

视觉运动信息处理神经系统*

郭雷 郭宝龙 陆心如

(西安电子科技大学, 西安 710071)

摘要 运动判决需要使用整体性质. 本文建立了一个神经系统, 并给出了该神经系统实现图象区域运动判决的原理. 该系统的特点是从整体性质出发来判断和表示运动信息. 这个系统主要采用的神经机制是: 信息分类表示和信息分类加工; 依据简单特征的运动匹配; 不同信息之间的竞争与合作; 神经活性的扩散、集中和阻挡.

关键词 视觉运动*, 神经网络, 神经机制*.

运动视觉的根本任务是依据投影图象中的二维运动, 判断和理解视场中三维物体的结构和运动. 根据这个观点, Ullman 将运动视觉的处理分为 2 个阶段^[1]: 首先是图象中二维运动处理(一致性对应), 然后依据二维运动推断三维结构. 本文主要讨论 Ullman 的第 1 步: 用神经系统判决图象区域的运动.

视觉刺激的局部信息在知觉上是模糊的, 例如刺激的位置, 亮度、定向^[2-5]及运动性质^[6]. 关于运动的一个著名的例子叫小孔问题(aperture problem), 它表明运动方向是整体性质^[7-9]. 这个问题可以从光流(速度场)和运动一致性两个方面解释. 最近一些研究者认为小孔问题中的条件是苛刻的, 实际情况不存在^[10]. 然而这种苛刻的条件是重要的情况, 因为人类视觉系统具有折扣(discount)亮度的功能: 将区域的亮度理解为均匀的^[11]. 这种理解过程被解释为亮度的填充而达到均匀的分布^[2,3,11,12], 这表明: 区域内亮度的弱变化对运动知觉并不影响, 并且填充能抑制局部噪声.

显然, 区域点的运动判决是比边界点的运动判决更高阶的知觉模糊问题. 在一个均匀亮度的区域内, 一个点可和它周围任一点亮度匹配, 特征上无法区分; 而一条直线的边界点只能和它周围边界上的点匹配, 因为边界点有边界强度及朝向的特征. 从速度场的观点来看, 在边界点上至少沿梯度方向的导数不为零; 而在区域点上, 所有方向的偏导数都为零; 区域的运动性质要受到包含它的轮廓的运动性质的限制.

小孔问题的本质不在于它的条件是特殊的, 而在于它表达了这样一个基本观点: 一个大尺寸的图象结构的运动判决是整体性质, 这是人类理解运动的基本点. 本文将沿着这样的观

* 本文 1992-01-04 收到, 1992-04-17 定稿

本文是国家自然科学基金和 863 高技术项目. 作者郭雷, 38 岁, 副教授, 主要研究领域为神经网络, 智能信息处理. 郭宝龙, 32 岁, 讲师, 主要研究领域为智能信息处理, 神经网络. 陆心如, 女, 62 岁, 教授, 主要研究领域为图像处理, 计算机视觉.

本文通讯联系人: 郭宝龙, 西安 710071, 西安电子科技大学 284 信箱

点,利用整体性质来判决运动.

1 基本原理

Ullman 认为图象结构的运动判决分为 2 个步骤:首先抽取特征,然后实现该特征在不同图象帧间的匹配,并计算运动参数,在这个过程中,误配是不允许的,因此所抽取的特征实际上应具有整体性质.因此特征的选择是运动一致性的难点.

本文将遵照一致性原理,然而方法却不同.我们将 Ullman 的原理分为 3 步:(1)抽取简单特征.这些特征是局部性质,如边界强度,边界朝向及点的亮度等,不要求具有与众不同的整体性质,因此易于实现特征抽取;(2)实现简单特征在不同帧间的匹配和对应.依据简单特征的匹配将不可避免地产生误配,因此,(3)消除误配.这三步合起来就完成了 consistency 匹配.不难看出这种匹配过程的关键在于允许误配发生,这就降低了对匹配特征的要求,使问题难度降低.

视觉运动信息是分类表示和分类加工的.根据神经生物学,一个神经元只对特定的信息敏感^[13,14].换句话说,一个神经元只表示和加工特定的视觉信息.从信息处理的观点来看,在神经系统中,视觉信息首先被分类,然后按照信息的类型沿不同通道区分地加工不同类型的视觉信息^[15]这样,视觉信息被区分且根据类型分别采用各自的加工方式.这是生物神经系统完成信息处理的一个重要特征.在人工神经系统中也延用了这个原理.例如,在边界轮廓系统中^[2-5],不同朝向的边界是区分地表示和加工的.在运动知觉系统中,不同性质的运动也是被区分地表示和加工^[6].这种具有生物学基础的信息加工方式将带来很大的益处,它在本文所建立的理论中占有十分重要的地位.下面,让我们来观察一个轮廓的运动.

考虑该轮廓上的点在两帧中的对应匹配.对于每一个轮廓点,可利用的特征只有 2 个:该点的边界强度和边界朝向,这些是简单特征.显然,第一帧中轮廓上的一点可以和第二帧的上许多轮廓点实现特征匹配.所以说,依据于简单特征,误配是不可避免的.

然而,在这些所有的匹配中,相互之间都有差别:每一个匹配都对应与其它匹配不同的运动方向和运动位移.这是一个本质性的差别,它反映了运动性质和匹配对应之间的内在关联性.因此,依据这个差别,我们不从每个轮廓点有什么样的特征对应这条思路去考虑问题,而是反过来,考虑在每一种可能的运动之下,该轮廓上有多少点实现运动匹配.也就是说:当给定一个运动(运动方向为 d 和运动位置 ΔX)之后,轮廓上的每一点沿方向 d 、距离 ΔX 是否能实现前后二帧的上述简单特征的匹配.十分明显,结论是只有当给定的运动正好和实际的运动吻合时,轮廓上的所有点均实现二帧中的特征对应,我们称之为满匹配;当给定的运动与实际运动不一致时,必然出现失配点.因此,当区分运动类型并对所有可能的运动分类匹配处理后,那么满匹配所对应的运动就是实际运动,满匹配给出运动判决.这样,依据运动信息分类,就把运动判决问题转化为发现满匹配问题.对于一个运动轮廓来说,满匹配意味着匹配点形成原来的封闭轮廓(运动轮廓).

对一个图象运动区域,本文将依据 3 个方面的因素确定实际运动:(1)区域轮廓点的匹配数量优势;(2)轮廓匹配点的轮廓封闭性;(3)区域点亮度特征的匹配数量优势.在噪声存在时,个别点的特征将引起奇变,绝对的满匹配将不存在.因此本文中采用匹配点数量优势(最多)作为判决依据.上述 3 个判决因素并不是相互独立实施的,它们将以综合的形式统一

完成.

2 神经系统组成和工作原理

运动判决神经系统结构如图 1 所示. 该系统主要由 3 个部分组成:(1)图象边界和区域生成系统(MP,BCS 和 RCS);(2)运动轮廓生成系统(EMD 和 MCS);(3)区域运动系统 FCS,EMD 和 RMS). 三部分最后在 RMS 中综合,给出区域运动的判决.

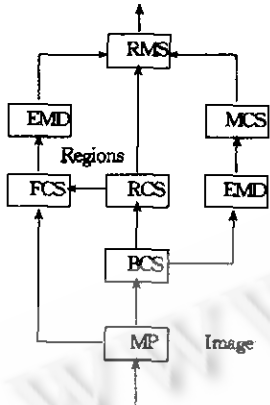


图1 神经系统组成

- MP(Monocular Preprocessing): 预处理系统
- BCS(Boundary Contour System): 边界轮廓系统
- FCS(Feature Contour System): 特征轮廓系统
- RCS(Region Contour System): 区域轮廓系统
- EMD(Elementary Motion Detector): 初级运动检测器
- MCS(Motion Contour System): 运动轮廓系统
- RMS(Region Motion System): 区域运动系统

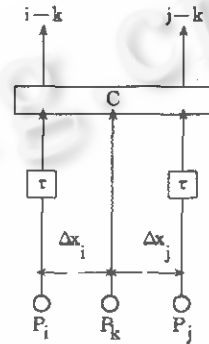


图2 EMD原理结构

依据神经心理学,神经活性的扩散和集中是脑的高级神经活动的基本规律之一^[17]. 扩散、集中加阻挡神经机制是本文依据的又一重要的原理. 扩散是指神经活性以流动的方式向周围相邻的神经元传播^[18,12,19]. 扩散的本质在于发现区域连通性. 在本文中扩散将用于发现运动.

扩散过程有几个方面的控制因素:(1)扩散信号,即扩散源. 它可采用亮度信号^[2-5,12], 外部注意输入^[19,20],和边界信号. 在本文中,扩散信号是匹配边界和匹配亮度.(2)扩散信号的注入位置,它决定扩散开始位置.(3)扩散阻挡信号,一般来源于边界信息^[12,19]及运动边界. 阻挡信号起限制扩散区域的作用.(4)阻挡位置,应为边界及运动边界的位置^[12,19].

如图 1 所示,本系统分为左中右三路. 左路是发现亮度特征的运动,中路提供区域位置,右路提供轮廓特征的运动,三者最终在区域运动系统 RMS 中综合. 下面就左右两路分别介绍,中路见文献^[19].

2.1 右路:依据轮廓的运动匹配

在本系统中,BCS 完成轮廓运动一致性匹配的第 1 步:发现边界特征,并作为右路 EMD 的输入. 图 2 中 P_k, P_i 和 P_j 为 BCS 中相应位置的边界信息. EMD 作为运动一致性的第 2 步,完成边界特征的初级运动匹配. EMD 的输出表示运动的当前位置,运动方向和运动速度(位移). 运动轮廓系统(MCS)是运动分层结构,每一层是 $k \times k$ 神经元陈列,用来表示和处理一种运动. EMD 将根据每个匹配的运动性质,将输出送至 MCS 的相应层及相应位置. 例如,在图 2 中,若 P_k 和 P_j 的特征相同,则送兴奋输出至 MCS 中处理 $j \rightarrow k$ 的运动层(该层专门处理运动方向为从 j 到 k 距离的为 Δx_i 的运动)中位于 k 位置的神经元. 具有不同的运动方向或运动位移的运动将送至 MCS 中不同的层. 在 MCS 中,同一层(同运动)中兴奋^的神经元之间是相互合作的,不同层(不同运动)但相同位置的神经元之间是相互抑制的. 这

样,在每一层中,兴奋(被 EMD 的输出激活)的神经元越多,相互合作越强大,结果是镇压掉合作相对弱的其它层神经元,正确运动在自身的合作和镇压误匹配运动的过程中,变得最为强大。

MCS 是完成运动一致性的第 3 步中的第 1 阶段:根据匹配点的数量优势消除误匹配。RMS 完成其第 2 阶段:根据运动轮廓的封闭性,进一步消除误匹配。

2.2 左路:依据亮度的运动匹配

右路是以轮廓形状为依据的运动匹配,然而它不能正确判别形状相同而亮度不同的 2 个物体的运动。因此亮度特征的利用是必然的。

在左路中,特征轮廓系统 FCS 作为亮度特征运动一致性的第一步:发现亮度特征。FCS 是一个单层 $k \times k$ 神经元阵列^[2-5,12],相邻神经元之间为兴奋型连接。亮度信息在阵列内扩散,BCS 中的边界抑制 FCS 中相应位置的权,用来阻挡扩散过程。左路 EMD 的输入为每个位置的均匀亮度,EMD 完成亮度匹配,并将亮度匹配信号依运动性质送至 RMS 中的相应层及相应位置。误匹配在 RMS 中消除。

3 系统方程

本节给出图 1 中各子系统的数学描述,它是本系统的细节表达。EMD,BCS,FCS 和 RCS 的方程见文献^[12,16,19]。

3.1 MCS 方程

设 MCS 共有 M 层神经元二维阵列,在第 P 层(i,j)位置的神经元的状态为 y_{ij}^p ,其方程为:

$$du_{ij}^p/dt = \sum_{i'} \sum_{j'} T^p(i,j,i',j') y_{i'j'}^p + \sum_{q \neq p} W_{ij}^{pq} y_{ij}^q - B y_{ij}^p + C S_{ij}^p, B > 0, C > 0 \quad (1)$$

其中 S_{ij}^p 为右路 EMD 中在(i,j)位置的第 P 群神经元的输出:

$$y_{ij}^p = \begin{cases} u_{ij}^p & u_{ij}^p > 0 \\ 0 & u_{ij}^p \leq 0 \end{cases} \quad (2)$$

$T^p(i,j,i',j') = T^p(i',j',i,j) \geq 0$ 为第 P 层内的连接权:

$$T^p(i,j,i',j') = \begin{cases} \alpha > 0 & S_{ij}^p > 0 \text{ 且 } S_{i'j'}^p > 0 \\ 0 & \text{其它} \end{cases} \quad (3)$$

$W_{ij}^{pq} \leq 0$ 表示 MCS 中两层 P 和 q 中在相同位置(i,j)之间的抑制连接。

3.2 RMS 方程

RMS 为 M 层神经元阵列,每一层有 $k \times k$ 个神经元。在阵列内的位置(i,j), ($2 \leq i \leq k-1, 2 \leq j \leq k-1$) 第 P 层的神经元服从扩散——集中方程:^[12,18,19]

$$dV_{ij}^p/dt = \sum_{i'} \sum_{j'} T_{+}^p(i,j,i',j') (z_{i'j'}^p - z_{ij}^p) + \sum_{i'} \sum_{j'} T_{-}^p(i,j,i'',j'') (Z_{i'j''}^p - Z_{ij}^p) - B V_{ij}^p + \sum_{q \neq p} W_{ij}^{pq} Z_{ij}^q + I_{ij}^p \quad (4)$$

式中 Z_{ij}^p 表示 RMS 中第 P 层在(i,j)位置的神经元的状态:

$$Z_{ij}^p = \frac{A}{1 + e^{-\beta V_{ij}^p}} \quad \beta > 0 \quad (5)$$

Z_{ij}^p 和 V_{ij}^p 之间是一个非线性关系. [21] (4)式中, $T_+^p(i, j, i', j') = T_+^p(i', j', i, j) \geq 0$ 为相邻神经元之间的连接权:

$$T_+^p(i, j, i', j') = Te^{-\lambda E_{ij}^p} \tag{6}$$

其中 E_{ij}^p 为 MCS 中的 P 层在 (i, j)和相邻位置之间的神经元状态之差. (4)中 $T_-^p(i, j, i'', j'') = T_-^p(i'', j'', i, j) < 0$ 表示 RMS 的 P 层中间隔神经元之间的连接, W_{ij}^p 为 RMS 中不同层但同位置的神经元之间的抑制连接($W_{ij}^p < 0$).

$$I_{ij}^p = \begin{cases} a_1 E_{ij}^p + a_2 S_{ij}^p & Q_{ij} \text{兴奋} \\ 0 & Q_{ij} \text{抑制} \end{cases} \tag{7}$$

上式 S_{ij}^p 为左路 EMD 的输出, Q_{ij} 为 RCS 在 (i, j)位置的神经元状态, Q_{ij} 兴奋表示 (i, j)位置为区域内部. a_1 和 a_2 为正的常数.

对于本文建立的运动判决神经系统,我们做了多组原理性计算机模拟实验,验证了其正确性.

参考文献

- 1 Ullman S. The interpretation of visual motion. MIT Press, 1979.
- 2 Grossberg S, Mingolla E. Psychol. Rev., 1985(92):173-211.
- 3 Grossberg S, Mingolla E. Percep. Psychophys, 1985(38):141-171.
- 4 Grossberg S. Percep. Psychophys, 1987(41):87-116.
- 5 Grossberg S. Percep. Psychophys, 1987(41):117-158.
- 6 Marshall J A. Neural Networks, 1990(3):45-74.
- 7 Marr D, Ullman S. Proc. R. Soc. Lond. B, 1981(211):151-180.
- 8 Hildreth E C. The measurement of visual motion. MIT Press, 1984.
- 9 Marr D. Vision, W H. Freeman and Company. San Francisco, 1982.
- 10 Uras S, Girosi F, Verri A et al. Biol. Cybern., 1988(60):79-87.
- 11 Paradiso M A, Nakayama K. Vision Rev., 1991(31):1221-1236.
- 12 Grossberg S, Todorovic D. Percep. Psychophys, 1988(43):241-277.
- 13 Hubel D H, Wiesel T N. J Physiology, 1962(160):106-154.
- 14 Hubel D H, Wiesel T N. Proc. R. Soc. Lond., 1977(198):1-59.
- 15 Livingstone M S. Scientific American, 1988(258):78-85.
- 16 Reichardt W. Sensory Communication. Rosenblith W A ed., Wiley, New York, 1961.
- 17 祝蓓里等. 心理学. 上海: 华东师范大学出版社, 1984.
- 18 Guo Lei, Biol. Cybern., 1990(64):61-67.
- 19 郭雷, 陆心如. 神经网络专集. 西安电子科技大学学报, 1991(增刊):96-110.
- 20 Kienker P K, Sejnowski T J. Perception., 1988(15):197-216.
- 21 Hopfield J J. Proc. Nat. Acad. Sci. USA, 81, 3088-3092, 1984.

NEURAL SYSTEMS FOR THE INFORMATION PROCESSING OF VISUAL MOTION

Guo Lei, Guo Baolong and Lu Xinru

(Xi'an University of Electronic Science and Technology, Xi'an 710071)

Abstract The global properties need to be used in motion decision. A neural system and its basic principle for motion decision are developed in this paper. The main specialties of the system are to utilize the regional global properties to decide and represent motion information. The essential neural mechanisms on which the system based are; representing and processing separately visual information; the motion matching according to local features; the cooperation and competition of different information, and the diffusion, concentration and blocking of neural activity.

Key words Visual motion, neural network, neural mechanism.