对极几何与三角测量整理

对于同一相机从不同角度拍摄的两幅图像，对其进行特征提取和匹配，就得到了若干对匹配特征；根据这些图像中二维坐标点的匹配关系，就可以恢复出两幅图像间摄像机的旋转和平移（即位姿的变化）。这是一个 2D 点对（即 2D-2D）间的运动估计问题，通常使用对极几何的方式得到相机相对运动的位姿。在 SLAM 实际应用中，通常被用于单目相机的初始化、双目相机的标定等方面。

一、 基本思想

如图 1，两个视角下的摄像机光心分别为$\boldsymbol{O}_1$和$\boldsymbol{O}_2$，图像所在的平面分别为$\boldsymbol{I}_1$和$\boldsymbol{I}_2$，三维空间中的点$\boldsymbol{P}$在两幅图像的投影分别为$\boldsymbol{p}_1$和$\boldsymbol{p}_2$，连接点$\boldsymbol{P}$、$\boldsymbol{O}_1$和$\boldsymbol{O}_2$所构成的平面为极平面（Epipolar plane），而连线$\boldsymbol{O}_1\boldsymbol{O}_2$则称为基线（Baseline），基线与成像平面的交点$\boldsymbol{e}_1$、$\boldsymbol{e}_2$为极点（Epipoles），连线$\boldsymbol{p}_1\boldsymbol{e}_1$、$\boldsymbol{p}_2\boldsymbol{e}_2$（即极平面与两个成像平面的交线）可记作$\boldsymbol{l}_1$、$\boldsymbol{l}_2$，称为极线（Epipolar line）。这样的几何约束关系称为对极几何约束。

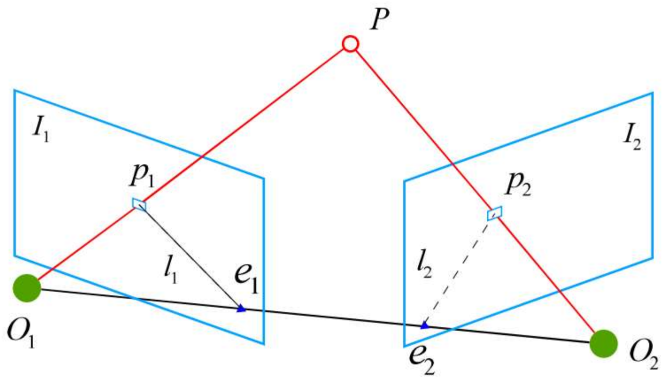

图 1 对极几何约束

我们可以通过这些点的共面关系恢复出摄像机间的相对位姿$\boldsymbol{R}$、$\boldsymbol{t}$。

而另一方面，由于极平面的存在，一幅图像中的特征的对应匹配点必然会出现在另一幅图的对应极线中。因此，在已经确定位姿关系的两幅图像中，匹配特征不必在整幅图像中搜索对应点，只需在对应的极线上搜索即可，这样就将二维搜索转化为一维搜索，在双目视觉 SLAM 中借助这一特性，可大大降低双目特征匹配时的计算量，同时也可提高特征匹配的准确率。

二、 基础矩阵和本征矩阵

两幅图像的位姿关系可由本征矩阵（Essencial Matrix）$\boldsymbol{E}$表述，本征矩阵仅包含二者之间的位姿信息；而基础矩阵（Fundamental Matrix）$\boldsymbol{F}$则在位姿信息的基础上包含了摄像机的内参数$\boldsymbol{K}$。本征矩阵在归一化的物理坐标系下操作，而基础矩阵在图像坐标系下进行操作。接下来从数学角度对$\boldsymbol{E}$和$\boldsymbol{F}$进行推导。

2.1 本征矩阵

已知三维空间中的点$\boldsymbol{P}=(X,Y,Z)^T$（在第一帧坐标系下），该点在两幅图像的投影分别为$\boldsymbol{p}_1$和$\boldsymbol{p}_2$，则有

$$s\boldsymbol{p}_1=\boldsymbol{KP}$$
$$s\boldsymbol{p}_2=\boldsymbol{K}(\boldsymbol{RP}+\boldsymbol{t})$$

取两个像素点$\boldsymbol{p}_1$、$\boldsymbol{p}_2$的归一化平面坐标为$\boldsymbol{x}_1$和$\boldsymbol{x}_2$，那么

$$\boldsymbol{x}_1=\boldsymbol{K}^{-1}\boldsymbol{p}_1$$

$$\boldsymbol{x}_2 = \boldsymbol{K}^{-1}\boldsymbol{p}_2$$

代入之后则有

$$\boldsymbol{x}_2 = \boldsymbol{R}\boldsymbol{x}_1 + \boldsymbol{t}$$

移项得

$$\boldsymbol{x}_2 - \boldsymbol{t} = \boldsymbol{R}\boldsymbol{x}_1$$

对于同一平面上的向量$\boldsymbol{x}$和$\boldsymbol{a}$，以及该平面的法向量$\boldsymbol{n}$，有约束条件

$$\boldsymbol{n} \cdot (\boldsymbol{x} - \boldsymbol{a}) = 0$$

而$\boldsymbol{x}_2$和$\boldsymbol{t}$都在极平面内，取$\boldsymbol{t} \times \boldsymbol{x}_2$为法向量，那么

$$(\boldsymbol{t} \times \boldsymbol{x}_2)^T(\boldsymbol{x}_2 - \boldsymbol{t}) = 0$$

代入

$$(\boldsymbol{t} \times \boldsymbol{x}_2)^T\boldsymbol{R}\boldsymbol{x}_1 = 0$$

化为反对称矩阵

$$([\boldsymbol{t}]_\times \boldsymbol{x}_2)^T\boldsymbol{R}\boldsymbol{x}_1 = 0$$

括号内展开

$$\boldsymbol{x}_2^T[\boldsymbol{t}]_\times^T\boldsymbol{R}\boldsymbol{x}_1 = 0$$
$$-\boldsymbol{x}_2^T[\boldsymbol{t}]_\times\boldsymbol{R}\boldsymbol{x}_1 = 0$$

去掉负号

$$\boldsymbol{x}_2^T[\boldsymbol{t}]_\times\boldsymbol{R}\boldsymbol{x}_1 = 0$$

最终可得到

$$\boldsymbol{x}_2^T\boldsymbol{E}\boldsymbol{x}_1 = 0$$

即本征矩阵$\boldsymbol{E}$为

$$\boldsymbol{E} = [\boldsymbol{t}]_\times\boldsymbol{R}$$

由$\operatorname{rank}([\boldsymbol{t}]_\times) = 2$易知，$\boldsymbol{E}$的秩为 2，同样是秩亏矩阵；$[\boldsymbol{t}]_\times$为反对称矩阵，$\boldsymbol{R}$为正交矩阵，因此$\boldsymbol{E}$的两个非零特征值相等。

由于对极约束是等于 0 的等式，因此对于$\boldsymbol{E}$而言，乘以任意非零常数后性质不变，因此$\boldsymbol{E}$具有尺度不确定性。

而另一方面，得到矩阵$\boldsymbol{E}$的两个分量平移$\boldsymbol{t}$和旋转$\boldsymbol{R}$都具有 3 自由度，共 6 自由度，再由于其尺度不确定性去掉 1 自由度，可知$\boldsymbol{E}$共有 **5 自由度**。

2.2 基础矩阵

令单应性矩阵$\boldsymbol{H}_{3\times 4} = \boldsymbol{K}[\boldsymbol{R} \mid \boldsymbol{t}]$，则有

$$s\boldsymbol{p}_1 = \boldsymbol{H}_1\boldsymbol{P}$$
$$s\boldsymbol{p}_2 = \boldsymbol{H}_2\boldsymbol{P}$$

其中

$$\boldsymbol{H}_1 = \boldsymbol{K}[\boldsymbol{I} \mid \boldsymbol{0}]$$
$$\boldsymbol{H}_2 = \boldsymbol{K}[\boldsymbol{R} \mid \boldsymbol{t}]$$

显然$\operatorname{rank}\boldsymbol{H} = 3$，那么矩阵$\boldsymbol{H}$的核空间（即右零空间）为一维空间，即

$$\boldsymbol{H} \cdot \boldsymbol{C} = \boldsymbol{0}$$

其中$\boldsymbol{C}$为4×1向量，可视作由齐次坐标表示的三维空间中的一点。

以$\boldsymbol{C}$为原点，与$\boldsymbol{P}$可确定一条射线，射线上的点$\boldsymbol{P}(\lambda)$可表示为（其中λ为比例系数）

$$\boldsymbol{P}(\lambda) = \lambda\boldsymbol{P} + (1-\lambda)\boldsymbol{C}$$

故

$$\begin{aligned}\boldsymbol{p}(\lambda) &= \lambda\boldsymbol{H}_1\boldsymbol{P} + (1-\lambda)\boldsymbol{H}_1\boldsymbol{C} \\ &= \lambda\boldsymbol{H}_1\boldsymbol{P} \\ &= \boldsymbol{p}_1\end{aligned}$$

也就是说，射线$\boldsymbol{CP}$上所有的点在第一幅图像上的投影都与点$\boldsymbol{P}$的投影$\boldsymbol{p}_1$重合，故而易知该射线通过摄像机 1 的光心；而点$\boldsymbol{P}$为空间中任意点，点$\boldsymbol{C}$是该射线上唯一固定点，可知点$\boldsymbol{C}$是摄像机 1 的光心$\boldsymbol{O}_1$。

反之，可求得通过光心$\boldsymbol{C}$和点$\boldsymbol{p}_1$的射线上的点可表示为

$$\boldsymbol{Q}(\lambda) = \boldsymbol{H}_1^{+}\boldsymbol{p}_1 + \lambda\boldsymbol{C}$$

其中$\boldsymbol{H}_1^{+}$为单应性矩阵的伪逆。

取该射线上两点$\boldsymbol{H}_1^{+}\boldsymbol{p}_1$和$\boldsymbol{C}$，分别投影到第二幅视图上，即为$\boldsymbol{H}_2\boldsymbol{H}_1^{+}\boldsymbol{p}_1$和$\boldsymbol{H}_2\boldsymbol{C}$，由之前的条件易知$\boldsymbol{H}_2\boldsymbol{C}$即为第二幅视图上的极点$\boldsymbol{e}_2$，$\boldsymbol{H}_2\boldsymbol{H}_1^{+}\boldsymbol{p}_1$在极线$\boldsymbol{l}_2$上，故而极线$\boldsymbol{l}_2$为

$$\begin{aligned}\boldsymbol{l}_2 &= \boldsymbol{e}_2 \times \boldsymbol{H}_2\boldsymbol{H}_1^{+}\boldsymbol{p}_1 \\ &= [\boldsymbol{e}_2]_\times \boldsymbol{H}_2\boldsymbol{H}_1^{+}\boldsymbol{p}_1\end{aligned}$$

而点$\boldsymbol{p}_2$也在极线$\boldsymbol{l}_2$上，则有

$$\boldsymbol{p}_2^T\boldsymbol{l}_2 = 0$$

故

$$\boldsymbol{p}_2^T[\boldsymbol{e}_2]_\times \boldsymbol{H}_2\boldsymbol{H}_1^{+}\boldsymbol{p}_1 = 0$$

即

$$\boldsymbol{p}_2^T\boldsymbol{F}\boldsymbol{p}_1 = 0$$

$$\boldsymbol{F} = [\boldsymbol{e}_2]_\times \boldsymbol{H}_2\boldsymbol{H}_1^{+}$$

将

$$\boldsymbol{H}_1 = \boldsymbol{K}[\boldsymbol{I} \mid \boldsymbol{0}]$$

$$\boldsymbol{H}_2 = \boldsymbol{K}[\boldsymbol{R} \mid \boldsymbol{t}]$$

$$\boldsymbol{H}_1^{+} = \begin{bmatrix}\boldsymbol{K}^{-1} \\ \boldsymbol{0}^T\end{bmatrix}$$

$$\boldsymbol{C} = \begin{bmatrix}\boldsymbol{0} \\ 1\end{bmatrix}$$

等代入，进而求出

$$\begin{aligned}\boldsymbol{e}_2 &= \boldsymbol{H}_2\boldsymbol{C} \\ &= \boldsymbol{K}[\boldsymbol{R} \mid \boldsymbol{t}]\begin{bmatrix}\boldsymbol{0} \\ 1\end{bmatrix} \\ &= \boldsymbol{K}\boldsymbol{t}\end{aligned}$$

及

$$\begin{aligned}\boldsymbol{F} &= [\boldsymbol{e}_2]_\times \boldsymbol{H}_2\boldsymbol{H}_1^{+} \\ &= [\boldsymbol{K}\boldsymbol{t}]_\times [\boldsymbol{R} \mid \boldsymbol{t}]\begin{bmatrix}\boldsymbol{K}^{-1} \\ \boldsymbol{0}^T\end{bmatrix} \\ &= [\boldsymbol{K}\boldsymbol{t}]_\times \boldsymbol{K}\boldsymbol{R}\boldsymbol{K}^{-1} \\ &= \boldsymbol{K}^{-T}[\boldsymbol{t}]_\times \boldsymbol{K}^{-1}\boldsymbol{K}\boldsymbol{R}\boldsymbol{K}^{-1} \\ &= \boldsymbol{K}^{-T}[\boldsymbol{t}]_\times \boldsymbol{R}\boldsymbol{K}^{-1}\end{aligned}$$

即，基础矩阵$\boldsymbol{F}$为

$$\boldsymbol{F} = \boldsymbol{K}^{-T}[\boldsymbol{t}]_\times \boldsymbol{R}\boldsymbol{K}^{-1}$$

或

$$\boldsymbol{F} = \boldsymbol{K}^{-T}\boldsymbol{E}\boldsymbol{K}^{-1}$$

与本征矩阵类似，基础矩阵同样具有尺度不确定的特性，即矩阵相乘一个非零常数之后不影响结果；$\mathrm{rank}(\boldsymbol{F}) = 2$，但是$\boldsymbol{F}$的两个非零特征值不相等。由

$$\boldsymbol{F} = [\boldsymbol{e}_2]_\times \boldsymbol{H}_2\boldsymbol{H}_1^{+}$$

其中$\boldsymbol{H}_2\boldsymbol{H}_1^{+}$可视作单应变换$\boldsymbol{H}_\pi$，有 8 自由度，而$\mathrm{rank}(\boldsymbol{F}) = 2$，为$\boldsymbol{F}$增加了一个约束，因此$\boldsymbol{F}$

有 **7 自由度**。

三、 求解基础矩阵

相比于本征矩阵，基础矩阵的约束是针对匹配点对的像素坐标的，相对而言更易获得，因此更多使用基础矩阵来估计两幅视角的位姿变化。由于基础矩阵$\boldsymbol{F}$具有 7 自由度，所以理论上通过 7 对对应点就可以解出$\boldsymbol{F}$，但是这一算法是基于$\boldsymbol{F}$秩为 2 的约束下的，共有 3 组不同的解。由于实际的图像中会有至少几十对匹配的特征，因此使用 7 点法的意义不大，更多的还是采用八点法的方式，通过线性方程组求解基础矩阵。而当所提供的匹配对数超过八对时，则可以可以求其最小二乘解。在 OpenCV 的相关函数中，除了 7 点算法和 8 点算法外，为增强算法的鲁棒性，还提供了 RANSAC、LMedS 等选项。

接下来简单介绍一下八点法求解基础矩阵$\boldsymbol{F}$的算法。假设有一对匹配的特征，他们的坐标分别为

$$\boldsymbol{p}_1 = [u_1 \quad v_1 \quad 1]^T$$
$$\boldsymbol{p}_2 = [u_2 \quad v_2 \quad 1]^T$$

用f_i表示$\boldsymbol{F}$中的每个元素，那么基础矩阵的约束关系可以表示为

$$[u_1 \quad v_1 \quad 1]\begin{bmatrix} f_1 & f_2 & f_3 \\ f_4 & f_5 & f_6 \\ f_7 & f_8 & f_9 \end{bmatrix}\begin{bmatrix} u_2 \\ v_2 \\ 1 \end{bmatrix} = 0$$

将$\boldsymbol{F}$展开为向量

$$\boldsymbol{f} = [f_1 \quad f_2 \quad f_3 \quad f_4 \quad f_5 \quad f_6 \quad f_7 \quad f_8 \quad f_9]^T$$

那么等式可以化为

$$[u_1u_2 \quad u_1v_2 \quad u_1 \quad v_1u_2 \quad v_1v_2 \quad v_1 \quad u_2 \quad v_2 \quad 1]\cdot\boldsymbol{f} = 0$$

对于其他的匹配特征也采用同样的方法表示，将所有方程放到一起，可得到一个线性方程组(其中u^i、v^i表示第i对匹配特征的坐标)

$$\begin{bmatrix} u_1^1u_2^1 & u_1^1v_2^1 & u_1^1 & v_1^1u_2^1 & v_1^1v_2^1 & v_1^1 & u_2^1 & v_2^1 & 1 \\ u_1^2u_2^2 & u_1^2v_2^2 & u_1^2 & v_1^2u_2^2 & v_1^2v_2^2 & v_1^2 & u_2^2 & v_2^2 & 1 \\ u_1^3u_2^3 & u_1^3v_2^3 & u_1^3 & v_1^3u_2^3 & v_1^3v_2^3 & v_1^3 & u_2^3 & v_2^3 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ u_1^8u_2^8 & u_1^8v_2^8 & u_1^8 & v_1^8u_2^8 & v_1^8v_2^8 & v_1^8 & u_2^8 & v_2^8 & 1 \end{bmatrix}\begin{bmatrix} f_1 \\ f_2 \\ f_3 \\ f_4 \\ f_5 \\ f_6 \\ f_7 \\ f_8 \\ f_9 \end{bmatrix} = \boldsymbol{0}$$

求解该方程组即可得到$\boldsymbol{F}$中各个参数的解。若该方程大于 8 行，即多于 8 组匹配的特征，则求其最小二乘解。

在实际计算中，由于噪声、数值舍入等原因造成误差，常常导致求出的基础矩阵不稳定。因此在解方程之前，先对匹配的坐标进行归一化，之后再行计算。

在 OpenCV 中使用的是各向同性归一化算法，即归一化之后各点到坐标原点的距离是$\sqrt{2}$，线性变换的过程为

$$\boldsymbol{p}_1' = \boldsymbol{T}_1\boldsymbol{p}_1$$
$$\boldsymbol{p}_2' = \boldsymbol{T}_2\boldsymbol{p}_2$$

由$\boldsymbol{p}_1'$和$\boldsymbol{p}_2'$（以及很多对其他对应点）求出解，得到$\boldsymbol{F}'$，进而将其还原为基础矩阵$\boldsymbol{F}$

$$\boldsymbol{F} = \boldsymbol{T}_2^T\boldsymbol{F}'\boldsymbol{T}_1$$

由之前的推导过程易知，第一幅视图中的每一点在第二幅视图中都有不同的极线与之对应，$\boldsymbol{p}_1$所对应的极线$\boldsymbol{l}_2$为

$$\boldsymbol{l}_2 = \boldsymbol{F}\boldsymbol{p}_1$$

在 OpenCV 的计算极线的函数中，输出的极线以(a, b, c)的形式进行编码，并且在输出前将其归一化，使得$a^2 + b^2 = 1$。

四、 三角测量

三角测量作为一种历史悠久的距离估算方法，广泛应用于航海、天文、测绘等领域。在得到相机的相对运动之后，可以使用这一方法，结合位姿来计算特征的深度，或者说是特征在空间中的三维坐标信息。

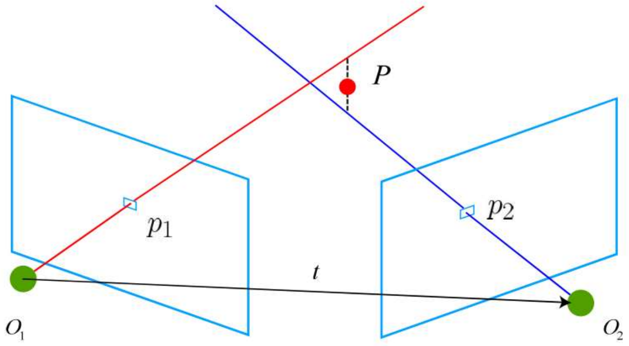

图 2 三角测量求取特征深度

如图 2，已知点的像素坐标$\boldsymbol{p}_i$、相机内参$\boldsymbol{K}$、以及摄像机相对位姿$\boldsymbol{R}$和$\boldsymbol{t}$，理论上直线$\boldsymbol{O}_1\boldsymbol{p}_1$和$\boldsymbol{O}_2\boldsymbol{p}_2$在场景中会与点$\boldsymbol{P}$相交，但是由于噪声的影响，两条直线往往无法相交。暂时不考虑误差，可以列出方程

$$s_1\boldsymbol{K}^{-1}\boldsymbol{p}_1 = s_2\boldsymbol{R}\boldsymbol{K}^{-1}\boldsymbol{p}_2 + \boldsymbol{t}$$

由像素坐标$\boldsymbol{p}_i$和相机内参$\boldsymbol{K}$，可以得出归一化坐标$\boldsymbol{x}_i$

$$\boldsymbol{p}_i = \boldsymbol{K}\boldsymbol{x}_i$$

方程可化简为

$$s_1\boldsymbol{x}_1 = s_2\boldsymbol{R}\boldsymbol{x}_2 + \boldsymbol{t}$$

其中s_i是方程的未知数，即需要计算的特征在相机坐标系下的深度（Z坐标），在两侧同时左乘$\boldsymbol{x}_1$的反对称矩阵$\boldsymbol{x}_1^\wedge$

$$s_1\boldsymbol{x}_1^\wedge\boldsymbol{x}_1 = s_2\boldsymbol{x}_1^\wedge\boldsymbol{R}\boldsymbol{x}_2 + \boldsymbol{x}_1^\wedge\boldsymbol{t}$$

这样方程左侧为 0，可以直接求出s_2的值，进而s_1的值也很容易求出。

在实际应用中，由于误差的存在，一般很难求出其精确解，因此常常会求其最小二乘解。

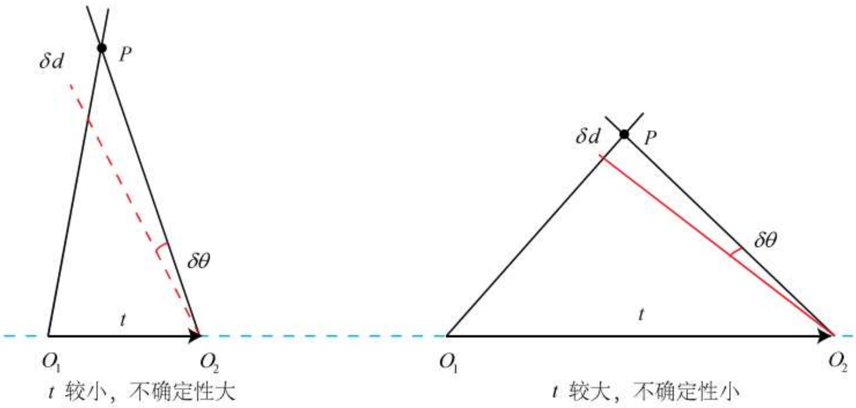

图 3 三角测量的矛盾

如图 3，三角测量的一个很重要的存在前提就是平移，如果没有平移，三角形无法成立，三角测量也就无从谈起了。而对于三角形而言，**增大平移**使得同样的角度误差$\delta\theta$下的深度误差δd更小，即三角测量**更加精确**。但是与此矛盾的是，增大平移的同时也会使得视角变化较大，摄像机采集到的图像因而产生较大的变化，造成**特征匹配困难**。这就是三角测量中存在的矛盾：大平移量易导致匹配失效，而小平移量易导致较大误差。

参考

[1] 高翔，张涛，刘毅，等. 视觉 SLAM 十四讲[M]. 北京：电子工业出版社, 2017.

[2] Kaehler A, Bradski G. 学习 OpenCV3[M]. 阿丘科技团队，译. 北京:清华大学出版社, 2018.

[3] Hartley R, Zisserman A. 计算机视觉中的多视图几何[M].韦穗，杨尚骏等，译. 合肥:安徽大学出版社, 2002.

[4] kokerf. 基本矩阵的基本解法之 8 点算法[EB/OL]. 2017[2019-3-7]. https://blog.csdn.net/kokerf/article/details/72630863.

[5] Brook_icv. SLAM 入门之视觉里程计(4)：基础矩阵的估计[EB/OL]. 2018[2019-3-7]. https://www.cnblogs.com/wangguchangqing/p/8214032.html.

[6] 语冰. 为什么本质矩阵 5 自由度，基础矩阵 7 自由度，单应矩阵 8 自由度？[EB/OL]. 2018[2019-3-7]. https://www.zhihu.com/question/270431743.