

一种基于细胞动力学的双目匹配方法*

彭雄宏 张翠 王正志

(国防科技大学自动控制系 长沙 410073)

摘要 双目匹配是计算机立体视觉的一个难点。以往的各种模型把注意力集中于将左视图像和右视图像中的要素特征唯一地进行双目匹配的任务上。有别于以往的这些模型,文章在 FACADE 模型框架的基础上,利用细胞动力学模型隐含了竞争和合作这一原理,实现了一个采用“方位合作,视差竞争”方法的双目匹配模型,从而探索一条利用生物视觉原理进行计算机立体视觉研究的可行途径。

关键词 双目匹配,细胞动力学,FACADE,生物视觉

分类号 TP391.41

A Method of Binocular Matching Based on Cell Dynamics

Peng Xionghong Zhang Cui Wang Zhengzhi

(Department of Automatic Control, NUDT, Changsha, 410073)

Abstract Binocular matching is a key problem in the field of computer vision. In this paper, we present a model based on cell dynamics, which is from the FACADE theory, to realize binocular matching by a method of “cooperation within orientation plane, competition between disparity plane”. The model helps us to realize a computer vision system based on biological vision theory. It is different from those normal models in the field of computer vision.

Key words binocular matching, neural dynamics, FACADE, biological vision

计算机立体视觉是计算机视觉的一个重要组成部分。计算机立体视觉伴随计算机视觉在 Marr 计算机视觉理论主导下发展已近二十年了。Marr 视觉理论中的立体视觉模型将注意力集中在左视图像和右视图像中的要素唯一地进行双目匹配的任务上^[1]。人们经过多年的研究发现,这些问题通常没有唯一解或者稳定解,亦或二者都没有。这当中的原因在于成像过程是多对一的映射,从而导致求逆过程没有唯一解。为了解决这个不适定的问题,人们引入了诸如平滑性约束等各种约束,以减少可能匹配的数目。但这些方法忽略了如下事实:自然景象通常并不一定满足类似于平滑性的条件。因此, Marr 视觉理论体系的局限性促使人们转而研究其它方法,其中,基于生物视觉的某些原理而提出的思路获得人们的广泛注意^{[2][3][4]}。

S. Grossberg 于 1994 年提出并于 1997 年扩展了 FACADE (Form-And-Color-And-Depth) 双目视觉理论框架^[4]。FACADE 是利用生物视觉系统研究的一些成果,在基于细胞动力学的基础上,对生物视觉系统进行了功能模拟。它用完全不同于传统计算机视觉理论的方法,完成了在特定图像上一些视觉功能的实现。但是, FACADE 理论是用于假定的理想观察条件下的规则景物的处理,对在日常场景中大量存在的诸如含噪、边缘模糊、结构非单一等图像的分析 and 处理则显得困难很大。本文所关注的问题是:如何充分利用 FACADE 理论的精髓——细胞动力学方程中所隐含的竞争和合作的作用,来实现非特定观察条件下非理想场景中景物的双目匹配。针对现实图像中观察条件不限定、存在噪声干扰和边界轮廓模糊等特点,有别于 FACADE 模型中双目匹配原理,我们提出了利用细胞动力学模型隐含的竞争和合作来实现的“方位合作,视差竞争”的双目匹配模型,以加强本模型对噪声的抗干扰能力和抗模糊能力,从而拓展了 FACADE 理论的应用范围。

* 1999 年 3 月 31 日收稿

第一作者:彭雄宏,男,1973 年生,博士生

1 基于细胞动力学的双目匹配原理及数学实现

双目匹配是计算机视觉中一个很重要的模块。它由人的双目产生立体感中获得启示,研究的是如何从左右两个摄像机所得的两幅图像中获取场景中物体的深度(距离)信息。基于细胞动力学的双目匹配是针对两个摄像机光轴平行的情况进行考虑和计算的。

基于细胞动力学的双目匹配原理是在充分吸收 FACADE 理论框架中的主要思想,对生物双目视觉的双目匹配过程建立了一个可用的模型,这在计算机视觉中实现后续的双目形状感知等生物视觉功能奠定了良好的基础。

基于细胞动力学的双目匹配系统由三大部分组成,它们分别是^[4]: ON 细胞和 OFF 细胞处理阶段;简单细胞处理阶段;双目合作竞争阶段。这几个阶段的动力学基础都是细胞动力学方程(也称膜电位方程或分流方程)。下面分别讨论这几个部分。

1.1 细胞动力学方程

一般来说,每个神经元细胞的功能实现都可以建模为一个变化的电压 $V(t)$ (也称它为该细胞的活性),它的动力学方程(亦称膜电位方程)如下^[4]:

$$C_m \frac{dV}{dt} = - (V(t) - E_{LEAK}) g_{LEAK} - (V(t) - E_{EXCIT}) g_{EXCIT}(t) - (V(t) - E_{INHIB}) g_{INHIB}(t) \quad (1)$$

这里参数 E_{LEAK} 、 E_{EXCIT} 、 E_{INHIB} 分别代表对应泄漏、激励和抑制的反电势, g_{LEAK} 是定常的泄漏电导,时变的电导 $g_{EXCIT}(t)$ 和 $g_{INHIB}(t)$ 分别是该细胞对应于激励和抑制的总输入的函数。在具体计算中,我们不考虑暂态过程的状态变化,因为我们假定所有状态达到稳态后才对活性进行计算。通过重新标定时间,电容项 C_m 可以设为 1。通过平移 $V(t)$, 泄漏逆电势可以设置为 0。这样变换后,抑制逆电势 E_{INHIB} 是非正的。

于是方程(1)改写为

$$\frac{dV}{dt} = - \alpha V + (U - V) g_{LEAK} - (V + L) g_{INHIB} \quad (2)$$

这里 $\alpha = g_{LEAK}$ 是定常的衰减率, $U = E_{EXCIT}$, $L = -E_{INHIB}$ 。

方程(2)的细胞动力学方程是 FACADE 理论(包括我们的双目匹配方法)的基础, FACADE 理论系统内的几乎所有方程都是以这个细胞方程为基础进行相应拓展的。在方程(2)中,如果引入“邻域”的激励和抑制信号,便能使方程隐含竞争和合作作用。在具体使用中, V 对应着图像的每一个象素位置, U 、 L 分别是相应功能细胞的最大活性值和最小活性值(通常都取 1.0), $g_{EXCIT}(t)$ 和 $g_{INHIB}(t)$ 以一定的策略和连接权值汇聚从前一处理阶段上对应邻域来的信号以形成激励输入和抑制输入。

1.2 ON 细胞和 OFF 细胞处理阶段

ON 细胞和 OFF 细胞处理阶段是由对应于视网膜和侧膝体的 ON 通道和 OFF 通道的中心-周边相互作用组成的。在生物视觉系统中,“中心激励,周边抑制”的 ON 细胞和“中心抑制,周边激励”的 OFF 细胞对照明的影响进行了补偿,它们敏感于每个细胞的感受野所张成的局部图像区域的对比度分布。相应地, FACADE 理论对 ON 细胞和 OFF 细胞处理阶段进行了建模,构造了对应于输入的左眼图像和右眼图像的 ON 和 OFF 通道。通道的响应被半波整流,以产生对有向简单细胞的输入信号。

$$\frac{do_{ij}^{+/-}}{dt} = - \alpha o_{ij}^{+/-} + (U - o_{ij}^{+/-}) g_{p,q}^{+/-}(p, q) I_{i+p, j+q} - (o_{ij}^{+/-} + L) g_{p,q}^{-/+}(p, q) I_{i+p, j+q} \quad (3)$$

上式与细胞膜电位方程相对应,其中: $o_{ij}^{+/-}$ 表示 ON/OFF 细胞的活性; I_{ij} 表示初始的图像输入; $g_{p,q}^{+}$ 和 $g_{p,q}^{-}$ 分别表示对初始图像信号进行信号汇聚所需的激励核和抑制核; U 代表细胞活性的上界, L 代表细胞活性的下界,下标 i, j 表示图像象素点的坐标。方程 3 对左视图像和右视图像均起作用。类似地,所有以下方程,除了右上角特别注明或标记 L, R 外,均同时对左视图像和右视图像操作。

1.3 简单细胞处理阶段

简单细胞处理阶段是由伸长的对比检测器组成的,它完成一定方向上的局部对比度检测,即有向滤波。简单细胞对预定的方位极性或对比度方向产生反应。简单细胞处理阶段输出经过半波整流后,被输

入到复杂细胞级(即双目合作竞争阶段)进行处理。在 FACADE 理论中,用 Gabor 滤波器来实现有向滤波的功能。它由一个正弦项和一个指数项的乘积组成。简单细胞处理阶段依方向的不同和对比极性的不同而形成多个方向上的正负极性的输出。

简单细胞处理阶段的输出标记为 $J_{ijk}^{+/-}$, 它表示位置为 (i, j) , 方向为 k 的正极性/ 负极性的简单细胞的输出。

1.4 双目合作竞争阶段

双目竞争和合作阶段是双目匹配中最关键的一步,它含有双目视觉系统中首次被双目信号激活的细胞: 复杂细胞。生物视觉系统正是通过双目竞争才能从左右两幅图像形成各个深度平面上的响应。

根据 Poggio, Richard 等人的生理学研究可知,包含在左视图像和右视图像中的视差信号被选择性地汇合到几组细胞中。不同的细胞组(我们称之为细胞池)对不同的视差范围进行调谐和编码。有证据表明: 每一个细胞池可以对特定的视差及该视差的邻近视差进行调谐和编码,一系列细胞池可以完成对整个视差范围进行调谐和编码,以达到准确地对场景中存在的连续视差处理的目的。为此,在实现中,我们可以采用有限的细胞池,按由远到近或由近到远的顺序,对感兴趣的范围内的视差在一定误差范围内进行调谐和编码。对应视差的细胞池数目越多,系统对连续视差的调谐和编码就越接近生物视觉^[4]。

首先,来自简单细胞的信号被以下述方式进行组合:

$$A^L = \sum_p C_{dp} J_{i,j+p,k}^L - \sum_p C_{dp} J_{i,j-p,k}^L \quad (4)$$

$$A^R = \sum_p C_{dp} J_{i,j+p,k}^R - \sum_p C_{dp} J_{i,j-p,k}^R \quad (5)$$

$$S_{ijkd} = h(A^L + A^R) \quad (6)$$

其中:

$$h(x) = \begin{cases} x - \Gamma & \text{若 } x > \Gamma \\ 0.0 & \text{其它} \end{cases} \quad (7)$$

在公式(4)~(7)中, C_{dp} 是将视差进行调谐和编码到不同的细胞池上的权函数(这一般是通过学习和经验来获得具体值); S_{ijkd} 是汇合到复杂细胞的输入信号; 下标 d 表示与不同的细胞池对应; $H(\cdot)$ 是具有阈值 Γ 的半波整流函数,从而保证了细胞本身活性在一定阈值之上才能对其它细胞产生有效输入。方程(6)中,对调谐后左右图像信号的汇合提供了阈值控制和非线性变换。

双目竞争和合作阶段细胞活性的动力学方程用下式表达:

$$\frac{db_{ijkd}}{dt} = -\alpha b_{ijkd} + (U_2 - b_{ijkd}) E_{ijkd} - (b_{ijkd} + L_2) I_{ijkd} \quad (8)$$

$$E_{ijkd} = S_{ijkd} + \sum_{p,q} g_{pqk} h(b_{i,j+p,q,k,d}) \quad (9)$$

$$I_{ijkd} = \sum_{p,m} h(b_{i,j+p,km}) F_{pkdm}^{left} + \sum_{p,m} h(b_{i,j+p,km}) F_{pkdm}^{right} \quad (10)$$

将公式(8)与细胞动力学方程比较,我们可以发现: E_{ijkd} 汇聚了对该细胞的所有激励输入,它包括了一个合作核 g_{pqk} 作用下邻域细胞对其的激励作用,通过合作核,我们可以将同一个细胞池上 (i, j) 领域中具有相同方向的活性进行合作以消除噪声; 抑制核 F_{pkdm}^{left} 和 F_{pkdm}^{right} 分别实现对来自于其它视差细胞池抑制的汇聚作用。在方程(9)中,我们利用合作核 g_{pqk} 达到在细胞池内进行“方位合作”的目的,在方程(10)中,我们利用抑制核 $F_{pkdm}^{left/right}$ 达到在多细胞池之间进行的“视差竞争”的目的。

细胞活性稳定后, b_{ijkd} 便表示在深度平面 d 上,方位为 k , 位置在 i, j 处的细胞的活性值。 b_{ijkd} 反映了在本阶段的“合作-竞争”过程中“局部优胜者”的分布情况,它基本上实现了左右图像中重要特征的匹配问题。在视差为 d , 方向对比为 k 的细胞活性平面上, b_{ijkd} 直接反映了具有该视差的细胞活性的全部分布。

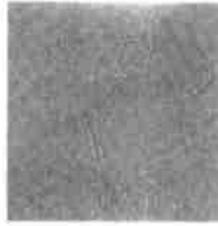
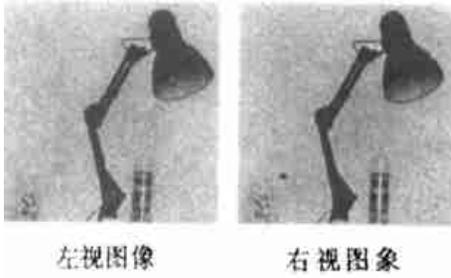


图1 初始立体图像对

Fig. 1 Initial stereo image

图2 ONL 细胞的输出(左图像)

Fig. 2 Output of ON Cell (Left image)

图3 简单细胞的输出(左图像, 竖直方向)

Fig. 3 Output of Simple Cell

(Left image, vertical)

2 基于细胞动力学的双目匹配的仿真

在实验中,我们使用了分别位于距观察点不同深度上的三个物体的立体图像对。如图1中,三个物体为“洗发剂”、“台灯”和“杀虫剂”(下面我们将分别用A、B、C来指代这三个物体)。其中,A距视点最近,B次之,C最远。在仿真中,我们采用了四个复杂细胞池,分别用 $Disp_0$ 、 $Disp_1$ 、 $Disp_2$ 和 $Disp_3$ 来表示,它们表示的顺序为由远到近。其中, $Disp_0$ 代表零视差距离上的细胞池细胞活性的分布(零视差距离表示双眼的凝视点所在的视差平面);在图像环境中, $Disp_0$ 与物体C所在的视差平面相对应, $Disp_1$ 与物体C和物体B之间的视差平面相对应(在此立体图像对中没有物体位于此视差平面上), $Disp_2$ 与物体B所在的视差平面相对应, $Disp_3$ 与物体A所在的视差平面相对应。由于立体图像对中的水平对比信号间的差别并不一定能确定这些水平信号在哪个视差层上,所以,在处理的过程中,FACADE理论认为对应各个视差的细胞池上的水平对比信号响应是一样的,视差主要是由图像中竖直特征在水平方向上移位来体现的,因此我们着重于细胞的竖直对比活性的竞争。

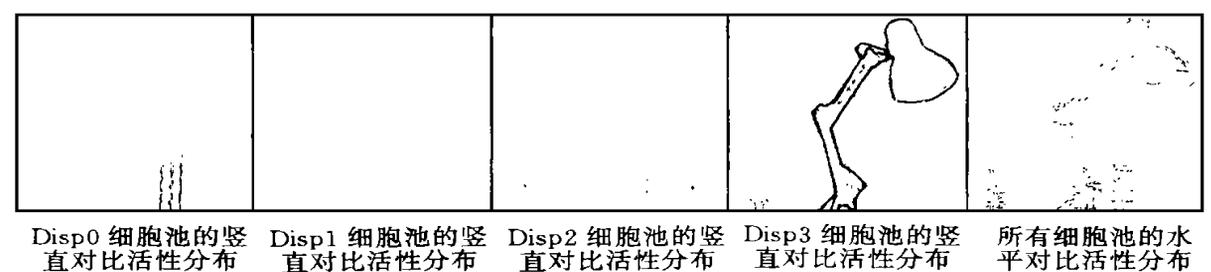


图4 视差反映在各个细胞池上的活性的分布

Fig. 4 Activity distribution of Disparity Cell Pools

由于ON细胞和OFF细胞模拟了生物视觉系统对渐变照明影响的消除,从ON细胞的输出可以看出:物体A的边缘,物体B的弯臂、灯罩和物体C三者之间的阴影,以及物体C的上部区域都被消除了。因为,在原始立体图像中,这三个区域的边缘很模糊,而且其亮度对比小,他们被ON细胞和OFF细胞当成照明渐变的情况而消除了,从而,这些特征无法在复杂细胞池中得到反映。

从图4可以看出,稳定后的各细胞池活性已经正确地反应了在各个视差范围上的竖直到对比信号,这为后续的双目视觉处理过程奠定了良好的基础。另外,图4中的 $Disp_3$ 中可以看出,竖直方向上没有什么断线,这是因为方向的合作减少了由于边缘模糊和噪声的影响而造成断线出现的可能。虽然在各个细胞池上仍有一些“斑点”,它们也是“方向合作、视差竞争”中的优胜者,但它们的获胜存在并不是我们所希望的。它们是误匹配。对于这些误匹配,由于我们已获得基本正确分配在各细胞池上的竖直到对比活性的分布,在后续的双目视觉处理过程中,我们将利用这些活性分布来完成边界优化,边界合作,水平活性的选取等功能,从而达到消除误匹配、恢复物体完整边界的目的。

3 结论

基于细胞动力学的双目匹配方法实现了一种完全不同于 Marr 体系的双目匹配技术。虽然在针对一些复杂的自然场景图像进行处理时,本方法也有困难,但与 Marr 视觉理论指导下的双目匹配方法相比,本方法在生物细胞动力学基础上模拟各级细胞的功能,从而抛开了传统的双目匹配技术中一些必不可少的约束,例如,视差平滑性。从另一个角度来说,我们可以将传统的双目匹配技术中一些约束转化为各阶段的细胞之间竞争和合作所需的激励和抑制,即化“约束”为“引导”,而且这种结构也有利于将来我们把已还原为细胞活性分布的人类的高层智能活动用来作为双目匹配、目标辨识的“引导”,从而为研究计算机立体视觉探索了一条新途径。同时,基于细胞动力学的双目匹配方法本质上是并行的,这在利用并行计算机或神经计算机来实现方面优于传统的特征串行匹配技术。

参考文献

- 1 Marr, 姚国正, 刘磊, 汪云九译, 视觉计算理论, 北京: 科学出版社, 1988
- 2 Ohzawa I, Deangelis G C & Freeman R D. Encoding of Binocular Disparity by Complex Cells in the Cat's Visual Cortex. *The Journal of Neurophysiology*, June 1997 77(6): 2879- 2909
- 3 Grunewald A & Grossberg S. Self-organization of Binocular Disparity Tuning by Reciprocal Corticogeniculate interactions. *Journal of Cognitive Neuroscience*, in press 1997, Technical Report CAS/ CNS-TR-96-027 Boston, MA: Boston University
- 4 Grossberg S & McLoughlin N P. Cortical Dynamics of Three-Dimensional Surface Perception: Binocular and Half-Occluded Scenic Images. *Neural Networks*, 1997, 10(9): 1583 ~ 1605