

基于线条特征的双目立体视觉系统

艾海舟 许万雍 周远清

(清华大学计算机系)

摘 要

本文阐述了一个以线条特征为匹配基元的双目立体视觉系统。匹配算法以环境相似性度量为核心,以多层表示为基础,采用由粗而精的控制策略,通过预测、归并、分解、传播四过程实现。匹配算法考虑到实际环境中的双目立体图象的非理想性及抽取出来的线条特征的非完整性,允许某种程度的纵向偏差,引入中心主轴概念,在分解过程中建立对应,计算视差。首先由前三个过程建立起可靠的粗层次的对应,称之为骨架对应,然后利用连续性约束假设将骨架对应的结果由传播过程逐级传播到整幅图象中去。

关键词——机器人学,立体视觉,匹配。

一、引 言

双目立体视觉理论建立在对人类视觉系统研究的基础上,通过双目立体图象的处理,获取场景的三维信息。其结果表现为深度图,实现二维图象到三维空间的重构, Marr 等

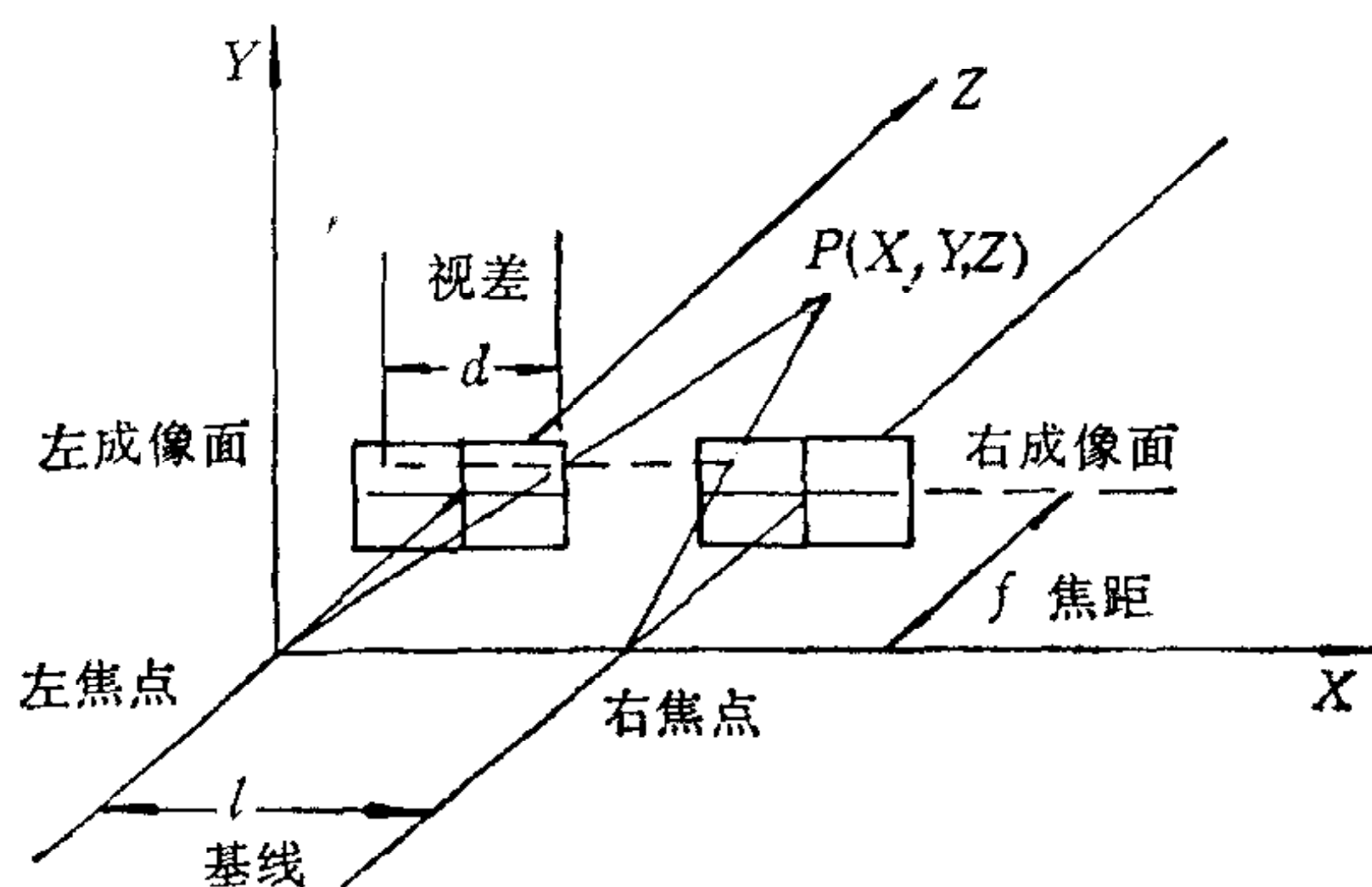


图1 双目理想透镜定位坐标模型

人^[1]最早提出并实现了一种基于人类视觉系统的计算视觉模型及算法。随着智能机器人技术的发展,这一课题受到极为普遍的重视,出现了各种形式的算法^[2-4]。解决对应点问题的匹配算法是双目立体视觉的关键。双目立体视觉模型中,双摄像机彼此参数一致,光轴平行且垂直于基线,构成一共极性结构,如图1所示。

二、线条基元的抽取

本算法的线条基元是通过对图象进行边缘检测、细化和符号化处理得到的,符号化是根据同一条线上边缘点的方向码几乎是相同的这一简单事实实现的^[5]。每个线条基元含有参数为两端点的坐标、方向角、长度及线上点的平均强度。

三、匹配算法

1. 预测过程

预测过程是找出每条线的最可能的对应线的过程,是依据如下条件进行的:(1)方向的一致性;(2)处在允许视差范围内;(3)周围环境的相似度。首先定义几个概念。

定义 1. “方向一致”是指两线条的方向角的差别在一定允许范围内。

定义 2. “以线条 p 为中心的窗口 (dx, dy) ”是指图 2 所示的范围。

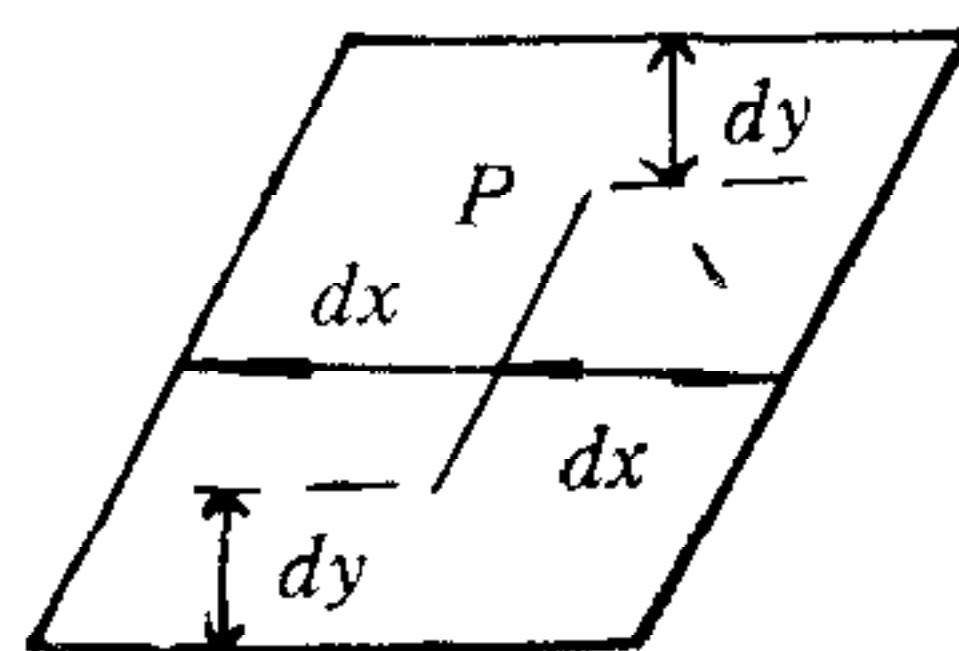


图 2 以 p 为中心的窗口 (dx, dy)

定义 3. “线条 p 与线条 q 相邻”的涵义是指 p 与 q 方向一致,且其中较短线的中点在以较长线为中心的窗口 (dx, dy) 中。

定义 4. “线条 p 在线条图 G 中的环境”是指线条集合

$$E(G, p) = \{e | e \in G, e \text{ 与 } p \text{ 方向一致且相邻}\}.$$

其中 $e \in E(G, p)$; $\text{length}(e)$ 为 e 线的长度; $d(p, e)$ 为线条 p 到线条 e 的平移距离; $V(p, e)$ 为线条 p 到线条 e 的平移交叠长度。

定义 5. “线条 p 在线条图 G 中的纵向分布度量”是指

$$D(G, p) = \sum_{e \in E(G, p)} d(p, e) \cdot \text{length}(e).$$

定义 6. “线条 p 在线条图 G 中的横向分布度量”是指

$$V(G, p) = \sum_{e \in E(G, p)} d(p, e) \cdot v(p, e).$$

定义 7. “线条 p 在线条图 G 中的环境度量”是指 $f(G, p) = W_1 \cdot D(G, p) + W_2 \cdot V(G, p)$ 。其中 W_1, W_2 为加权系数。

定义 8. “线条图 G_1 的线条 p 与线条图 G_2 的线条 q 的环境相似性度量”是指

$$g(G_1, G_2, p, q) = |f(G_1, p) - f(G_2, q)|.$$

预测过程描述如下: 设双目立体图象的符号化描述是 G_1 与 G_2 , $G_1 = \{e_1, e_2, \dots, e_m\}$, $G_2 = \{e'_1, e'_2, \dots, e'_n\}$ 。先计算出基于总体分布中心的总体视差 $C1(X_{c1}, Y_{c1})$,

$C2(X_{c2}, Y_{c2})$, $\Delta x = X_{c2} - X_{c1}$. 用 Δx 作为视差的粗估计, 作为确定对应候选集的开窗依据. 考虑到实际中的双目立体图象并非只有水平视差, 在设计算法中, 允许纵向有一个小的波动范围. 在总体视差基础上, 根据处理的环境特点, 确定出允许最大水平视差 ($\pm \max dx$) 及允许纵向波动范围 ($\pm \max dy$).

设 G_1 的任意一条线为 e , 将 e 平移 Δx 得 e' , 在 G_2 中以 e' 为中心的窗口 ($\max dx$, $\max dy$) 范围里确定 e 的对应候选集 $F \subseteq G_2$, $F(G_1, e) = \{e'' | e \in G_1, e'' \in G_2, \text{将 } e \text{ 平移 } x \text{ 得 } e', e'' \text{ 与 } e' \text{ 方向一致且相邻}\}$. 根据周围环境的相似度, 确定 F 中最合适的对应. 取加权系数 $W1 = V(G_1, e)$, $W2 = D(G_1, e)$,

$$e'_k = \min_{e'' \in F(G_1, e)} g(G_1, G_2, e, e'') = \min_{e'' \in F(G_1, e)} |f(G_1, e) - f(G_2, e'')|,$$

则称 $e'_k \in G_2$ 是 $e \in G_1$ 的预测对应线, 记作 $e \in G_1 \rightarrow e'_k \in G_2$.

2. 归并过程

归并过程是将双向预测的结果归并成对应的组, 每组由两个对应的线集合表示, 可以认为其中任意一集合中的任一线的对对应线存在于另一线集合中. 归并的结果相当于两幅图象在线条意义下的一次对应分割, 这一过程描述如下: 设预测结果是

正向预测: $\forall e \in G_1 \rightarrow p_1(e) \in G_2$;

反向预测: $\forall e \in G_2 \rightarrow p_2(e) \in G_1$.

首先形成初始归并 S_1 与 S_2 .

$$\forall e \in G_1, S_1(e) = \{e' | e' \in G_2, p_2(e') = e\},$$

$$\forall e \in G_2, S_2(e) = \{e' | e' \in G_1, p_1(e') = e\}.$$

然后将 S_1 和 S_2 归并起来形成若干互不相关的对应集合组, 每个对应集合组由两个互相对应的集合组成, 不妨设有 n' 个这样的组, 记作

$$T(i) = \{M_1(i), M_2(i)\}, \quad i = 1, 2, \dots, n'.$$

归并算法如下:

1) 扫描预测结果, 形成初始归并

$$S_1(e_1), S_1(e_2), \dots, S_1(e_m),$$

$$S_2(e'_1), S_2(e'_2), \dots, S_2(e'_n);$$

2) $i := 1, M_1 := \{e_1\}, M_2 := S_1(e_1)$, 给 e_1 打标;

3) 在 M_2 中找未打标的元素 e' , 若存在, 则做 $M_1 := M_1 \cup S_2(e')$, 给 e' 打标, 返回 3); 若不存在, 执行 4);

4) 在 M_1 中找未打标的元素 e , 若存在, 则做 $M_2 := M_2 \cup S_1(e)$, 给 e 打标, 返回 4); 若不存在, 执行 5);

5) 若 M_2 中存在未打标的元素, 则返回 3); 否则, 形成一对对应集合组 $T(i) := \{M_1, M_2\}$, 再执行 6);

6) 在 G_1 中找未打标的元素 e , 若存在, 则做给 e 打标, $i := i + 1, M_1 := \{e\}, M_2 := S_1(e)$, 返回 3); 若不存在, 则停机.

3. 分解过程

首先来考虑计算视差的问题, 假若对应 $e \in M_1 \longleftrightarrow e'_k \in M_2$, 那么如何计算视差(d)?

对于不接近于水平的线 e 和 e' ，利用对应点在同一水平扫描线上的性质，视差就是从 e 沿水平扫描线到 e' 的距离。但对于接近于水平的线就不能这样做了。因为在这种情况下不能再用水平扫描线确定对应点计算视差。但根据图象特征分布的相似性特点，可以用如下的估计方法：

设已知 m 组对应线 ($e_i \leftrightarrow e'_i, i = 1, 2, \dots, m$) 及其平均视差 ($d_i, i = 1, 2, \dots, m$)。 $p_i(x_i, y_i)$ 与 $p'_i(x'_i, y_i)$ 是 e_i 和 e'_i 上的一组对应点。现在寻找对应的两条近水平方向线中较短线中点 (x_{mid}) 的对应点。设 x 为较长线段上的任意一点 $\forall x \in [x_f, x_t]$ ，计算

$$l_i = x_{mid} - x_i, l'_i = x - x'_i, f(x) = \sum_{i=1,2,\dots,m} |l_i - l'_i| = \sum_{i=1,2,\dots,m} |x_{mid} + d_i - x|.$$

设 $f(x_0) = \min\{f(x) | x \in [x_f, x_t]\}$ ，把 x_0 作为所找的对对应点，视差 $d = x_0 - x_{mid}$ 。

对应集合组的某一集合的中心主轴是指通过其分布中心，沿着该集合中的线的方向，不超出该集合范围的控制线。下面描述分解过程。

第一步将所有对应集合组重新排序，将非近水平方向的线放在前面首先处理。处理完所有非近水平方向的线后，再处理近水平方向的线。

对于每一对应集合组 $T = \{M_1, M_2\}$ ，一般可以认为两中心主轴是对应的，故可利用每条线与各自中心主轴的平移距离 (d) 将该对应集合组分解，采用三分法分解，如图 3 所示。

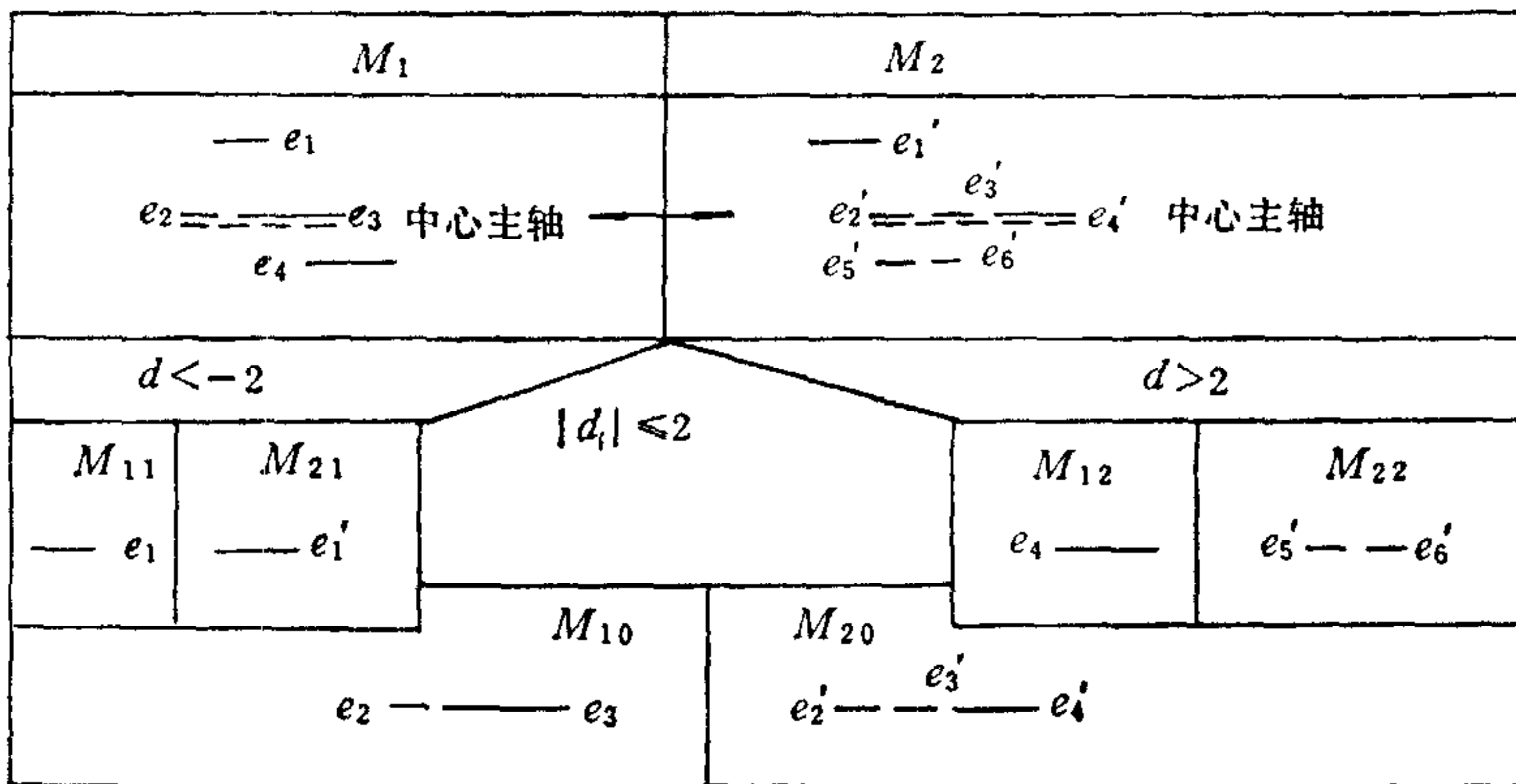


图 3 分解过程图示

分解得 $(M_{11}, M_{21}), (M_{10}, M_{20}), (M_{12}, M_{22})$ 三部分。根据分解的不同类型，分别采取如下处理方法：

1) $(M_{10} \neq \emptyset, M_{20} \neq \emptyset)$ 。 (M_{10}, M_{20}) 一般代表场景中一条线在两幅图象中都断成了若干段的情况。这时认为它们的中心主轴是对应的，计算基于中心主轴的对应视差，赋予集合 M_{10} 中的每条线。将 $(M_{11}, M_{21}), (M_{12}, M_{22})$ 中不存在非空集合的对应组加入待继续分解的对应集合组之列，以便进一步分解对应。

2) $(M_{20} = \emptyset, M_{11} = \emptyset, M_{12} = \emptyset)$ 。 M_1 集中于中心主轴，代表一条线，而 M_2 的中心主轴并不反映对应情况，是一种例外。这时在 M_2 中选择 M_1 的中心主轴在环境相似性意义下的最佳对应线，计算视差，赋给 M_1 的每条线。

3) ($M_{10} = \emptyset, M_{21} = \emptyset, M_{22} = \emptyset$). M_2 集中于中心主轴, 代表一条线, 而 M_1 的中心主轴并不反映对应情况. 类似于 2) 的情况, 在 M_1 中找 M_2 的中心主轴的对应线, 计算视差, 赋给该对应线.

4. 传播过程

这一过程的目的在于利用上述过程的结果引导较全局的匹配. 可以认为, 对于较长线条的符号化描述层次, 上述三过程后的结果很可靠, 但对于短线条的描述层次, 由于所

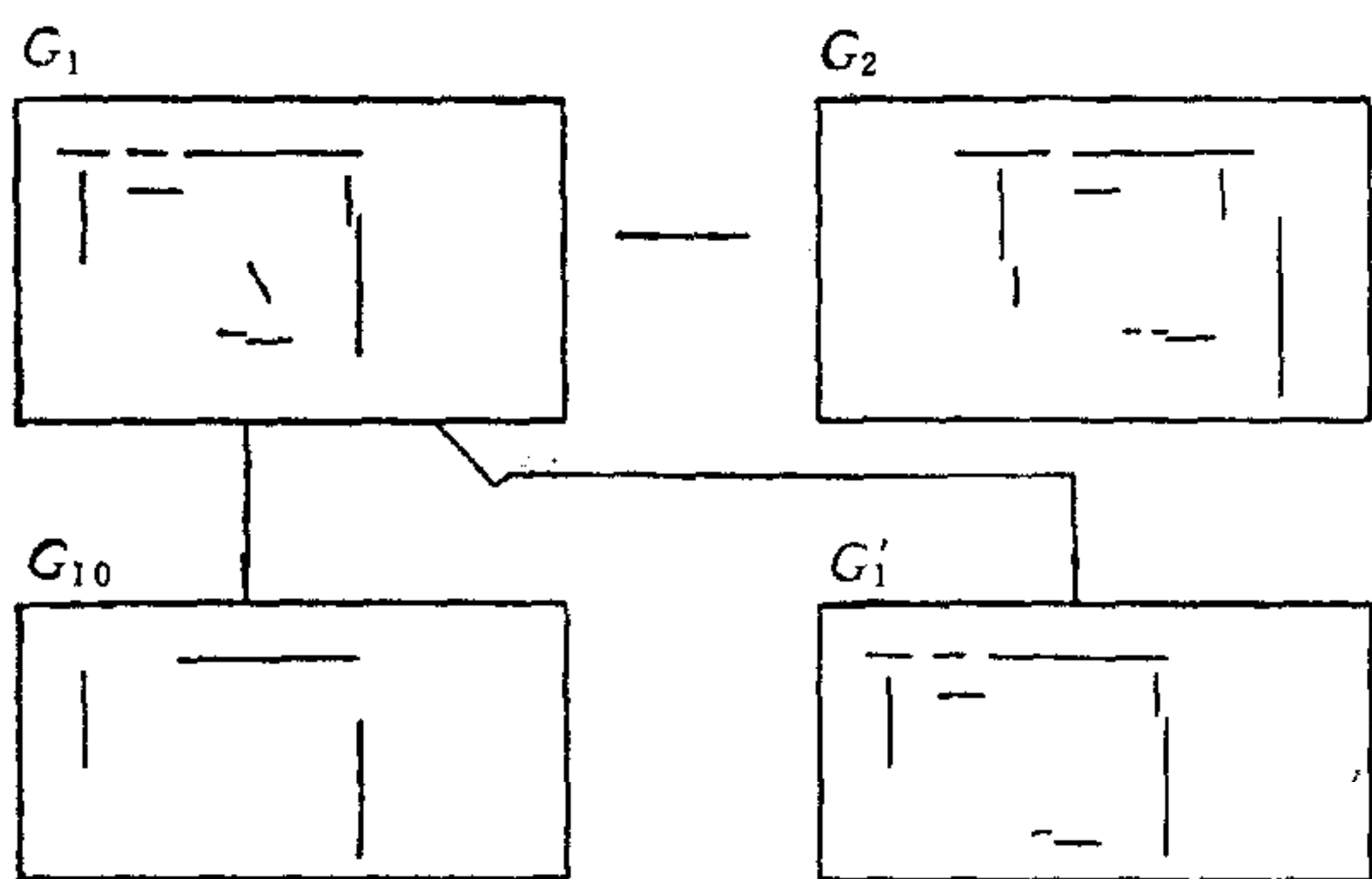


图 4 传播过程图示

定义的环境相似性度量的局限性, 上述结果难以适用. 根据这一性质, 可以建立图象的基于最小线条长度的不同层次的描述. 首先用前三个过程获取主导特征的对应关系, 建立起较粗层次描述的对应, 再利用该结果引导较精层次描述的对应. 传播过程是基于这种思想建立起来的, 类似于由粗而精方式的匹配过程, 但又具有自己的独特性, 它是将骨架的可靠对应结果传播到各个分支以至整体中去.

传播过程的具体实现描述如下:

设较长线条层次的对应结果已知, 体现在 G_{10} 图上. 现在寻求较短线条层次描述 G_1 与 G_2 的对应, 如图 4 所示.

对于 G_1 中的任意一线 ($\forall e \in G_1$), 在 G_{10} 中找与其具有相近方向而且平移距离 (d) 最小的线 ($e' \in G_{10}$). 如果该平移距离在允许误差范围内 (如 $|d| \leq 1$), 则直接接收该线 (e') 的视差数据作为线 (e) 的视差; 否则, 以该线 (e') 的视差作为在 G_2 中开窗的依据, 在所确定的小窗中用预测过程中求最佳可能对应的同样方法寻求对应线, 计算视差, 赋给线 (e). 与预测过程中的窗相比, 这时可以选择更小的窗, 如减小一半, 从而在很大程度上缩小搜索空间, 减少误匹配率.

四、距离计算

在图 1 所示的 (X, Y, Z) 坐标系中, 三维信息的计算简单直接.

$$(X, Y, Z) = L/d \cdot (x, y, f).$$

其中基线长 $L = 14.60$ 厘米, $f = 525.514$ 象束. 将匹配算法得到的每条线两端点的视差数据, 经上述计算过程就形成了空间中的线, 完成了由图象空间到三维空间的映射.

五、示例与误差分析

作者在实验中处理了多组图象, 其结果是令人满意的. 限于篇幅, 这里仅给出一组图象的实验结果及部分图片 (图 5).

距离计算的结果简单说明如下: 所处理的图象是实验室中拍摄的终端桌图象, 终端

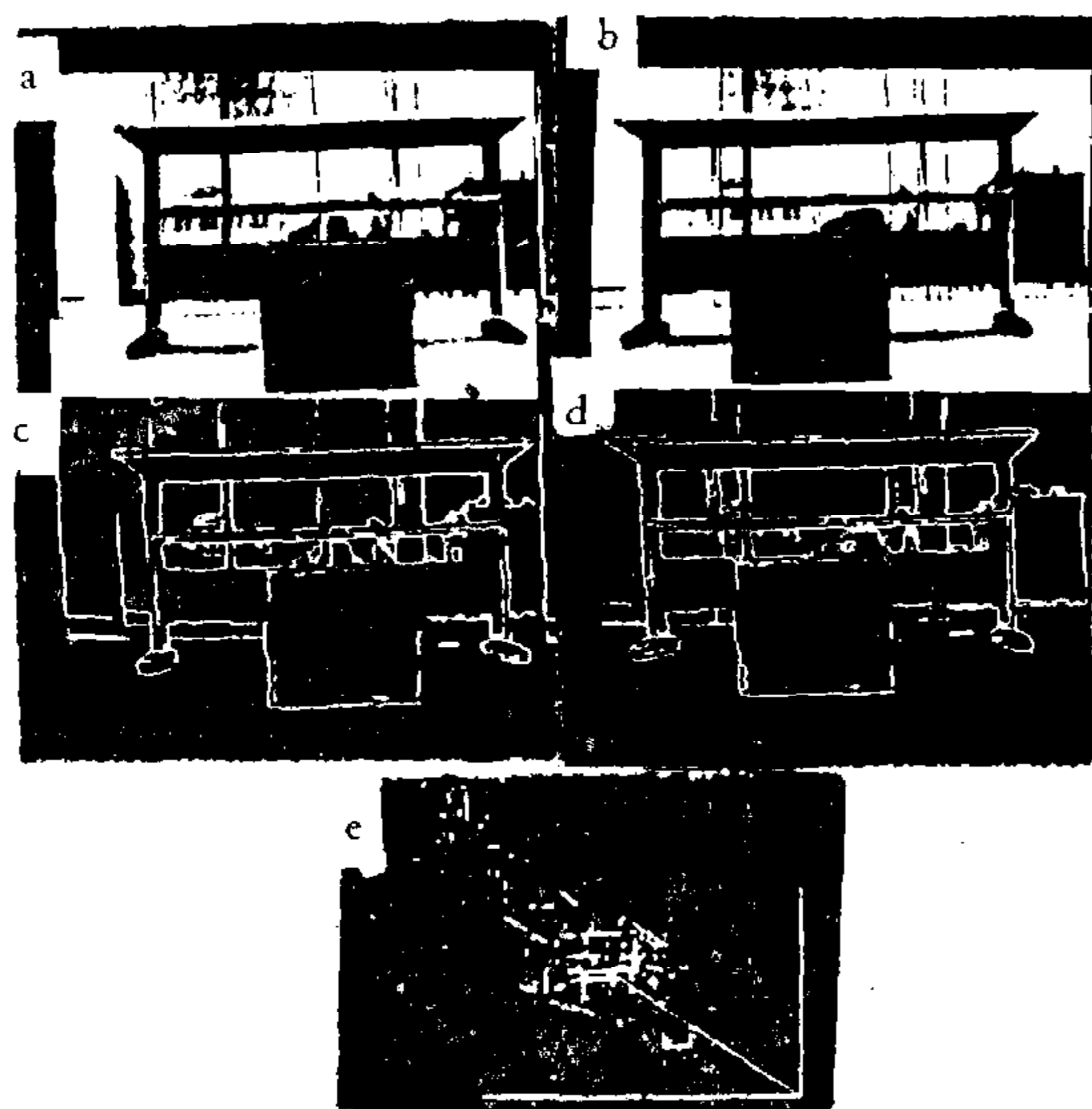


图5 a, b: 立体图象对; c, d: 最高层线条图对; e: 恢复三维信息(透视变换)

桌 (3.00—3.50m) 前放了一个盒子 (2.00m), 背景是玻璃隔层 (6.00m); 玻璃隔层外的书桌 (6.00—7.40m), 墙面 (10.00m) 前的柜子 (9.40—10.00m)。计算出的距离集中在上面列出的数值上, 说明视差基本上是可靠的。对于主导特征而言, 距离计算很准确, 几乎与测量结果完全一致, 误差在 5% 以内。但事实上, 将距离计算的结果通过透视变换恢复出来的场景与原来场景相差还是比较远的, 这主要是由近水平方向对应线的视差的非精确性造成的。

六、小 结

这项研究的目的是为自主车配视觉, 环境是简单的室内场所。室内环境是以线特征为主导特征的, 而且其中水平线和垂直线居大多数。该算法比较适合于自主车的较长距离的制导过程。

匹配算法的关键问题出现在计算近水平方向对应线的视差上。可以设想再增加一个摄像机, 使得它与原来两摄像机中的一个构成另一双摄像机组。与原来那组不同, 这组所拍摄的图象之间没有水平视差, 而只有纵向视差。将这样的两组图象 (共三幅不同的图象) 的对应计算视差的结果结合起来, 就可以取得比较理想的结果。

参 考 文 献

- [1] Eric, W. and Grimson, L., Computational Experiments with a Feature Based Stereo Algorithm, *IEEE Trans., PAMI-7*(1985), 17—34.
- [2] Medioni, G. and Nevatia, R., Segment-Based Stereo Matching, *Computer Vision, Graphics and Image Processing*, **31**(1985), 2—18.
- [3] Medioni, G. and Nevatia, R., Matching Images Using Linear Features, *IEEE Trans., PAMI-6*(1984), 675—685.
- [4] Price, K. and Reddy, R., Matching Segments of Image, *IEEE Trans., PAMI-1*(1979), 110—116.
- [5] 艾海舟等, 一种简单的线条特征抽取算法, *机器人*, **3**(4)1989, 32—35。

LINE-SEGMENT-BASED BINOCULAR STEREO VISION SYSTEM

AI HAIZHOU XU WANYONG ZHOU YUANQING
(*Tsinghua University*)

ABSTRACT

An line-segment-based binocular stereo vision system is illustrated in this paper. The matching algorithm is based on the region similarity criterion by a coarse-to-fine strategy in multilevel representation. The algorithm is composed of four steps: predicting, appending, splitting, and spreading. Coping with the uncertainty of real stereo images and the incompleteness of extracted features, the algorithm accepts vertical diversions to some extent. A concept of the principal axis is introduced and used for matching which takes place in the step of splitting. A reliable skeleton matching can be built at the end of the first three steps then it is spread over the whole image sequentially.

Key words —Robotics; stereo vision; matching.