эти сцены снимались на столе, на моделях. Иногда, для субъективного восприятия масштабов, туда потом наложением помещали изображение человека или иного объекта с субъективно понятными размерами. Субъективно, сцены из такого фильма воспринимались как содержащие очень крупные объекты, несмотря на миниатюрный размер моделей, никакого противоречия не ощущалось. Другой пример – фраза, которую говорит Малыш из одноименной повести Стругацких : «…раньше корабль был огромный, потом съежился…» - но это не корабль уменьшился, а Малыш вырос.

Другое соображение – это то, что при постановке эксперимента по восстановлению трехмерных сцен надо оставаться в рамках разумного. Это означает, что количество используемых видео камер или количество обнаруженных точек на объекте не может составлять миллионы. Обычно камер 2 или больше, а точек 10-100. Это приводит к тому, что не следует ожидать какого-либо драматического увеличения точности измерений за счет накопления статистики. Точность результатов будет определяться заложенными базовыми принципами измерений. В задаче восстановления трехмерных сцен по набору цифровых изображений практически существует только один принцип, известный еще древним Египтянам. – триангуляция, восстановление высоты треугольника по его стороне и двум прилежащим углам. С точностью до множителя, не сильно отличающегося от единицы, точность измерений с помощью любого алгоритма восстановления будет определяться решением этой классической задачи.

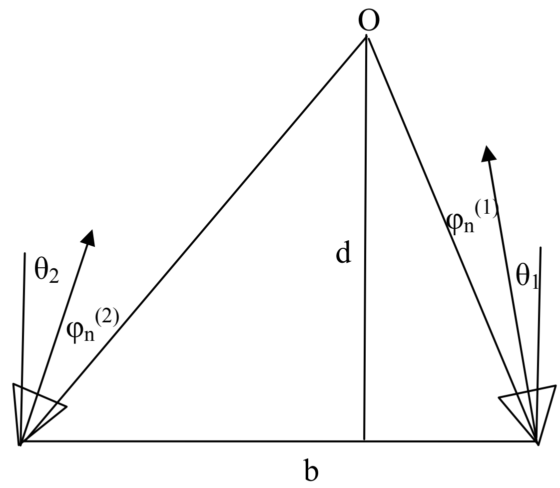
Рассмотрим задачу триангуляции подробнее. Начнем рассмотрение с цифровой камеры, как главного измерительного прибора.



Камера представляет собой матрицу фоточувствительных элементов, помещенную в фокальной плоскости объектива , на которой формируется изображение. На рис.1. плоскость изображения I (матрица) изображена впереди оптического центра O камеры, так как это не влияет на угловые зависимости, но упрощает запись формул и рисунки. Камера характеризуется максимальным углом обзора φ_{max} и количеством пикселей в строке N и столбце M~N, т.е количеством фоточувствительных элементов матрицы MxN. Размер каждого фоточувствительного элемента одинаков и они образуют прямоугольную сетку. Это обеспечивается с высокой точностью технологиями изготовления микросхем. Направление луча, проходящего через n-й пиксель, $N/2<n<N/2$ задается углом φ_n, который может быть выражен как

$$tg\varphi_n = \frac{tg\dfrac{\varphi_{\max}}{2}}{N/2}n$$

Перейдем теперь к рассмотрению задачи триангуляции.



Камеры расположены на расстоянии b друг от друга (стерео база). Направления оптичиеских осей камер задаются углами θ_1 и θ_2 относительно перпендикуляра к направлению b. Камеры видят точку O объекта соответственно в пикселе n_1 и n_2 направления на эту точку задаются углами $\varphi_{n_1}^{(1)}$ и $\varphi_{n_2}^{(2)}$ соответственно. Расстояние до точки O от базовой линии b равно d. Обозначая расстояния от первой и второй камер до точки O через a_1 и a_2, можно записать:

$$d = a_1 \cos(\theta_1 + \varphi_{n_1}^{(1)})$$

$$d = a_2 \cos(\theta_2 + \varphi_{n_2}^{(2)})$$

$$b = a_1 \sin(\theta_1 + \varphi_{n_1}^{(1)}) + a_2 \sin(\theta_2 + \varphi_{n_2}^{(2)})$$

Избавляясь в этих формулах от a_1 и a_2, получим:

$$\frac{b}{d} = tg(\theta_1 + \varphi_{n_1}^{(1)}) + tg(\theta_2 + \varphi_{n_2}^{(2)}) = \frac{tg(\theta_1) + tg(\varphi_{n_1}^{(1)})}{1 + tg(\theta_1)tg(\varphi_{n_1}^{(1)})} + \frac{tg(\theta_2) + tg(\varphi_{n_2}^{(2)})}{1 + tg(\theta_2)tg(\varphi_{n_2}^{(2)})}$$

В случае, если погрешность определения координат точки на изображении составляет 1 пиксель, углы $\varphi_{n_1}^{(1)}$ и $\varphi_{n_2}^{(2)}$ имеют погрешность:

$$\Delta\varphi_{n_1}^{(1)} = \frac{tg\,\frac{\varphi_{\max}^{(1)}}{2}}{N_1/2}$$

поэтому погрешность определения величины b/d равна:

$$\Delta\frac{b}{d} = \frac{b\Delta d}{d(d+\Delta d)} \approx \frac{b\Delta d}{d^2} = tg(\theta_1 + \varphi_{n_1}^{(1)} + \Delta\varphi_{n_1}^{(1)}) - tg(\theta_1 + \varphi_{n_1}^{(1)}) + tg(\theta_2 + \varphi_{n_2}^{(2)}) + tg(\theta_2 + \varphi_{n_2}^{(2)} + \Delta\varphi_{n_2}^{(2)}) =$$

$$= \frac{\sin(\Delta\varphi_{n_1}^{(1)})}{\cos(\theta_1 + \varphi_{n_1}^{(1)} + \Delta\varphi_{n_1}^{(1)})\cos(\theta_1 + \varphi_{n_1}^{(1)})} + \frac{\sin(\Delta\varphi_{n_2}^{(2)})}{\cos(\theta_2 + \varphi_{n_{21}}^{(2)} + \Delta\varphi_{n_2}^{(2)})\cos(\theta_2 + \varphi_{n_2}^{(2)})} \approx$$

$$\approx \frac{\Delta\varphi_{n_1}^{(1)}}{\cos^2(\theta_1 + \varphi_{n_1}^{(1)})} + \frac{\Delta\varphi_{n_2}^{(2)}}{\cos^2(\theta_2 + \varphi_{n_2}^{(2)})}$$

Здесь было использовано приближение, что погрешность угла мала по сравнение с остальными углами, т.е. разрешение камеры достаточно велико. Полагая, что обе камеры имеют одинаковое разрешение и максимальный угол обзора отсюда можно записать:

$$\frac{b\Delta d}{d^2} \approx \frac{tg\,\frac{\varphi_{\max}}{2}}{N\;/2}\left(\frac{1}{\cos^2(\theta_1 + \varphi_{n_1}^{(1)})} + \frac{1}{\cos^2(\theta_2 + \varphi_{n_2}^{(2)})}\right)$$

Выражение в скобках не может быть меньше, чем 2, поэтому

$$\frac{b\Delta d}{d^2} > \frac{tg\,\frac{\varphi_{\max}}{2}}{N}$$

Окончательно получаем, что ***<u>относительная погрешность определения расстояния не может быть лучше, чем задаваемая формулой:</u>***

$$\frac{\Delta d}{d} > \frac{1}{N}\frac{d}{b}tg\frac{\varphi_{\max}}{2}$$

Например, если угол зрения камеры составляет 45°, то $tg\frac{\varphi_{\max}}{2} = 0.4$ и для камеры разрешением 1000 пикселей в строке (столбце) при расстоянии между камерами 1м и расстоянии до объекта 1м погрешность составит 1/1000*0.4=0.0004м=0.4мм. Если же расстояние между камерами равно 1м, а расстояние до объекта равно 10м, то погрешность измерения расстояния составит уже 0.04м=4см , при увеличении расстояния до объекта вдовое (до 20 м) точность упадет до 16см. Вообще говоря как видно из полученной формулы, погрешность пропорциональна квадрату расстояния до объекта.

Из этой же формулы можно получить оценку диапазона расстояний вне которых измерения вообще не возможны, это расстояние $d_{\max}$, начиная с которого погрешность превышает само расстояние.

$$\frac{b}{d_{\max}} = \frac{1}{N}tg\frac{\varphi_{\max}}{2}$$

Т.е. если отношение базы к расстоянию образует угол, сопоставимый или меньший, чем угловое разрешение камеры, то стерео эффект отсутствует и измерение расстояния невозможно принципиально.

Приведенные оценки максимально оптимистичные и получены в предположении наилучшей из возможных геометрии съемки. При использовании различных алгоритмов стерео имеются факторы, приводящие как к улучшению так и ухудшению полученных оценок. К некоторому улучшению ведет возможность определения координат точки на изображении с субпиксельной точностью. С другой стороны, возможность нахождения ошибочных соответствий, шумы, радиальная дисторсия камеры, погрешности калибровки и неточность фокусировки приводят к ухудшению полученных оценок. В сумме следует ожидать, что полученные оценки будут выполняться с точностью до некоторого множителя, который не будет отличаться от 1 больше чем на порядок, и , скорее всего, будет около единицы.