

一个两级迭代体视匹配算法¹⁾

李文彪 潘士先

(北京航空航天大学电子工程系 100083)

摘 要

本文给出了一个两级迭代体视匹配算法。算法以单个 $\nabla^2 G$ 通道输出中那些可因相似的方向连接成段的零交为匹配基元。第一级受体视 Marr-Poggio 内聚性的视差平滑性规则的控制,第二级由第一级的输出驱动并受体视 Mayhew-Frisby 形状连续性的视差连续性规则控制,从而使匹配结果具有在上述两个体视一般物理约束意义上的局部和全局一致性。文中给出了实验结果。

关键词: 体视匹配,视差平滑性,视差连续性。

一、前 言

体视匹配是视差体视理论的核心问题。任何体视匹配算法必须解决如下三个问题:

P1. 表像选择: 匹配赖以进行的图象特征的选择。

P2. 匹配规则: 依所选表像将物理世界的某些固有约束表示成匹配必须服从的若干规则。

P3. 算法结构: 利用适当的数学方法设计出能实现匹配规则又能在计算机上快速稳定地执行的算法。

Marr-Poggio 的“人的体视计算理论^[4]”在视觉和机器视觉研究中具有重要和广泛的影响。但作为这一理论重要组成部分的多通道协同处理迄今未获得完整而有效的结果。Mayhew 和 Frisby 的体视理论^[5]在体视表像和物理约束方面对 Marr-Poggio 理论提出了补充和商榷,并有相当重要的影响。随着机器视觉的发展,现已有了许多面向机器视觉的体视匹配算法。这些算法在表像选择和匹配约束方面以不同程度和方式借鉴了 Marr-Poggio 理论中的 $\nabla^2 G$ 零交、唯一性和内聚性以及 Mayhew-Frisby 的形状连续性 (figural continuity) 等概念,而在算法上却十分多样化。典型的如 Baker 的 Vertabi 算法^[2], Ohta 和 Kanade 的动态规划法^[7], Barnard 和 Thompson 的松弛标记法^[3]等。这些算法各具特色,在解除模糊和抗干扰能力、复杂程度和计算效率等方面各有千秋,但它们的集合则可以说是为设计更为完善的体视匹配算法提供了重要的借鉴。

本文报道一种计算机体视匹配算法。这种算法以单个 $\nabla^2 G$ 通道的具有某种可连接

本文于 1991 年 1 月 7 日收到。

1) 国家自然科学基金资助项目。

性的零交作为体视表像,用体现 Marr-Poggio 内聚性的视差变化平滑性或最小视差原则以及体现形状连续性的沿连续零交段的视差连续性构成强有力的匹配规则。算法结构则综合利用了 Medioni 和 Nevatia 的最小视差迭代^[6]以及一种由姜勤等改造过用于实现形状连续性匹配的 Barnard-Thompson 松弛标记法^[1],构成一种两级迭代算法。算法稳定可靠,收敛快,且适于并行处理。

二、算法的设计策略

下面就体视匹配的三个基本问题(P1,P2,P3)说明本算法的设计策略。

表像选择。在 Marr-Poggio 理论中,体视表像是若干个 $\nabla^2 G$ 通道的零交。然而如在前面提到的多个通道的协同处理仍是一个难题。近来的体视匹配算法几乎都采用“边缘”或其某种组合作为表像。所谓边缘基本上是图象二阶差分的零交即灰度突变点。笔者倾向于采用单个 $\nabla^2 G$ 通道——相当于 Marr-Poggio 理论中最高分辨率通道——的零交作为表像。但是,对于任何一幅哪怕是内容极为简单的图象,高分辨率通道输出大量零交,其中大部分并不对应于物理边界。如果认为物理边界一般说来具有一定的空间延展性和连续性,那么只有那些属于连续零交段的零交才可能对应于物理边界。基于此种考虑,希望采用在某种可信度上对应于真实物理边界的零交作为表像。为此,可对离散图象定义连续零交段如下:

连续零交段:相邻行上列坐标相差 $\pm J$,方向差小于一给定阈值 δ 且长度超过一给定阈值 L 的同号零交的集合。

本文用所有属于零交段的零交作为体视表像(基元)。注意表像仍是单个零交而非零交段,即匹配是在单个零交而不是在零交段之间进行的。这种差别是重要的。例如,有的算法用直线化的零交段作为匹配基元^[6],因此匹配结果自然具有视差连续性,即沿任一直线段的视差是连续变化的。这种做法实际上是用二维连续性去强迫三维连续性。但二维连续性并不一定意味着三维连续性,人们需要做的恰恰是通过匹配获取三维连续的物理边缘。

匹配规则。根据迄今的研究,唯一性(物体表面上给定的一点在任一给定的时刻在空间有唯一的位置)、内聚性(一个场景由若干个物体所组成,其中每一物体的表面相对于它和观察者之间的距离而言是平滑的)和形状连续性(物体的边缘是连续的,可以理解为内聚性用于物体边缘的结果)也许可以说是最一般因而也是最少的体视物理约束(任何其它的约束或者是主观的或者需要附加某种先验知识)。据此,可提出如下四条匹配规则·

- R1. 相容性: 同号且方向和斜率在某种程度上相似的零交方能匹配。
 - R2. 唯一性: 任一零交有唯一对应点¹⁾。
 - R3. 视差平滑性: 视差沿内极线(epipolar line)方向的变化应最为平缓。
 - R4. 视差连续性: 零交段如果确是三维边缘的投影,其上视差应是连续的。
- 严格执行这些规则将保证匹配在上述物理约束意义上的局部和全局一致性。

1) 不考虑双目感知中的一些特殊情形。

算法结构。实现上述匹配规则的算法有许多选择。基于有效性和计算量兼顾的考虑,本文构造了如下的一个两级迭代匹配算法。

对于满足 R1 和 R2 的所有可能匹配,进行第一级迭代:

S1. 视差差评价。借用 Medioni 和 Nevatia 的最小视差差迭代算法^[6]。视差差实际上是视差平滑性的一种量度。原算法用于直线化零交段的匹配,现在用于零交的匹配,所以有所简化和修改。这一级迭代的结果是所有可能匹配以视差差函数为量度的平滑性评价,以及在这一量度上的最佳匹配——优先匹配。

选择 Medioni-Nevatia 算法的一个重要原因是它实际上只需迭代两次,并且所有各匹配对的视差差函数值可以并行计算。

优先匹配只有同时满足 R4,即它所在零交段上的视差满足某种连续性量度时,才被接纳为正确的匹配。因此有第二级迭代:

S2. 视差连续性评价。利用姜勤等的一种基于形状连续性的匹配算法^[1]。此算法系由 Barnard 和 Thompson 的一种概率模型松弛标记视差分析方法^[3]改造而来。它结构简单,稳定,初始收敛速率高(头几次迭代便可达到相当的收敛水平),且可并行处理。初始概率估计由 S1 输出的视差差评价价值计算。概率更新则完全由一个视差连续性评价函数控制,使视差连续性评价价值高的匹配的概率趋于增大。这一级迭代输出在视差连续性评价上是可信的匹配。

最后,取 S1 输出的优先匹配和 S2 输出的视差连续匹配的交集为匹配结果。

三、算 法

首先定义可能匹配。对左图坐标为 (i, j) 的零交 Z_l , 右图对应行上以 (i, j) 为中心, 宽为 $2d_{\max} - d_{\max}$ 为最大视差搜索范围——窗口内所有与 Z_l 同号且方向差不超过一阈值 θ 的零交都是 Z_l 的可能匹配。同样地可定义右图零交 Z_r 的可能匹配。当 d_{\max} 比 $\nabla^2 G$ 滤波器口径大得多时,方向相似程度并不是判定可能匹配似然率的可靠依据,因此方向相似性约束宜松。

以下分别描述两级迭代。

S1. 视差差评价

设左图零交 Z_l 有 M_l 个可能匹配,记作 $A_l(k)$, $k = 1, \dots, M_l$ 。相应的视差记为 $d_l(k)$, $k = 1, \dots, M_l$ 。令 $A_l(M_l + 1)$ 为 Z_l 的无匹配标记,则 Z_l 具有匹配标记集 $\{A_l(k): k = 1, \dots, M_l, M_l + 1\}$ 。

对于左图零交 $Z_l(i, j)$, 定义右图上的一个区域 $W_R(Z_l)$ 为其视差差评价域, $W_R(Z_l)$ 是以 (i, j) 为中心的一个 $2d_{\max} \times h$ 的矩形窗。 $W_R(Z_l)$ 内的零交数目记为 $N_R(Z_l)$ 。

同理,对右图零交 Z_r 可定义 $\{A_r(h): h = 1, \dots, M_r + 1\}$, $\{d_r(h): h = 1, \dots, M_r\}$ 及 $W_L(Z_r)$ 和 $N_L(Z_r)$ 。

设 Z_l 和 Z_r 是一对可能匹配,且 $Z_r = A_l(k)$, $Z_l = A_r(h)$ 。则定义此匹配对的视差差评价价值为

$$V_i^{(t+1)}(k) = \left\{ \sum_{Z_{i'} \in W_L(Z_i), A_{i'}(k') \in C_1(Z_{i'})} \min |d_i(k) - d_{i'}(k')| \right\} / N_L(Z_i) \\ + \left\{ \sum_{Z_{r'} \in W_R(Z_i), A_{r'}(h') \in C_2(Z_{r'})} \min |d_r(h) - d_{r'}(h')| \right\} / N_R(Z_i), \quad (1)$$

式中 t 为迭代次数, $C_1(Z_{i'})$ 和 $C_2(Z_{r'})$ 说明于后. 对左图任意零交 Z_i , 可得其所有可能匹配的视差差评价值 $V_i^{(t+1)}(k)$, $k = 1, \dots, M_i$. 同样地, 对右图零交 Z_r , 可得其所有可能匹配的视差差评价值 $V_r^{(t+1)}(h)$, $h = 1, \dots, M_r$. 由此可定义优先匹配.

优先匹配: 设 Z_i 与 Z_r 是一可能匹配对, $Z_i = A_r(h)$, $Z_r = A_i(k)$. 若对任意的 $Z_{i'} = A_i(k')$, $k' \neq k$ 有 $V_i^{(t)}(k) < V_i^{(t)}(k')$ 且对任意的 $Z_{r'} = A_r(h')$, $h' \neq h$ 有 $V_r^{(t)}(h) < V_r^{(t)}(h')$, 则称 Z_i 与 Z_r 是 t 次迭代时的优先匹配. 记为 $Z_r = Q^{(t)}(Z_i)$, $Z_i = Q^{(t)}(Z_r)$. 注意优先匹配是一一对应的.

$C_1(Z_{i'})$ 定义如下: $Z_{i'}$ 与 Z_i 不同行时, 若 $Q^{(t)}(Z_{i'})$ 非空, 则 $C_1(Z_{i'}) = Q^{(t)}(Z_{i'})$, 否则 $C_1(Z_{i'}) = \{A_{i'}(k): k = 1, \dots, M_{i'}\}$; $Z_{i'}$ 与 Z_i 同行时, 若 $Q^{(t)}(Z_{i'})$ 非空且不等于 Z_r , 则 $C_1(Z_{i'}) = Q^{(t)}(Z_{i'})$, 否则, 当 $\{A_{i'}(k)\} \neq Z_r$ 时, $C_1(Z_{i'}) = \{A_{i'}(k)\} - Z_r$; 当 $\{A_{i'}(k)\} = Z_r$ 时, 只得令 $C_1(Z_{i'}) = Z_r$, 以保证迭代进行.

同样可定义 $C_2(Z_{r'})$.

迭代从 $t = 0$ 开始. 对任意零交 x , $Q^{(0)}(x)$ 为空. 迭代至优先匹配不再改变为止(实际上只需两次)并且输出: 1) 最小视差差匹配集合 A_i ; 2) 每一可能匹配对的视差差评价值(用视差差评价值的最大值归一化).

S2. 视差连续性评价

首先计算初始概率估计. 令 Z_i 和 Z_r 是可能匹配对, $Z_r = A_i(k)$, $Z_i = A_r(h)$. 定义 Z_i 和 Z_r 的相似系数为

$$w_i(k) = \frac{1}{1 + CV_i(k)} = w_r(h), \quad (2)$$

式中 C 为一正常数.

令 Z_i 无匹配的概率估计为

$$P_i(M_i + 1) = 1 - \max_k [w_i(k)], \quad (3)$$

则 Z_i 所有可能匹配的初始概率可取为

$$P_i(k) = [1 - P(M_i + 1)] \frac{w_i(k)}{\sum_{i=1}^{M_i} w_i(i)}, \quad k = 1, \dots, M_i. \quad (4)$$

对 Z_r 同样可得匹配初始概率 $P_r(h)$, $h = 1, \dots, M_r$.

接着进行视差连续性评价. 设 $Z_r = A_i(k)$ 和 $Z_i = A_r(h)$. 对 Z_i 取一 $\sigma \times \sigma$ 矩形窗, 它包含 Z_i 且使窗口内过 Z_i 的零交段 ZS_i 有最大长度. 由 ZS_i 对该匹配对作出的视差连续性评价为

$$q_i^{(t)}(k) = \sum_{Z_{i'} \in ZS_i} \left[\sum_{m \in |d_i(k) - d_{i'}(m)| \leq 1} P_{i'}^{(t)}(m) \right], \quad (5)$$

t 表示迭代次数。同样可对 Z_r 定义 ZS_r , 并计算 $q_{2r}^{(t)}(h)$ 。该匹配对总的视差连续性评价值为

$$q_i^{(t)}(k) = q_{1i}^{(t)}(k) + q_{2r}^{(t)}(h) = q_r^{(t)}(h). \quad (6)$$

然后可以进行概率更新如下:

$$S_i^{(t+1)}(k) = P_i^{(t)}(k)[A + Bq_i^{(t)}(k)], \quad k = 1, \dots, M_i, \quad (7)$$

$$S_i^{(t+1)}(M_i + 1) = P_i^{(t)}(M_i + 1), \quad (8)$$

$$P_i^{(t+1)}(k) = \frac{S_i^{(t+1)}(k)}{\sum_{j=1}^{M_i+1} S_i^{(t+1)}(j)}, \quad k = 1, \dots, M_i + 1. \quad (9)$$

A, B 为更新速率控制参数。同样地更新 $P_r(h)$ 。若干次迭代后, 若匹配 $\{Z_l = A_r(h), Z_r = A_l(k)\}$ 满足 $P_l(k) \geq T$ 和 $P_r(h) \geq T$, (T 为一适当的阈值, $0 < T < 1$), 则接纳 (Z_l, Z_r) 为视差连续匹配集 A_2 的元素。

最后输出匹配集 $A = A_1 \cap A_2$ 。

四、实 验

实验所用的立体图对如图 1 所示, 数字图象为 $140 \times 140 \times 8\text{bit}$ 。(c), (d) 是连续零交图, 这里取 $J=2$, $\delta = \pi/2$, $L=4$ (因此这里所谓连续, 其条件是十分宽松的)。图 (e) 是匹配结果, 即左图中每个获得匹配的零交偏移相应视差值后所得的二值图。所用参数值为: $d_{\max} = 15$, $\theta = \pi/4$, $h = 5$, $c = 10$, $A = 0.3$, $B = 3$, $\sigma = 10$, $T = 0.7$ (注意, 视差搜索范围很大: ± 15 , 因此假目标相当多)。视差差评价迭代 2 次, 视差连续性评价迭代 4 次。在进行视差连续性评价时, 零交的连续条件较获取表像时为严: $J=1$, $\delta = \pi/4$ 。

从匹配结果看, 由噪声引起的零交大都因不能连结成段或未获匹配而被去除, 对应于真实物理边缘

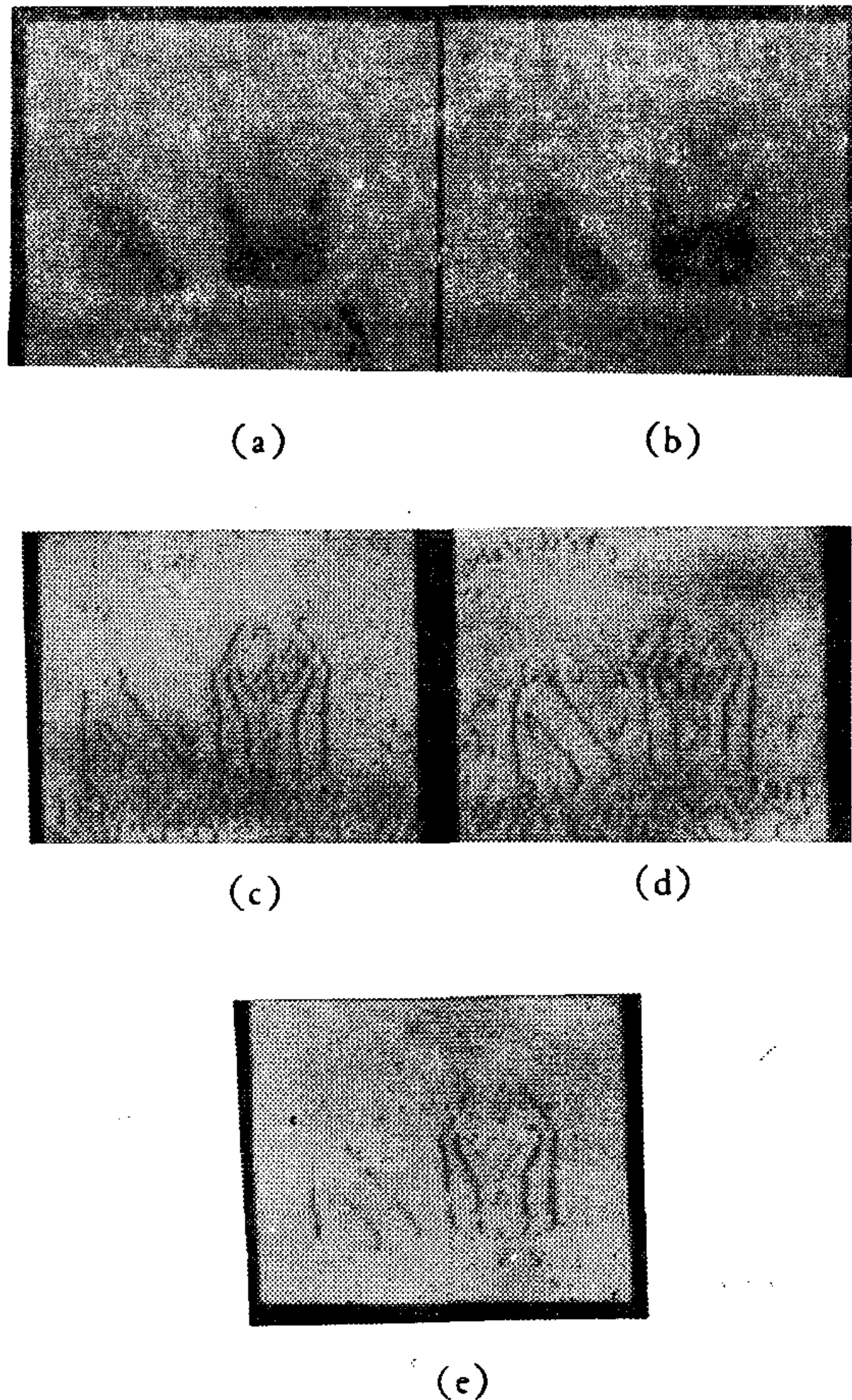


图 1

(a) 和 (b) 实验立体图对, (c) 和 (d) 用于匹配的左右零交图, (e) 匹配结果

的零交则大都获得了正确的匹配。

参 考 文 献

- [1] 姜 勤、潘士先, 一种松弛标记体视匹配算法, 信号处理, 4(1988), No. 4.
- [2] Baker, H. H., Depth from Edge and Intensity Based Stereo, Stanford Artificial Intelligence Lab., Tech. Rep., AIM-347, 1982.
- [3] Barnard, S. T. and Thompson, W. B., Disparity Analysis of Images, *IEEE Trans. PAMI*, 2(1980), 333—340.
- [4] Marr, D. and Poggio, T., A Computational Theory of Human Stereo Vision, *Proc. R. Soc. Lond. B.* 204, 301—328.
- [5] Mayhew, J. E. W. and Frisby, J. P., Psychophysical and Computational Studies towards a Theory of Human Stereopsis, *Art. Intell.*, 17(1981), 349—385.
- [6] Medioni, G. and Nevatia, R., Segment-based Stereo Matching, *Computer Vision, Graphics and Image Processing*, 31(1985), 2—18.
- [7] Ohta, Y. and Kanade, T., Stereo by Intra and Inter-scanline Search Using Dynamic Programming, *IEEE Trans. PAMI*, 7(1985), 139—154.

A TWO-STAGE ALGORITHM FOR STEREO MATCHING

LI WENBIAO PAN SHIXIAN

(Dept. of Electronic Engineering, Beijing University of Aero. & Astro., Beijing 100083, P. R. C.)

ABSTRACT

A two-stage iterative stereo matching algorithm is proposed. Primitives used for matching are those zerocrossings from a single ∇^2G channel that form connected edges due to their similarity in sign and orientation. The first stage is controlled by a disparity smoothness rule which substantiates Marr-Poggio's cohesiveness of the physical world. The second stage, driven by the output of the first stage, is governed by a disparity continuity rule that represents Mayhew-Frisby's figural continuity. Therefore, the algorithm yields matches that are locally and globally consistent in the sense of the two general constraints of stereo vision mentioned above.

Key words: Stereo matching; disparity smoothness; disparity continuity.



李文彪 1965 年出生, 云南昆明人。1990 年在北京航空航天大学获得工学硕士学位, 现为北京航空航天大学电子工程系博士研究生。研究方向为机器视觉和人工神经网络。



潘士先 1935 年出生, 浙江湖州人。1956 年毕业于天津大学电力系。1981—1983 年为美国 Purdue 大学访问学者。现为北京航空航天大学电子工程系教授。曾获 IEEE ASSP (声学、语音和信号处理) 学会 1984 年 Senior Award。研究领域为信号处理、图象重建和计算机视觉。