

PanoVideo: 摄像机运动建模及从视频估计 摄像机运动参数的一种方法*

张茂军 胡晓峰 库锡树

(国防科技大学七系, 长沙 410073)

摘要 通过给摄像机平移、旋转与变焦等运动建模,并把运动模型与基于象素点亮度变化的方法相结合来估计摄像机运动参数。用得到的运动参数可以把视频构造成一幅全景图,全景图可广泛应用于视频压缩与检索。实验表明,该方法可成功地应用于视频会议系统中的视频压缩与视频检索。

关键词 摄像机运动估计,全景图,基于内容检索,视频压缩

1 引言

多媒体研究中有两个领域备受关注:基于内容检索与视频的超低比特率压缩。对视频进行基于内容的检索是一件艰难的事情。早期,检索视频主要依靠手工给视频的关键画面定义关键词,由检索关键词来定位到用户所需的画面。Otsuji^[1]提出自动场景检测技术,通过该技术把视频自动地分割成一系列片段,由于每个片段中场景变化较小,可以用每个片段的第一帧(即关键帧)代表该片段。于是检索视频的任务被简化为检索视频中的关键帧。另一类方法是利用视频中的运动信息检索视频。Teodosio^[2]提出 Salient Still 方法,通过估计摄像机运动参数,把视频的连续多帧画面拼接成全景图(Panoramic Image)。一幅全景图往往可以代表连续的几十帧甚至几百帧图象。这样,检索视频时只需从全景图中检索所需画面。这种方法效率很高。

利用摄像机的运动参数还可以实现很高的视频压缩比,在文献[3]中,Irani 通过估计摄像机的运动参数构造出 Mosaic 图象,并用 Mosaic 图象补偿各帧的运动图象。在有些场合应用这种方法可以把视频压缩到 64kbps 以下。

估计摄像机运动参数的关键是摄像机运动建模。摄像机运动模型简单与否直接关系到运动估计的效率。[3]中把摄像机的运动建成一个 8 参数的运动模型。[2]中采用 6 参数的摄像机运动模型。本文在详细分析摄像机工作原理的基础上,构造了摄像机的 4 参数三维运动模型(T_x, T_y, W_z, α),其中 T_x, T_y 分别是摄像机沿 X 轴与 Y 轴平移参数, W_z 是摄像机沿 Z 轴的旋转参数, α 反映摄像机的焦距变化。该模型具有无失真且速度快的优点。为估计大位移的摄像机运动及消除物体运动对估计过程的影响,文章中采用了构造图象塔法及选择迭代法。由于我们主要把上述方法应用于构造全景图,所以称之为 PanoVideo(即 Panoramic Video)。

2 摄像机运动模型

自然场景经过摄像机透视投影产生视频。为了说明这种投影关系,我们建立两个坐标系:摄像机坐标系与图象空间坐标系。摄像机坐标系是一个三维坐标系,原点位于摄像机棱镜的中心。 z 轴与棱镜表平面垂直。图象空间坐标系是一个与 z 轴垂直的平面,原点为 z 轴与图象空间坐标系的交点。图 1 说明了摄像机坐标系与图象空间坐标系间的关系,两坐

* 本项目得到“九五”国防预研(项目编号 9.5.5.5)的资助。

收稿日期:1996-12-09;收到修改稿日期:1997-01-20

标系原点间的距离等于摄像机的焦距 f 。

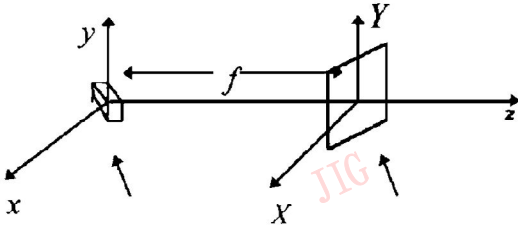


图 1 摄像机坐标系与图象空间坐标系的关系

Fig.1 The relationship between camera coordinate system and image space coordinate system

设某点在摄像机坐标系坐标为 (x, y, z) , 通过透视投影, 该点被投影到图象空间坐标系, 坐标为 (X, Y) 。 (x, y, z) 与 (X, Y) 间的关系为:

$$X = f \cdot \frac{x}{z}; \quad Y = f \cdot \frac{y}{z} \quad (1)$$

2.1 摄像机变焦

当摄像机焦距变化而位置不变时, 各点在摄像机坐标系中的坐标保持不变 (即 $s=s'$)。但它们在图象空间坐标系中的投影点 $S(X, Y)$ 变为:

$$X' = \frac{f'}{f} \cdot X; \quad Y' = \frac{f'}{f} \cdot Y \quad (2)$$

其中 f 为原焦距, f' 为变化后的焦距。定义焦距变化因子 $\alpha = \frac{f'}{f}$

2.2 摄像机平移与旋转

摄像机一般可以做 6 个自由度的运动: 平移 (T_x, T_y, T_z) , 旋转 (W_x, W_y, W_z) 。由附录的证明可知: 摄像机沿 z 轴向摄像机平移 T_z , 等效于摄像机不动而焦距由 f 增加到 $\frac{d}{d-T_z} f$, 其中, d 指自然场景到摄像机的距离。所以摄像机沿 z 轴的平移参数 T_z 可以用变焦因子 α 替代。

摄像机旋转也会导致象素点坐标的变化。这种变化可以由一个 3×3 矩阵描述。例如摄像机沿 x 轴旋转 W_x 角度, 坐标变化为:

$$(x', y', z') = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos W_x & \sin W_x \\ 0 & -\sin W_x & \cos W_x \end{bmatrix} \cdot (x, y, z)^T \quad (3)$$

由于视频序列中相邻两帧间的变化发生在极短的时间内 (典型情况是 1/30 秒), 所以可以认为帧间摄像机旋转角度 W_x 是很小的。因而, $\cos W_x \approx 1$ 且

$\sin W_x \approx W_x$, 帧图象中各象素点的位置变化为:

$$X' = f \cdot \frac{x'}{z'} = \frac{X}{1 - \frac{Y}{f} \cdot W_x}$$

$$Y' = f \cdot \frac{y'}{z'} = \frac{Y + f \cdot W_x}{1 - \frac{Y}{f} \cdot W_x}$$

由于 W_x 很小, 可以假定 $\frac{Y}{f} \cdot W_x$ 远小于 1, 因而有

$$X' = X, Y' = Y + f \cdot W_x \quad (4)$$

类似地, 当摄像机沿 y 轴旋转 W_y 时, 有:

$$X' = X + f \cdot W_y, Y' = Y \quad (5)$$

由式(4)、(5)可知, 摄像机沿 x 轴与 y 轴的旋转可以用摄像机的平移参数 (T_x, T_y) 近似。

设摄像机沿 z 轴旋转的角度为 W_z , 由文献[4], 得点在摄像机坐标系的坐标变化规律为:

$$(x', y', z')^T = \begin{bmatrix} 1 & -W_z & 0 \\ W_z & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot (x, y, z)^T$$

所以,

$$X' = f \cdot \frac{x'}{z'} = f \cdot \frac{x - W_z \cdot y}{z} = X - W_z \cdot Y \quad (6)$$

$$Y' = f \cdot \frac{y'}{z'} = f \cdot \frac{y + W_z \cdot x}{z} = Y + X \cdot W_z \quad (7)$$

这样, 当摄像机沿 X 轴、 Y 轴与 z 轴旋转与平移后, 图象中的象素点 (X, Y) 将按式(8)变为 (X', Y') :

$$\begin{bmatrix} X' \\ Y' \end{bmatrix} = \alpha \cdot \begin{bmatrix} 1 & W_z \\ W_z & 1 \end{bmatrix} \cdot \begin{bmatrix} X \\ Y \end{bmatrix} + \begin{bmatrix} T_x \\ T_y \end{bmatrix} \quad (8)$$

式中, α, W_z, T_x, T_y 为图象的 4 个全局运动参数。

3 摄像机运动参数估计公式

设点 $(X(t), Y(t))$ 表示点 S 于 t 时刻在图象坐标系中的坐标。在 $t+dt$ 时刻, 由式(8)得到点 S 的坐标为:

$$\begin{bmatrix} X(t+dt) \\ Y(t+dt) \end{bmatrix} = \alpha \cdot \begin{bmatrix} 1 & -W_z \\ W_z & 1 \end{bmatrix} \cdot \begin{bmatrix} X(t) \\ Y(t) \end{bmatrix} + \begin{bmatrix} T_x \\ T_y \end{bmatrix}$$

这样, 点 S 的运动速度可以表为:

$$\begin{bmatrix} \frac{dX}{dt} \\ \frac{dY}{dt} \end{bmatrix} = \begin{bmatrix} \alpha - 1 & -\alpha \cdot W_z \\ \alpha \cdot W_z & \alpha - 1 \end{bmatrix} \cdot \begin{bmatrix} X(t) \\ Y(t) \end{bmatrix} + \begin{bmatrix} T_x \\ T_y \end{bmatrix} \quad (9)$$

令: $u = \frac{dX}{dt}$ $v = \frac{dY}{dt}$ 。由光流约束方程^[6], 设像素点 (X, Y) 经过 dt 的时间间隔运动到 $(X + dX, Y + dY)$, 则有:

$$I(X + dX, Y + dY, t + dt) = I(X, Y, t) \quad (10)$$

式中, $I(X, Y, t)$ 指 t 时刻像素点 (X, Y) 的亮度。

由 Taylor 公式展开 $I(X + dX, Y + dY, t + dt)$, 可得:

$$\begin{aligned} & I(X + dX, Y + dY, t + dt) \\ &= I(X, Y, t) + I_X * u + I_Y * v + I_t + \text{高阶项} \end{aligned} \quad (11)$$

$$\text{式中, } I_X = \frac{dI}{dX} \quad I_Y = \frac{dI}{dY} \quad I_t = \frac{dI}{dt}$$

显然, (u, v) 是点 (X, Y) 的运动速度。 I_X, I_Y, I_t 分别是对 I 的时空 (spatiotemporal) 偏导。如果忽略高阶项, 由式 (10) 和式 (11) 可得式 (12):

$$I_X * u + I_Y * v + I_t = 0 \quad (12)$$

令 $\beta = \alpha - 1, \gamma = \alpha \cdot W_z$, 把式 (9) 代入式 (12) 中可得:

$$\begin{aligned} & (I_X \cdot X + I_Y \cdot Y)\beta + (I_Y \cdot X - I_X \cdot Y)\gamma + \\ & I_X \cdot T_X + I_Y \cdot T_Y = -I_t \end{aligned} \quad (13)$$

考虑视频序列中相邻的两帧图象 I_1, I_2 , 设每帧图象有 K 个像素点, 由于摄像机运动时, 图象中所有点的位置都会发生变化。所以可以假定这 K 个像素点具有相同的运动参数 α, W_z, T_X, T_Y , 于是有:

$$F \cdot A = B \quad (14)$$

式中, F 是一个 $K \times 4$ 的矩阵, 每一行等于 $(I_X \cdot X + I_Y \cdot Y, I_Y \cdot X - I_X \cdot Y, I_X, I_Y)$, B 是一个 $K \times 1$ 的向量, 每个元素等于 $(-I_t)$, $A = (\beta, \gamma, T_X, T_Y)$ 。

在等式 (14) 中, 由于 I_X, I_Y, I_z 均可求, 所以 F 和 B 已知, 只有 A 未知。采用最小二乘法, 求得:

$$A = (F^T \cdot F)^{-1} \cdot F^T \cdot B \quad (15)$$

4 实现中的一些问题

以上讲了 PanoVideo 算法所依据的基本原理。由式 (15) 可以计算出在相邻两帧间摄像机的运动参数。但在实际应用时存在两种问题: 一是物体运动的干扰。在视频序列中除摄像机运动外, 物体可能也在运动。在估计摄像机运动时, 必须消除物体的运动, 否则参与物体运动的点将对公式 (15) 造成干扰。二是当摄像机运动的位移较大时, 式 (11) 中的高阶项的值也较大, 这时式 (12) 不再成立, 所以应用式 (15)

便有很大误差。PanoVideo 算法通过构造图象塔与选择迭代法来解决上述两问题。

4.1 构造图象塔

设源图象 I 包含 R 行 C 列个像素点。定义 I 的图象塔 $P = \{p_0, p_1, \dots, p_N\}$, P 按下列方式构造: 源图象 I 作为图象塔的第 0 级, 第 1 级图象 p_1 是从 p_0 中抽样形成的与 p_0 相似但尺寸小的图象, 第 2 级图象 p_2 又是从 p_1 抽样但比 p_1 小的图象, 如此类推, 一直到构造出 p_N 。显然, P 中 p_0 的运动参数即是源图象的运动参数。随着级数的增加, 图象被缩小, 图象的运动参数也变小。因而对图象塔中级数较高的图象来说, 因运动参数较小 (典型情况是远小于 1), 所以式 (2) 中 Taylor 展开式的高阶项便可以消去。PanoVideo 采用 Gaussian 图象塔构造法^[6]。 p_0 等于源图象 I , p_1 到 p_N 由式 (16) 定义:

在第 i 级图象 $p_i (1 \leq i \leq N)$ 中, 像素点 (X, Y) ($0 \leq X < C_i, 0 \leq Y < R_i$) 的亮度 $p_i(X, Y)$ 为:

$$p_i(X, Y) = \sum_{m=-2}^2 \sum_{n=-2}^2 e(m, n) p_{i-1}(2X + m, 2Y + n) \quad (16)$$

这里 C_i 和 R_i 为第 i 级图象的宽和高。 $e(m, n)$ 为权重因子, $e(m, n)$ 需满足下列条件:

$$\begin{aligned} & e(m, n) = \hat{e}(m) \cdot \hat{e}(n) \\ & \sum_{m=-2}^2 \hat{e}(m) = 1 \text{ 且 } \hat{e}(m) = \hat{e}(-m) (m = 0, 1, 2) \end{aligned}$$

为减少算法的计算量, 我们定义 $e(m, n) = \frac{1}{5}$, 即在相邻区域内取均值。

若已求出第 i 级摄像机运动参数 $(\alpha^i, W_z^i, T_X^i, T_Y^i)$, 则第 $i-1$ 级运动参数 $(\alpha^{i-1}, W_z^{i-1}, T_X^{i-1}, T_Y^{i-1})$ 为:

$$\begin{aligned} & T_X^{i-1} = 2 \cdot T_X^i, \quad T_Y^{i-1} = 2 \cdot T_Y^i, \\ & \alpha^{i-1} = \alpha^i, \quad W_z^{i-1} = W_z^i \end{aligned} \quad (17)$$

4.2 选择迭代法

为了消除物体运动的干扰。PanoVideo 算法是这样运转的: 首先计算第 N 级运动参数 $(\alpha^N, W_z^N, T_X^N, T_Y^N)$, 把得到的参数代入等式 (13) 中, 满足等式 (13) 的像素点组成集合 Ω_1 , 其它像素点组成集合 Ω_2 。集合 Ω_1 与 $(\alpha^N, W_z^N, T_X^N, T_Y^N)$ 被用作计算第 $N-1$ 级运动参数, 如此类推.....。

基于以上两点, PanoVideo 算法可归纳为 (如图 2):

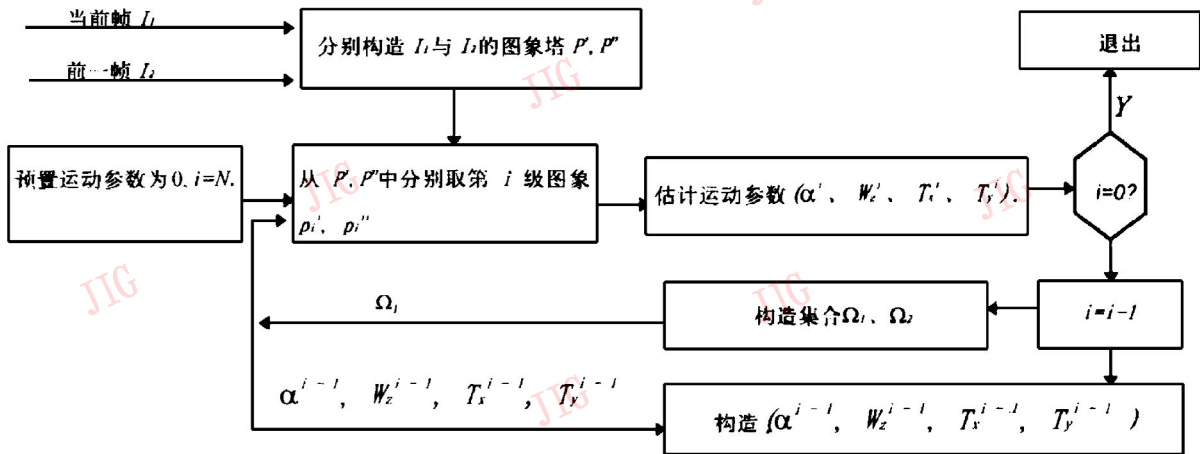


图 2 PanoVideo 算法的流程图

Fig. 2 The flow chart of panoVideo algorithm

(1) 首先构造 I_1 的图象塔 P' , I_2 的图象塔 P'' 。初始化: $i=N, \Omega_1^N = \{(X, Y) | 0 \leq X < C, 0 \leq Y < R\}$ 。并把 A 的初始值 $(\overline{\beta^N}, \overline{\gamma^N}, \overline{T_x^N}, \overline{T_y^N})$ 置为 0。把式(17)稍加修改: 设第 i 级运动参数 $A^i = (\beta^i, \gamma^i, T_x^i, T_y^i)$, A^i 的初始值 $\overline{A^i} = (\overline{\beta^i}, \overline{\gamma^i}, \overline{T_x^i}, \overline{T_y^i})$, 令 $C^i = A^i - \overline{A^i}$, 则:

$$C^i = (F^T F)^{-1} \cdot (F^T \cdot B - F^T F \overline{A^i}) \quad (18)$$

(2) 分别从 P', P'' 中取出第 i 级图象 P'_i, P''_i , 对 Ω_1^i 中的所有象素点应用式(18), 求出运动参数 C^i 。

$$(3) A^i = C^i + \overline{A^i}.$$

(4) 若 $i=0$, 退出 PanoVideo 算法, 摄像机运动参数 (α, W_z, T_x, T_y) 按下列方式计算:

$$\alpha = \beta^i + 1,$$

$$W_z = \frac{\gamma^i}{\alpha},$$

$$T_x = T_x^i,$$

$$T_y = T_y^i$$

若 i 不等于 0, 把 A^i 代入式(13)。对 Ω_1^i 中的每个象素点, 计算 I_x, I_y, I_z, X, Y , 满足(13)式的象素点被放入集合 Ω_1^{i-1} 。

计算 A^{i-1} 的初始值 $\overline{A^{i-1}}$: 由式(17)可知, $\overline{\beta^{i-1}} = \beta^i, \overline{\gamma^{i-1}} = \gamma^i, \overline{T_x^{i-1}} = 2 \cdot$

设当前正在计算两帧图象 I_1, I_2 间摄像机运动参数 (α, W_z, T_x, T_y) 。

$$T_x^i, \overline{T_y^{i-1}} = 2 \cdot T_y^i.$$

(5) i 减 1, 转到第(2)步执行。

4.3 测试与评价

以上提出的 PanoVideo 算法已在实际的视频中测试过。图 3 是我们用摄像机在实验室拍下的几幅测试图象。其中, (a) 和 (b) 是场景不动摄像机平移时拍摄的平移前后的两幅图象, 由 PanoVideo 算法求得平移参数为 $(T_x, T_y) = (6.0, -0.5)$ 的象素



图 3 PanoVideo 算法的实验结果

Fig. 3 The experimental result of PanoVideo algorithm

点。把(a)按(6.0, -0.5)进行运动补偿得到的图象与(b)基本吻合。

(c)和(d)是场景不动,摄像机焦距变化前后的图象,用 PanoVideo 算法估计结果为 1.033。在(e)和(f)图象中,除摄像机平移导致图象背景的移动外,图象中的人脸部分也同时运动,采用 PanoVideo 算法可估计出摄像机运动参数 $(\alpha, W_z, T_x, T_y) = (1.002, 0.027, 4.0, -1.6)$,若在应用 PanoVideo 算法时不采用选择迭代法,估计结果为 $(\alpha, W_z, T_x, T_y) = (1.028, 0.148, 6.3, 4.8)$ 。

由于在给摄像机运动建模时,我们把摄像机的旋转运动近似为平移,这种近似实际上就假定在摄像机旋转时景深(Z坐标)变化很小。如果摄像机离物体很近时做旋转运动,每个象素点将不仅是X与Y坐标且Z坐标的变化也表现的很明显,因而应用式(4)时的误差较大,从而导致应用整个算法有误差。因而本算法只适用于摄像机旋转时,视频中各象素点景深变化相对较小的场合。实验表明,该算法可成功地应用于典型的电视会议场景中视频的压缩与检索。

5 应用实例

图 4 是由摄像机运动参数从视频中构造出的一幅全景图。摄像机在我们实验室从左至右平移 5 秒钟得到一段测试视频。采用 PanoVideo 算法计算出在视频相邻两帧间摄像机的运动参数,并由这些参数把各帧图象拼接成一幅无缝的全景图。构造的全景图可应用于 3 个目的:



图 4 由摄像机运动参数造出的全景图

Fig. 4 The constructed panoramic image

(1)视频浏览:有了全景图,浏览整段视频不再需要逐帧地看,只需看一下全景图,便对视频的内容一目了然。

(2)基于内容检索:比如需要检索测试视频中有椅子的图象帧。我们只需从全景图中找到椅子,并由各帧在全景图中的位置即可确定各帧是否有椅子。

(3)视频压缩:由于整段视频可以由一幅全景图来代表,所以可以实现很高的压缩比。在上述的例子中,全景图的宽与高分别为 260 与 120 个象素点。每帧图象长与宽分别为 160 和 120 个象素点。5 秒钟的视频共有 150 帧,所以由全景图可实现的压缩比为 $\frac{150 \times 160 \times 120}{260 \times 120} \times C_{\text{JPEG}}$ 。 C_{JPEG} 是指采用 JPEG 对全景图压缩的压缩比。设 $C_{\text{JPEG}} = 30$,则总压缩比等于 2769 倍。

6 总结

PanoVideo 算法主要用于从视频序列中估计摄像机的三维运动参数,它是我们正在研究的基于结构描述的视频压缩算法的关键技术之一。PanoVideo 算法依据 3 个基本思想:(1)摄像机三维运动建模。(2)构造图象塔。(3)选择迭代。在某些应用场合,如可视电话会议系统中,对于典型的脸——肩视频序列,PanoVideo 算法在使用时与人脸的三维模型及三维运动模型^[7]相结合,可以获得更理想的效果。

参考文献

- 1 Otsuji K, Tonomura T. Projection Detecting Filter for Video Cut Detection. ACM Multimedia'93 Conference Proceedings, 1993: 251~257.
- 2 Laura Teodosia, Walter Bender. Salient Video Stills: Content and ConteXt Preserved, ACM Multimedia'93 Conference, 1993: 39~46.
- 3 Michal Irani. Video Compression using Mosaic Representations. Signal Proceedings: Image Communication 7, 1995: 529~552.
- 4 Netravali A N, Salz J. Algorithms for Estimation of Three Dimension Motion. AT&T Technical journal, 1985. 2, 64(2).
- 5 Horn B K P, Schunck B. Determining Optical Flow. Artificial Intelligence 17. 185~203.
- 6 Peter J Burt, et al. The laplacian pyramid as a compact image code. IEEE Transactions on communications. 1993. 4.
- 7 Chang Seok Choi. Analysis and Systhesis of Facia Image Sequences in Model-based Image Coding. IEEE Transactions in Circuits and Systems for Video Technology. 1994. 6.

附录 A

证明式(7)。式(7)可以转化为如下命题:物体 AB 向摄像机方向平移 T_z , 新位置上物体被标记为 CD, CD 的像为 C_1D_1 。在这过程中,摄像机的焦距保持不变。求证:若物体 AB 不移动,但改变摄像机的

焦距 $f_{AB} = \frac{d}{d - T_z} \cdot f_{CD}$ (d 为 AB 的物距), 那么, 此时 AB 的像 A_1B_1 等于 CD 的像 C_1D_1 , 即 $A_1B_1 = C_1D_1$ 。

证明:

由透镜成像公式可得:

$$\frac{1}{d} + \frac{1}{v_{AB}} = \frac{1}{f_{AB}} \tag{A1}$$

把 $f_{AB} = \frac{d}{d - T_z} \cdot f_{CD}$ 代入式(A1)可得:

$$\frac{1}{d} + \frac{1}{v_{AB}} = \frac{d - T_z}{d \cdot f_{CD}} = \frac{u_{CD}}{d \cdot f_{CD}} \tag{A2}$$

式(A2)两边同乘以 $\frac{d}{u_{CD}}$, 有:

$$\frac{1}{u_{CD}} + \frac{d}{v_{AB}} \cdot \frac{1}{u_{CD}} = \frac{1}{f_{CD}} \tag{A3}$$

把式(A3)减 $\frac{1}{u_{CD}} + \frac{1}{v_{CD}} = \frac{1}{f_{CD}}$, 得:

$$\frac{d}{v_{AB}} \cdot \frac{1}{u_{CD}} = \frac{1}{v_{CD}}$$

$$\text{所以, } \frac{d}{u_{CD}} = \frac{v_{AB}}{v_{CD}} \tag{A4}$$

图 A1 是依据透镜成像原理绘制的。显然, 三角形 OCD 相似于

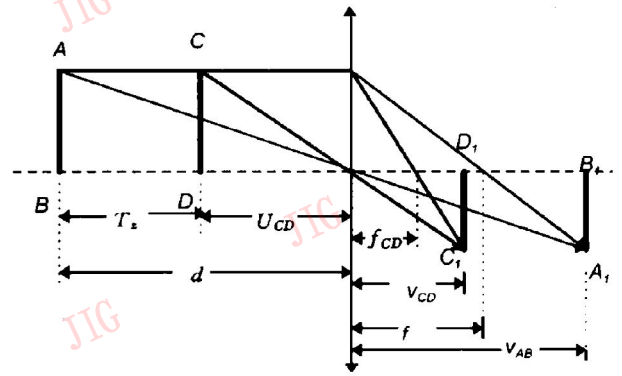


图 A1 摄像机成像原理图

Fig. A1 The Principle of camera Imaging

OC_1D_1 , 三角形 OAB 相似于 OA_1B_1 , 所以,

$$\frac{d}{AB} = \frac{v_{AB}}{A_1B_1}; \frac{u_{CD}}{CD} = \frac{v_{CD}}{C_1D_1} \tag{A5}$$

由已知条件可知: $AB = CD$, 于是:

$$\frac{A_1B_1}{C_1D_1} = \frac{v_{AB}}{v_{CD}} \cdot \frac{v_{CD}}{d} = 1$$

因此 $A_1B_1 = C_1D_1$, 得证。

张茂军, 国防科技大学系统工程系博士研究生。曾获部委级一等奖, 共发表各类论文 20 余篇。目前主要研究方向为多媒体会议系统、压缩视频流处理技术与虚拟现实系统。

PanoVideo: A Efficient Approach for Motion Modelling and Motion Estimation of the Camera

Zhang Maojun, Hu Xiaofeng, Ku Xishu

(Dept. 7, National University of Defense Technology, Changsha 410073)

Abstract This paper studies an algorithm for estimating camera zooms, rotations and pans from video sequences. Three new options are presented: a simple 3-D model of camera motion, image pyramid, and selective iteration. This algorithm can be used in video compression, panoramic imaging, content-based retrieval for video, and virtual reality.

Keywords Camera motion estimation, Panoramic imaging, Content-based retrieval, Video compression