

# 从表观运动来看运动的基本约束<sup>1)</sup>

傅 浩 吴立德

(复旦大学计算机科学系, 200433)

## 摘 要

本文综述了生物学和心理物理学中有关运动视觉的发现, 并将它们总结为层次结构基元模型和九条运动基本约束。给出了一个以此为基础的运动分析算法和初步的实验结果。

**关键词** 运动, 对应, 另交叉, 基本约束, 心理物理学。

## 1 引言

最近的视觉生理心理学研究表明<sup>[1-5]</sup>人类的视觉系统由多条通道组成, 其中之一专用于处理运动信息, 因而可以相对独立地对运动视觉进行研究。

运动感知中的表观运动是由间隙变化的画面产生的运动。它实际上是一系列静止画面的快速变换。当人们在观察它们时, 自动进行插值而形成连续运动的错觉。计算机上所能获得或生成的运动都是表观运动。它主要在早期视觉中形成, 避免了高层知识的参与, 从而相当简化了问题的研究, 是个理想的研究对象。

本文第二节综述有关表观运动的一些新发现, 第三节据此引出层次基元模型和九条运动的基本约束。第四节给出了一个以此为基础的运动分析算法和初步的实验结果。

## 2 生物学上的一些发现

要产生表观运动, 其前提是找到“对应”。对应涉及两个基本问题: 对应所依据的特征和对应所采用的技术。人类视觉系统所选用的用于对应的一批特征和所采用的一批对应技术, 是由几百万年人类进化过程中经自然选择过程而积累起来的, 是很有效的, 因而值得借鉴。

### 2.1 特征选择<sup>[2, 6-8]</sup>

视觉系统分层次地选择特征。在不同的阶段选择不同的特征。首先选择的是一些空间低频区域(亮斑或暗斑), 其次选择明显的轮廓或边缘, 最后才选择特征点。

作为上述这些特征(区域、轮廓、点)的基础, 另交叉和纹理是两种重要而特殊的特征。

另交叉受到偏爱, 不仅因为其封闭性、不必细化、信息量丰富(如指纹定理)、“单调性”等一系列优越性<sup>[4, 9, 10]</sup>, 还因为生理心理学都证明了另交叉确实在人类视觉中存在并起着

本文于1991年10月21日收到。

1) 国家自然科学基金与高校博士点基金资助项目。

重要作用<sup>[4]</sup>.

纹理是另外一类重要特征. 实验表明, 仅仅依靠纹理也能进行对应<sup>[1,2]</sup>.

## 2.2 对应技术<sup>[2]</sup>

人类的视觉系统采用一系列的技术以帮助实现快速的对应.

第一种技术是先抽取特征, 在以后的分析中仅注意这些特征的变化, 从而大大减少了可能的匹配数.

第二种技术是将匹配限制在现实的三维世界中最可能出现的运动上. 作为这一技术的实例, 有如下的一些规则. 一是视觉系统通常假定正在运动着的物体继续保持匀速直线运动. 二是假定运动着的物体通常都是刚体, 从而其上的点享有共同的运动, 这也称为“运动捕获”. 三是假定前后两时刻中最相近的物体通常也是最可能对应的物体等等.

第三种技术是假定运动着的物体将逐步地遮住和放开背景的一部份. 它实际上是上述两种技术结合的一种推论.

第四种技术是在感知多物体运动时采用的. 这时视觉系统在没有明显的反证据时, 通常假定各物体以相同的方式运动. 这也就是格式塔心理学派所谓的“全场效应”或“乘车运动”.

下节中我们将对这些技术进行某种程度的形式化, 以便应用于计算机视觉中.

## 3 层次结构基元和运动的基本约束

### 3.1 层次结构基元

基于上述发现, 我们提出一个层次结构基元模型: 区域——轮廓——特征点<sup>[7,8]</sup>.

首先以背景中独立的不包含于任何其它区域的区域为特征基元, 获取这些区域的对应以得到初步的运动信息. 然后对每一对匹配了的区域找轮廓 (包括围成区域本身的轮廓和勾勒区域内部细节的轮廓), 并对这些轮廓找对应, 从而得到进一步的运动信息. 最后进一步找出轮廓上的曲率极值点, 进行特征点对应, 得到更精细的运动描述. 这一模型完全遵循了由粗及精的运动感知过程, 与人类视觉系统相一致.

### 3.2 运动的基本约束

无论是区域对应, 还是轮廓对应, 或者是点对应, 都会出现多对多匹配, 我们也必须像视觉系统一样, 引入基本约束.

为表述简洁, 先引入几个记号. 令  $V_i^n$  为第  $n$  幅图象上第  $i$  个特征基元 (区域, 轮廓或点, 下同) 的速度.  $C_i^n$  为第  $n$  幅图象上第  $i$  个特征基元的性质 (如形状、灰度、结构等). 九条运动的基本约束现在可以描述如下:

#### 3.2.1 基元的分层法则

“对应”由区域开始, 经轮廓而到特征点.

#### 3.2.2 基元性质相似性法则

对应的基元的性质要相似, 即

$$\sum_{i,j \text{ 为匹配对}} \|C_i^n - C_j^{n+1}\| = \min$$



其中  $\|C_i^n - C_i^{n+1}\|$  表示  $C_i^n$  和  $C_i^{n+1}$  之间差异的某种测度。

### 3.2.3 运动捕获法则

同一区域或轮廓内的特征享有共同运动,即

$$\sum_{\substack{\text{特征基元 } j \in \\ \text{特征基元 } i}} \|V_j^n - V_i^n\| = \min$$

### 3.2.4 平滑运动法则

在没有明显的反证据时,物体保持原有运动,即

$$\sum_n \|V_i^n - V_i^{n+1}\| = \min$$

### 3.2.5 乘车运动法则

在没有明显的反证据时,假定不同的物体有相同的运动,即

$$\sum_{i \neq j} \|V_i^n - V_j^n\| = \min$$

### 3.2.6 最近法则

$n$  时刻画面上的特征基元通常在下一时刻画面上的相近位置处出现,即假定运动不大,

$$\sum_i \|V_i^n\| = \min$$

### 3.2.7 背景遮挡法则

在对应过程中,常常会发生不完全对应或无对应情况,遮挡这时是一种很好的解释,非常符合人类视觉的习惯。它通常要求满足下列两个条件:

(1) 满足运动捕获法则和平滑运动法则,即

$$\sum_{\substack{\text{在图象 } n \text{ 上,特征基元 } j \in \text{特征基元 } i \\ \text{在图象 } n+1 \text{ 上,特征基元 } j \text{ 不出现}}} (a\|V_j^n - V_i^n\| + (1-a)\|V_j^n - V_j^{n+1}\|) \\ + \sum_{\substack{\text{在图象 } n \text{ 上,特征基元 } j \text{ 不出现} \\ \text{在图象 } n+1 \text{ 上,特征基元 } j \in \text{特征基元 } i}} a\|V_j^n - V_i^n\| = \min$$

其中  $a \in [0, 1]$ , 它的取值决定于匹配情况,当  $i$  失配时,  $a = 0$ , 当  $i$  匹配较理想时取较大值。

(2) 若  $j$  为被遮挡特征基元, 则要求  $V_j^n$  使得  $j$  处于其它特征基元的覆盖之下。

### 3.2.8 多分辨率另交叉法则

区域——轮廓——特征点是对应过程中的一个层次。不同分辨率的另交叉是对应过程中的另一个层次。每个物体的大小、灰度和复杂程度是不同的,因此要求的观察精度也不同。于是要选择不同分辨率的  $\sigma$  来作另交叉,最后再根据对应的好坏来选择。

### 3.2.9 整体法则

运动捕获是一个全局法则,而最近法则是一个局部法则。它们看似矛盾,其实则不然。因为这些法则虽同时存在,但在不同场合其优先级不同,作用强弱不同,执行时序也不同。

整体法则为视觉系统中的每一法则分配一个权  $w_i$  ( $0 \leq w_i \leq 1, \sum_i w_i = 1$ ), 根据每一法则的优先级不同,  $w_i$  的取值不同.  $w_i$  还可以是时间的函数, 以体现同一法则在不同时刻的作用强度.

## 4 一个基于人类视觉系统的运动分析算法

给出一个图象序列, 我们对  $n$  时刻和  $n+1$  时刻的图象进行分析. 首先选择  $K$  个分辨率作另交叉, 建立区域——轮廓——特征点结构. 然后分三个层次进行匹配.

### 4.1 区域匹配

区域匹配对每一分辨率进行. 设时刻  $n$  时的图象在第  $\sigma_k$  分辨率下共有  $I1_k$  个区域. 记第  $i1$  个区域为  $G(n, k, i1)$ , 其性质为  $C_{i1}^n(k)$ . 类似地, 时刻  $n+1$  时的图象在第  $\sigma_k$  分辨率下共有  $J1_k$  个区域, 记第  $j1$  个区域为  $G(n+1, k, j1)$ , 其性质为  $C_{j1}^{n+1}(k)$ . 这里区域的性质综合了形状, 结构和灰度等方面的信息.

根据运动捕获法则, 假设每个区域享有共同的运动. 若区域  $G(n, k, i1)$  和区域  $G(n+1, k, j1)$  对应, 则它们中心的位移除以时间间隔  $\Delta t$  所得的速度, 记为  $V1_{i1}^n(k, j1)$ , 即为所要求的速度.

为找到对应, 要从

$$V1_{i1}^n(k, j1), 1 \leq i1 \leq I1_k, 1 \leq j1 \leq J1_k, 1 \leq k \leq K.$$

中挑选一组最佳的速度表示

$$V1_{i1}^n(k_r, j1_r), 1 \leq i1 \leq I1_{k_r}, j1_r \in [1, J1_{k_r}], k_r \in [1, K]$$

并用它们来产生最终的对应和速度  $\{V1_{i1}^n\}$ .

对应首先要求得到尽可能多的区域匹配, 即失配率  $N1$  要尽可能小. 若将一些区域合并作为一个整体来处理时能降低失配率, 则应这样做, 并同时对它们的轮廓也作恰当的合并. 对应还要满足前面提到的基本约束, 特别是基元性质相似法则, 平滑运动法则, 最近法则和乘车运动法则.

最后按整体法则, 可在  $K$  个分辨率之间由以下要求获取最佳对应和最佳速度表示  $\{V1_{i1}^n\}$ :

$$\begin{aligned} & w_1 \cdot N1 + w_2 \cdot \sum_{i, j \text{ 匹配}} \|C_i^n(k) - C_j^{n+1}(k)\| + w_3 \cdot \sum_i \|V1_i^n\| \\ & + w_4 \cdot \sum_{i1} \|V1_{i1}^n - V1_{i1}^{n+1}\| + w_5 \cdot \sum_{i1 \neq i1'} \|V1_{i1}^n - V1_{i1'}^n\| = \min \end{aligned}$$

这里  $w_i$  为权系数, 可根据不同的场景、不同的需要来控制. 无信息时, 不妨取  $w_i = \frac{1}{5}$ ,

$1 \leq i \leq 5$ .

对于失配的区域用背景遮挡法则求取速度.

### 4.2 轮廓匹配

进行了区域匹配之后, 继续对围成区域的轮廓和表示区域内部细节的轮廓进行匹配.

由于我们采用了六邻域四个池的跟踪方式<sup>[7]</sup>，所得到的轮廓都是封闭且互不相交的。每个区域的轮廓以包含关系为父子关系形成一个树状结构。匹配在两个树状结构间进行。

若区域  $G(n, k, i1)$  和区域  $G(n + 1, k, j1)$  相匹配,且区域  $G(n, k, i1)$  共有  $I2_{i1}$  个轮廓,其第  $i2$  轮廓  $L(n, k, i1, i2)$  的性质为  $C_{i2}^n(k, i1)$ 。类似地,区域  $G(n + 1, k, j1)$  共有  $J2_{j1}$  个轮廓,其第  $j2$  轮廓  $L(n + 1, k, j1, j2)$  的性质为  $C_{j2}^{n+1}(k, j1)$ 。

与区域匹配相类似,假设每个轮廓享有共同的运动。若轮廓  $L(n, k, i1, i2)$  和轮廓  $L(n + 1, k, j1, j2)$  相匹配,则它们的中心的位移除以  $\Delta t$  所得的速度,记为  $V2_{i2}^n(k, i1, j2)$ ,即为所要求的轮廓  $L(n, k, i1, i2)$  的速度。

与区域匹配时相仿,可按下列要求从

$$V2_{i2}^n(k, i1, j2), 1 \leq i2 \leq I2_{i1}, 1 \leq j2 \leq J2_{j1}$$

中挑出一组最佳的速度表示

$$V2_{i2}^n(k, i1, j2_r), 1 \leq i2 \leq I2_{i1}, j2_r \in [1, J2_{j1}]$$

并用它们来产生最终的对应和速度  $\{V2_{i2}^n\}$ 。

$$\begin{aligned} & w1 \cdot N2 + w2 \cdot \sum_{\substack{i1 \text{ 与 } j1 \text{ 匹配} \\ i2 \text{ 与 } j2 \text{ 匹配}}} \|C_{i2}^n(k, i1) - C_{j2}^{n+1}(k, j1)\| \\ & + w3 \cdot \sum_{\text{轮廓 } i2 \in \text{区域 } i1} \|V1_{i1}^n - V2_{i2}^n\| + w4 \cdot \sum_{i2 \neq i2'} \|V2_{i2}^n - V2_{i2'}^n\| \\ & = \min \end{aligned}$$

其中  $N2$  是轮廓的失配率,  $w_i, 1 \leq i \leq 4$  是权系数。

对于失配的轮廓也用背景遮挡法则来求取速度。

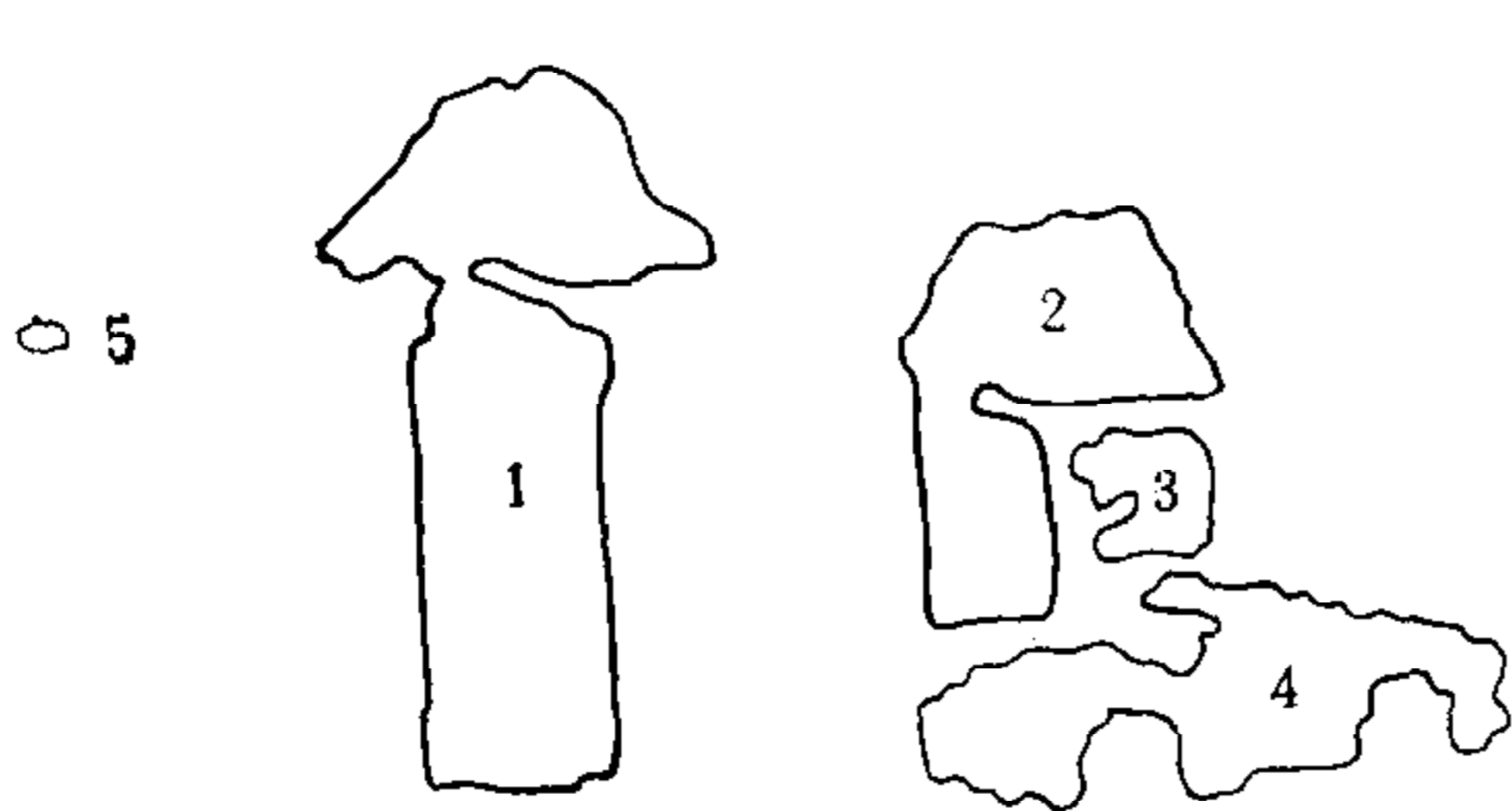


图 1 a

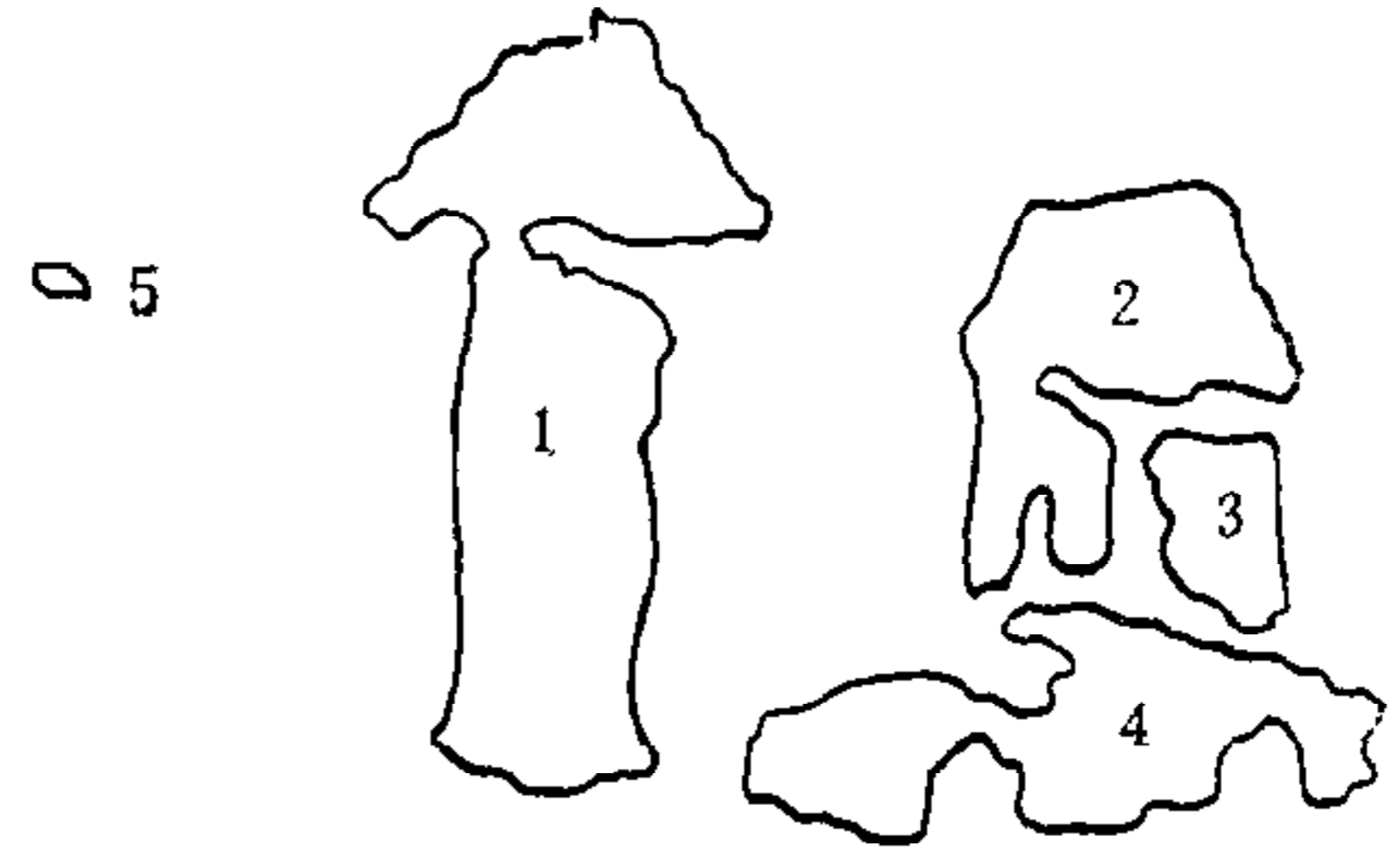


图 1 b

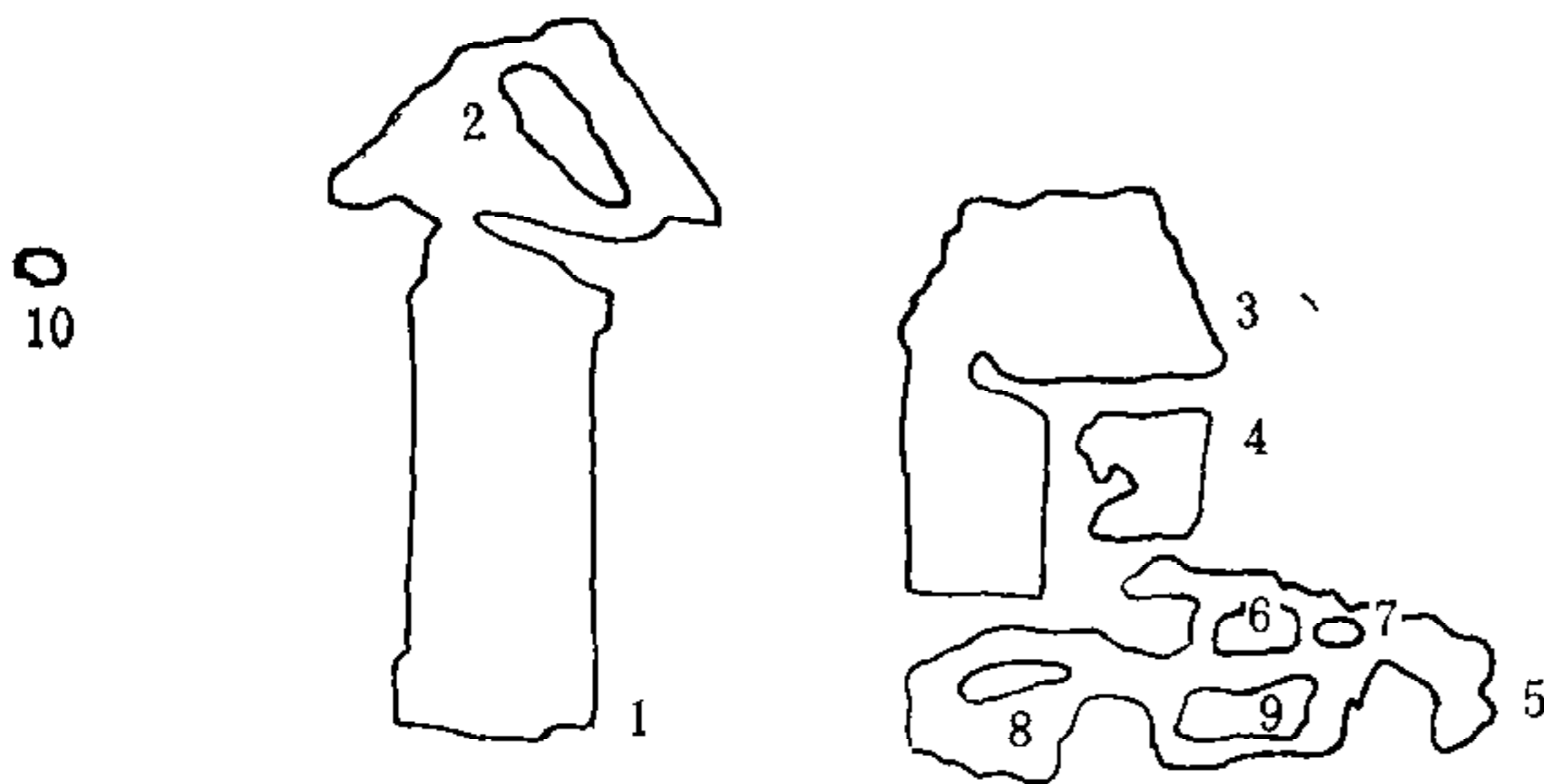


图 2 a

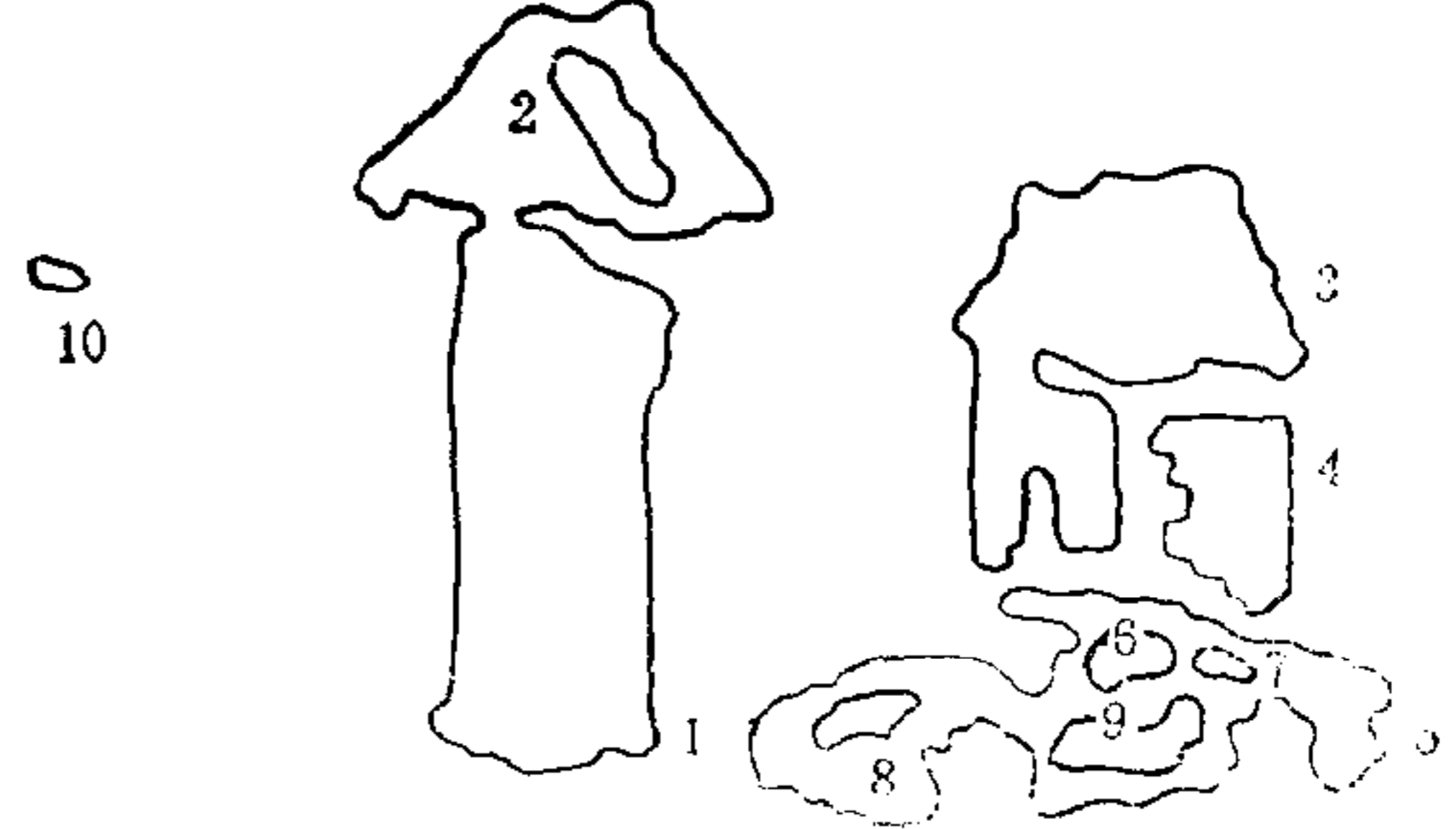


图 2 b



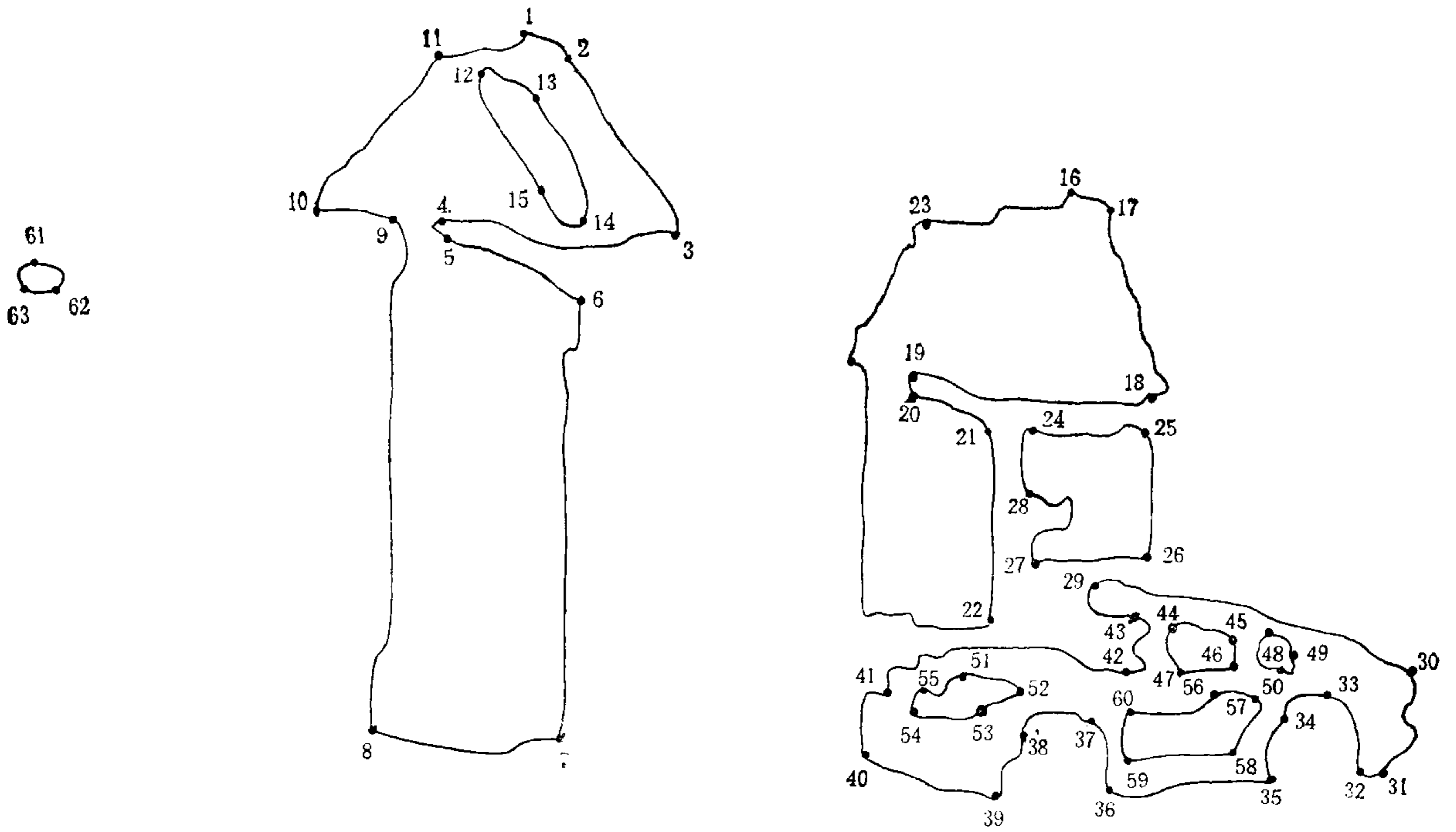


图 3 a

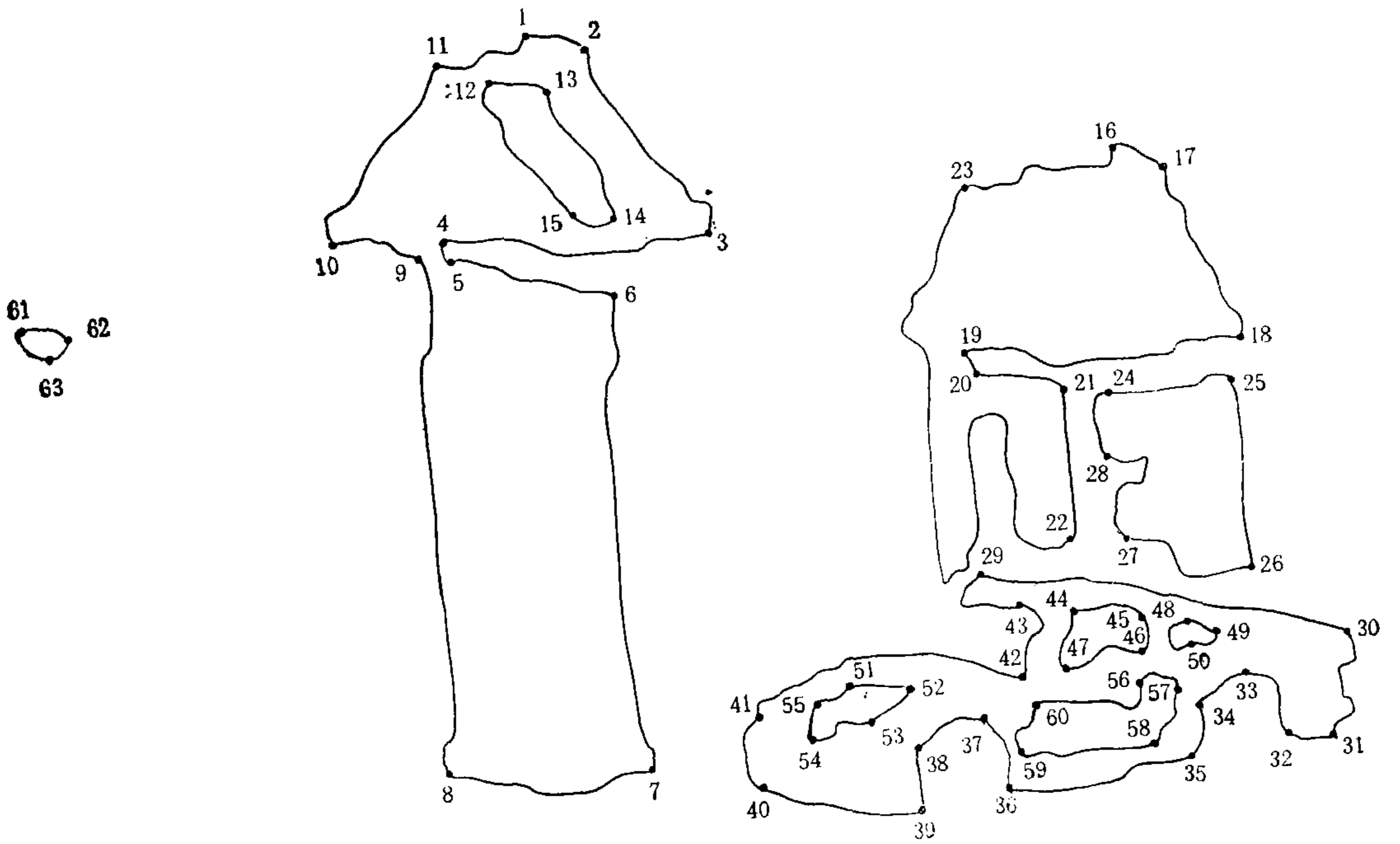


图 3 b

### 4.3 特征点匹配

进行了轮廓匹配后,继续对轮廓上的特征点进行匹配。

若轮廓  $L(n, k, i_1, i_2)$  和轮廓  $L(n+1, k, j_1, j_2)$  相匹配,且轮廓  $L(n, k, i_1, i_2)$  共有  $I_{3i_2}$  个特征点,其第  $i_3$  特征点  $P(n, k, i_1, i_2, i_3)$  的性质为  $C_{i_3}^n(k, i_1, i_2)$ ; 轮廓  $L(n+1, k, j_1, j_2)$  共有  $J_{3j_2}$  个特征点,其第  $j_3$  特征点  $P(n+1, k, j_1, j_2, j_3)$  的性质为  $C_{j_3}^{n+1}(k, j_1, j_2)$ 。

若特征点  $P(n, k, i_1, i_2, i_3)$  与特征点  $P(n+1, k, j_1, j_2, j_3)$  相匹配,则它们的位移除以  $\Delta t$ , 记为  $V_{3i_3}^n(k, i_1, i_2, j_3)$ , 即为所要求的特征点  $P(n, k, i_1, i_2, i_3)$  的速度。

与区域和轮廓匹配时相仿,可按下列要求从  $V_{3i_3}^n(k, i_1, i_2, j_3)$ ,  $1 \leq i_3 \leq I_{3i_2}$ ,  $1 \leq j_3 \leq J_{3j_2}$  中挑出一组最佳的速度表示:

$$V_{3i_3}^n(k, i_1, i_2, j_3), 1 \leq i_3 \leq I_{3i_2}, j_3 \in [1, J_{3j_2}].$$

并用它们来产生最终的对应和速度  $\{V_{3i_3}^n\}$ :

$$\begin{aligned} w_1 \cdot N_3 + w_2 \cdot \sum_{\substack{i_2 \text{ 与 } j_2 \text{ 匹配} \\ i_3 \text{ 与 } j_3 \text{ 匹配}}} \|C_{i_3}^n(k, i_1, i_2) - C_{j_3}^{n+1}(k, j_1, j_2)\| \\ + w_3 \cdot \sum_{\substack{\text{特征点 } i_3 \in \text{轮廓 } i_2}} \|V_{2i_2}^n - V_{3i_3}^n\| + w_4 \cdot \sum_{i_3 \neq i_3'} \|V_{3i_3}^n - V_{3i_3'}^n\| = \min \end{aligned}$$

其中  $N_3$  是特征点的失配率,  $w_i, 1 \leq i \leq 4$  是权系数。

对于失配的特征点用背景遮挡法则求速度。

以上给出了一个基于人类视觉系统的运动分析方法的总体思路,其实施细节将另文发表。图 1、图 2 和图 3 分别是 从一对运动图象序列中用本方法求得的匹配了的区域,匹配了的轮廓,以及匹配了的特征点。实验的结果是令人满意的。

## 5 结论

计算机视觉是建立在人类视觉系统和计算机科学技术上的一门学科,生物学和心理物理学对于人类视觉系统的研究,无疑对于计算机视觉的研究起着重要作用。本文针对计算机视觉中的一个模块——运动分析模块作了详细研究,从而获得了具有生物学和心理物理学基础,并且实验结果良好的层次结构基元模型和一系列运动的基本约束。

### 参 考 文 献

- [1] Livingstone, M. S., Art Illusion and the vision system, *Scientific American*, 1988, 258(1):78—85.
- [2] Ramchandran, V. and Anstis, S., The Perception of Apparent Motion, *Scientific American*, 1986, 254(6):80—86.
- [3] Hildreth, E. C., *The Measurement of Visual Motion*, MIT Press, Cambridge, Mass., 1984.
- [4] Marr, D., *Vision*, W. H. Freeman and Company, San Francisco, 1982.
- [5] Ullman, S., *The Interpretation of Visual Motion*, MIT Press, Cambridge, Mass., 1973.
- [6] 傅洁,吴立德,计算机视觉中运动分析的连续处理方法综述,模式识别与人工智能,1991,4(1): 91—100.
- [7] 傅洁,吴立德,基于围线的二维运动计算,模式识别与人工智能,1991,4(2): 1—10.
- [8] Wang, X., Fu, J., Wu L., A Matching Algorithm Based on Hierchical Primitive Structure, Proc.

10th ICPR, Atlantic City, NJ., 285—288.

[9] Wu, L., Xia, S., Scaling Theorems for Zero-Crossings, *IEEE Trans. PAMI*, 1990, 12(1):46—54.

[10] 谢绍辉, 吴立德, 关于  $\nabla^2 G$  算子的一些研究, *自动化学报*, 1990, 16(3): 193—202.

## ON THE BASIC MOTION CONSTRAINTS FROM APPARENT MOTION

FU JIE WU LIDE

(Department of Computer Science, Fudan University, Shanghai 200433)

### ABSTRACT

In this paper the discoveries on motion vision in biology and psychophysics are reviewed and summarized as an hierarchical primitive model with nine basic motion constraints. A motion analysis algorithm based on the model and constraints is presented.

**Key words:** motion; correspondence; zero-crossings; Basic Constraints; psychophysics.

傅 洁 照片及简介见本刊第19卷第4期。



吴立德 简介见本刊第19卷第4期。