

· 生物电子学 ·

## 仿视皮层机制的随机点视频序列运动特征提取

许悦雷, 吕超, 马时平, 李帅, 邹洪中, 张文达, 辛鹏

(空军工程大学航空航天工程学院 西安 710038)

**【摘要】**为了探索视皮层对多方向刺激的运动信息处理机制,依据生物实验中部分中颞叶区(MT)细胞对随机点视频刺激运动感知的侧偏特性,提出了一种仿视皮层机制的随机点视频序列运动特征提取模型。首先,采用Von Mises函数拟合初级视皮层(V1)简单细胞的感受野,再结合非线性能量模型、调谐和非调谐正则化方法等模拟实现V1复杂细胞对运动信息的方向感知;其次,提出了级联前馈方法对V1复杂细胞响应进行线性加权求和,得到MT细胞的响应输出;最后,利用生物实测MT细胞响应数据对新模型的有效性进行了实验验证。仿真结果表明模型的输出结果与恒河猴的运动感知实验数据基本吻合,能够较好地模拟视皮层对不同夹角的多方向随机点视频序列的运动感知机制,为提取复杂运动序列的特征提供新的方法,进而为视觉类脑计算奠定基础。

**关键词** 中颞叶区; 运动特征; 运动感知; 侧偏; 视皮层

**中图分类号** TP391, R338 **文献标志码** A **doi**:10.3969/j.issn.1001-0548.2017.04.025

## Motion Feature Extraction of Random-Dot Video Sequences with the Visual Cortex Mechanism

XU Yue-lei, LÜ Chao, MA Shi-ping, LI Shuai, ZOU Hong-zhong, ZHANG Wen-da, and XIN Peng

(Aeronautics and Astronautics Engineering College, Air Force Engineering University Xi'an 710038)

**Abstract** According to side-bias characteristics of motion perception towards random-dot video stimuli among some middle temporal cortex (MT) cells in the experiment, a model for motion feature extraction of random-dot video sequences with the visual cortex mechanism is proposed for exploring the visual cortex mechanism underlying multi-direction motion information processing. First, Von Mises function is used to model receptive fields of simple cells in primary visual cortex (V1). A nonlinear energy model untuned and tuned normalization operation is adopted to simulate the direction perception of V1 complex cells; then, with the feedforward and cascade method, the responses of V1 model are pooled through a set of linear weights, thus giving rise to the responses of MT cells. The simulation results are generally consistent with the experimental data of rhesus monkeys. It can model the motion perception and feature extraction mechanism underlying multi-direction random-dot video sequences at different direction separations in visual cortex. So that it can provide new ideas for extracting features of complex motion sequences and lay the foundation for the visual brain-like computation for the further step.

**Key words** middle temporal cortex (MT); motion feature; motion perception; side-bias; visual cortex

随着欧美等国“脑计划”的实施,类脑计算将成为模式识别和人工智能的优先发展方向之一,其发展可为人工智能等提供新的思路<sup>[1]</sup>。生物视觉具有计算机视觉不可比拟的优势,其更高效、更准确,且需要的先验知识很少,使得生物视觉的研究成为了一大热门。视觉类脑计算即仿生物视觉信息处理机制的运算,其目的是将生物视觉的巨大优势引入计算机视觉领域,发展由生物启发式的计算机视觉,实现对视觉目标信息的高效处理<sup>[2]</sup>。对自然图像、

复杂运动序列等视觉目标的类脑感知,构建可提取出复杂随机点视频序列运动特征的仿生视觉模型,将是计算机视觉要解决的基本问题之一。人类在感知和认识世界的过程中,有80%以上的信息是通过视觉系统获得<sup>[3]</sup>,视觉系统是绝大多数生物最重要的感觉系统。通过视觉系统,可以获取目标的亮度、色彩、运动方式、形状等物理信息。研究表明,初级视皮层(primary visual cortex, V1)和中颞叶区(middle temporal cortex, MT)是生物视觉系统处理运

收稿日期: 2016-11-16; 修回日期: 2017-02-24

基金项目: 国家自然科学基金(61372167, 61379104)

作者简介: 许悦雷(1975-),男,博士,教授,主要从事脑认知与智能信息处理方面的研究。

动信息最主要的区域, V1作为大脑视皮层中最先接收到视觉信息的区域, 对纹理和边缘等信息较为敏感, 并能初步感知运动信息, 通过MT细胞的进一步整合, 则可获取速度和方向等运动信息。

目前, 研究人员基于视皮层机制提出了多种运动特征提取模型, 典型的有以下3种。1) HS模型(Heeger and Simoncelli model)<sup>[4]</sup>, 该模型V1和MT阶段均对各自输入进行线性加权、矫正和正则化操作。运用不同时空方向的三维滤波器组来模拟具有方向和空间频率选择性的V1简单细胞, 再对V1简单细胞的响应加权求和以获取V1复杂细胞的响应, 最后通过对V1复杂细胞的输出加权求和来模拟MT细胞的响应; 2) RMM模型(recurrent motion model)<sup>[5]</sup>, 一种基于递归神经网络的运动检测模型。它依据实验中MT细胞对于偏好方向(preferred direction, PD)和反PD方向响应的实验数据, 采用Elman模型, 有效模拟了MT神经元对于PD和反PD这一简单方向随机点的运动特征提取能力; 3) FFV1MT模型(feedforward V1-MT model)<sup>[6]</sup>, 该模型通过三维Gabor时空滤波器模拟简单细胞的感受野, 再由V1细胞响应加权组合及正则化模拟得到MT细胞模型, 并将其成功地应用于光流估计和运动估计<sup>[7]</sup>。以上3种模型在一定条件下较好地模拟了视皮层V1和MT区细胞对运动信息的处理机制, 可以提取光栅、栅格等连续序列和简单随机点的运动信息, 但由于多方向复杂随机点视频序列运动的不连续性和高频特点, 以上模型很难模拟视觉皮层对它们的区分性特征提取的能力。

鉴于此, 本文根据生物实验数据, 深入探索视皮层V1和MT的感受野特性, 构建仿视皮层机制的运动感知模型, 并将该模型的仿真结果与恒河猴(rhesus monkeys)的生物实验数据进行比较分析, 以论证模型的合理性。本文以一定方向夹角(direction separation, DS)运动的双方向随机点视频序列为输入刺激, 对比生物实验中的MT呈侧偏性(side-biased)的细胞响应, 在剖析视皮层运动信息处理机理的基础上, 建立视觉类脑运动感知模型, 实现了对复杂多方向随机点视频序列的运动特征提取。另外, 本文对生物视觉系统信息处理机理进行的初步探索, 可为类脑计算模型的构建奠定基础, 具有一定的开创性。

## 1 生物视觉感知系统

灵长类动物的视觉系统是一个多层次的复杂神经网络结构, 主要由视网膜、侧膝体(lateral geniculate nucleus, LGN)和视皮层3大部分组成<sup>[8]</sup>。视皮层包含

V1和纹外视皮层, 纹外视皮层中较重要的区域有V2、V3、V4、MT, V2接受来自V1的前馈输入, 并投射到V3、V4、MT, 同时对V1有反馈输入。在人类视觉系统中存在两条通路, 均以视网膜为起点, 并经过LGN到达V1, 如图1所示。其中一条为背侧通路(dorsal stream), 也称Where通路或motion通路, 沿着枕顶叶分布, 由V1、V2、V3经MT、内上颞叶(medial superior temporal cortex, MST)到达顶叶, 主要处理空间和运动信息, 实现视觉系统对空间信息和动作信息的具体感知<sup>[9]</sup>; 另一条为腹侧通路(ventral stream), 也称What通路或form通路, 沿着大脑皮质的枕颞叶分布, 从枕叶的V1、V2经V4投射到下颞叶(inferior temporal cortex, IT), 主要处理颜色、形状和纹理等信息, 实现对物体的识别与感知<sup>[10]</sup>。当前, 关于运动感知和特征提取的仿生视觉模型多基于背侧通路, 且绝大部分都是围绕V1和MT阶段展开。

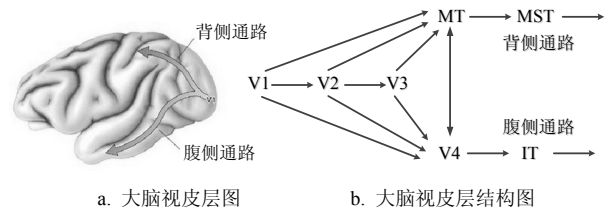


图1 视皮层两条通路

V1是生物视觉信息处理的初始区域, 其视野集中于局部区域, 一般分为简单细胞和复杂细胞两类。简单细胞的感受野呈狭长形, 对空间频率及朝向敏感, 复杂细胞以简单细胞输出的叠加为输入, 其感受野由简单细胞感受野叠加而成, 对特定方向的条状刺激敏感。MT具有全局感知的能力, 可整合经多V1细胞处理的局部运动信息, 其对运动的速率和方向较为敏感, 感受野约为V1的100倍左右<sup>[11-12]</sup>。图2为V1和MT的感受野层级关系。

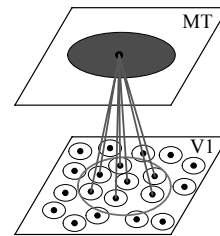


图2 V1和MT感受野层级关系

## 2 生物实验及运动感知模型

### 2.1 生物实验描述

本文实验是在美国威斯康辛州大学麦迪逊分校的神经科学系, 与他人合作完成, 相关信息与文献

[13]所述相同。生物学研究表明,恒河猴大脑视皮层区与人脑视皮层区相似度达90%以上<sup>[14]</sup>。实验对象为图3a所示的两只成年健康的恒河猴。

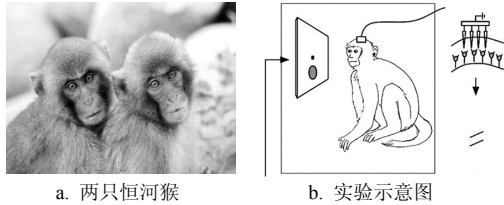


图3 实验对象

实验中所用刺激为运动的无色差随机点视频序列,在静止的圆形孔径内呈现,并通过显示屏向恒河猴视网膜投影。视频序列中同一方向上的所有点均以相同的速率移动,具有运动一致性。基于控制变量的思维,通过多次给两只恒河猴视网膜投射单一方向或不同DS的随机点视频序列,实时采集响应。DS采取45°、60°、90°、135°等典型度数,具体随机点视频序列刺激如图4所示。本文将钨电极直接连接视觉皮层MT细胞中的神经元,如图3b所示,实时采集响应数据。本实验以前200 ms为预备时间,实验所处环境、所用设备的型号参数、随机点亮度和速度以及其中所需的医学准备、生物电子学记录等与文献[13, 15]描述相同。当DS=90°时,实测相应数据如图5所示,图中右侧为响应值与颜色深度的对应图。

## 2.2 运动感知模型

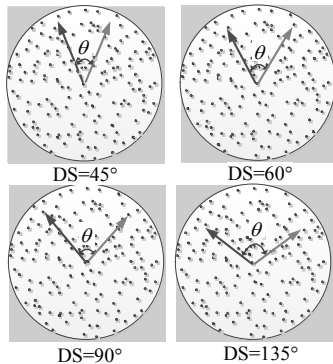


图4 随机点视频刺激

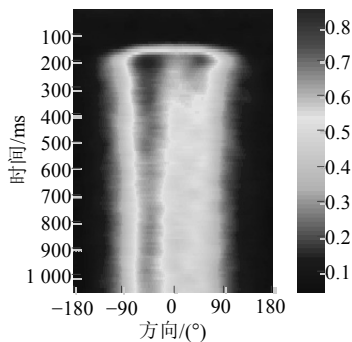


图5 实验中MT响应(DS=90°)

针对V1简单细胞的感受野特性,本文采用Von Mises函数描述其方向调谐曲线。Von Mises函数与圆高斯函数类似,它是一种描述方向数据的钟形模型,在生物神经学、天文学、地理学等领域有着重要的作用<sup>[16]</sup>。V1细胞的PD方向与接收的刺激方向形成一个类似钟形的感受野模型,当刺激方向与PD方向一致时,调谐曲线达到最大值。这与Von Mises函数的吻合度较高。首先定义参数调谐带宽控制系数 $a$ ,其表达式为:

$$a = \frac{1}{TW^2} \quad (1)$$

式中,  $TW$  为方向调谐带宽,即细胞的最大响应与最小响应之差的一半所对应的宽度<sup>[17]</sup>。当  $TW$  越大时,  $a$  越小。V1细胞方向调谐曲线为:

$$R_n(t_i) = e^{\alpha[\cos(D_i - S_i) - 1]} \quad (2)$$

式中,  $S_i$  为随机点刺激的运动方向;  $D_i$  为V1细胞的PD方向。图6为V1细胞方向调谐曲线简图。

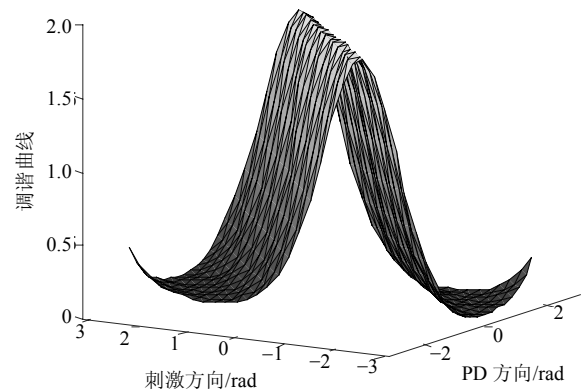


图6 V1简单细胞调谐曲线示意图

从图6可以看出,由对于细胞而言,当刺激运动方向  $S_i$  与PD方向  $D_i$  一致时,调谐曲线达到最大值,此时细胞呈兴奋状态;当刺激运动方向与PD方向相反,即相差180°时,调谐曲线达到最小值,此时细胞呈抑制状态。而这与生理学研究结果保持一致<sup>[9]</sup>。

V1简单细胞的响应由刺激  $T(S_i, t_i)$  和方向调谐曲线  $R_n(t_i)$  的内积计算得到,即:

$$L_n(t_i) = \sum_m R_m(t_i) T(S_i, t_i) \quad (3)$$

结合非线性的能量模型<sup>[18]</sup>,进行非调谐正则化操作,得到:

$$P_n(t_i) = \frac{L_n^2(t_i)}{\sum_k L_k^2(t_i) + \sigma_1^2} \quad (4)$$

由于非调谐正则化对V1细胞方向感知能力表征不足,故采取调谐正则化,即结合静态非线性的自正则化(self-normalization)操作。得到V1复杂细胞

的响应为:

$$V_n(t_i) = \frac{P_n(t_i)}{P_n(t_i) + \sigma_2} \quad (5)$$

式中,  $\sigma_1$  和  $\sigma_2$  为极小的正值常数, 以避免“白墙”问题, 导致分母为零的情况。

在MT阶段, 由图2所示的感受野层级特性, V1-MT之间的连接采用前馈级联方法, 即:

$$M(t_i) = \sum_k w_k V_k(t_j) \quad (6)$$

得到MT细胞的响应。为准确获取式(6)中的权值系数  $w_k$ , 采用岭回归(ridge regression)方法, 根据V1复杂细胞的响应及实验中恒河猴的MT细胞响应, 定义该系数的误差:

$$E(\mathbf{w}) = \|\mathbf{V}_w - \mathbf{M}_r\|^2 + \lambda \|\mathbf{w}\|^2 \quad (7)$$

式中,  $\mathbf{w}$  是权值系数  $w_k$  的向量形式;  $\mathbf{V}$  是由V1的输出响应数据构成的矩阵, 矩阵中的各行代表V1的输出响应;  $\mathbf{M}_r$  为实验中记录的恒河猴MT输出响应向量;  $\lambda$  为岭参数, 是一个极小的正值常数。

为求式(7)的最小值, 对其求导, 并令导数为零, 得:

$$\mathbf{w} = (\mathbf{V}^T \mathbf{V} + \lambda \mathbf{I})^{-1} \mathbf{V}^T \mathbf{M}_r \quad (8)$$

式中,  $\mathbf{I}$  为单位矩阵。由  $\mathbf{w}$  获得最佳权值系数  $w_k$ , 将  $w_k$  代入式(6)即可求出本文模型的MT输出响应。

### 3 实验及模型结果分析

#### 3.1 V1阶段

本文所提的视皮层模型中, V1阶段采用Von Mises函数表征V1响应的调谐曲线, 并进行调谐和非调谐正则化处理, 得到V1复杂细胞的响应, 其响应曲线如图7所示。由图7可以看出, 整体而言, V1复杂细胞在随机点运动的方向响应最大, 偏离运动随机点方向响应逐渐变小, 这与生物实际是相符的<sup>[9]</sup>。当DS为45°和60°时, 模型对两随机点的区分性较小, 此时V1响应呈单峰特性; 随着DS的逐渐增大, 模型对两随机点的区分能力渐渐增强, 当DS为90°和135°时, V1响应呈双峰特性。模型V1阶段基本实现了V1细胞对方向的敏感特性, 可较好区分出DS较大的不同随机点视频序列, 是整个模型的基础。

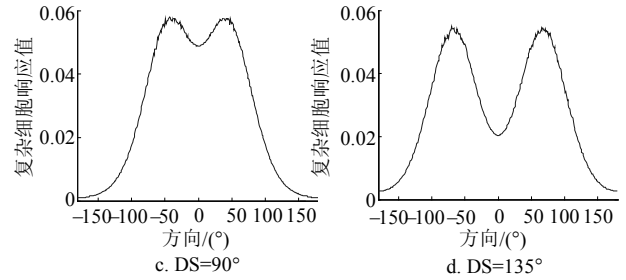
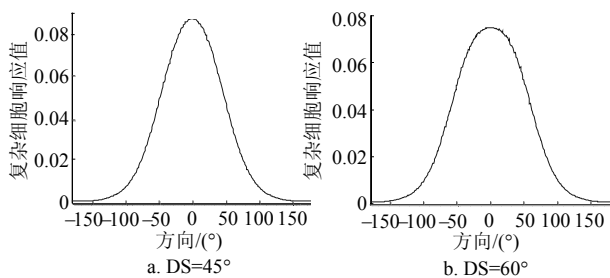


图7 V1复杂细胞响应

#### 3.2 MT阶段

##### 3.2.1 对比分析

针对生物实验中不同MT细胞响应呈现的特点, 文献[13]将采集的响应分为均衡型, 双峰型及侧偏型, 比例分别为42%、19%、39%, 对应3种不同方向选择性的MT细胞。本文主要研究实验结果中的侧偏型细胞。图8a为本文各响应的对比, 从图8a第一列可以看出, 在165 ms后响应逐渐较强, 而随着200 ms时刻的随机点视频序列产生, 实验中恒河猴的MT细胞响应达到峰值, 之后响应整体呈由强到弱的变化规律。整体看来, 侧偏特性随着DS的逐渐增大而增强, 当DS为45°和60°时, 起初呈对称的单峰特性, 随时间推移, 分别在500 ms和400 ms后逐渐略呈侧偏特性; 当DS为90°和135°时, 在分别保持100 ms和50 ms左右的双峰特性后, 响应渐渐偏向一侧, 偏向性较强。实验中各DS响应虽然在侧偏的时间起点及侧偏程度上存在差异, 但均偏向同一侧, 且DS越大, 响应的侧偏特性呈现得越快, 侧偏程度越明显; 而从图8a第二列可看出, 将两单向随机点刺激分开实验并得到的平均响应呈明显的对称性, 无侧偏性, 且强度明显小于双方向随机点刺激下的响应; 从图8a第三列看, 在预备时间后模型在各DS输出的响应, 随时间而逐渐变弱, 且均呈现出侧偏特性, 在DS为90°和135°时尤为明显。当DS为45°和60°时, 前400 ms左右略呈现双峰特性。对比图8a, 本文模型的输出响应与生物实际较为吻合, 较两单向随机点平均响应而言, 不同DS下均体现出了侧偏特性。实验数据表明侧偏型MT细胞对于多刺激的响应与各分刺激响应的与明显区别。

为直观理解, 取800 ms时刻的生物实际响应与本文输出, 不同DS下响应曲线如图8b所示。左侧为生物实验中的响应, 右侧为本文模型的输出响应。从图中可看出, 各DS情况下, 响应曲线均与生物数据相似, 本文模型达到较好的模拟效果。

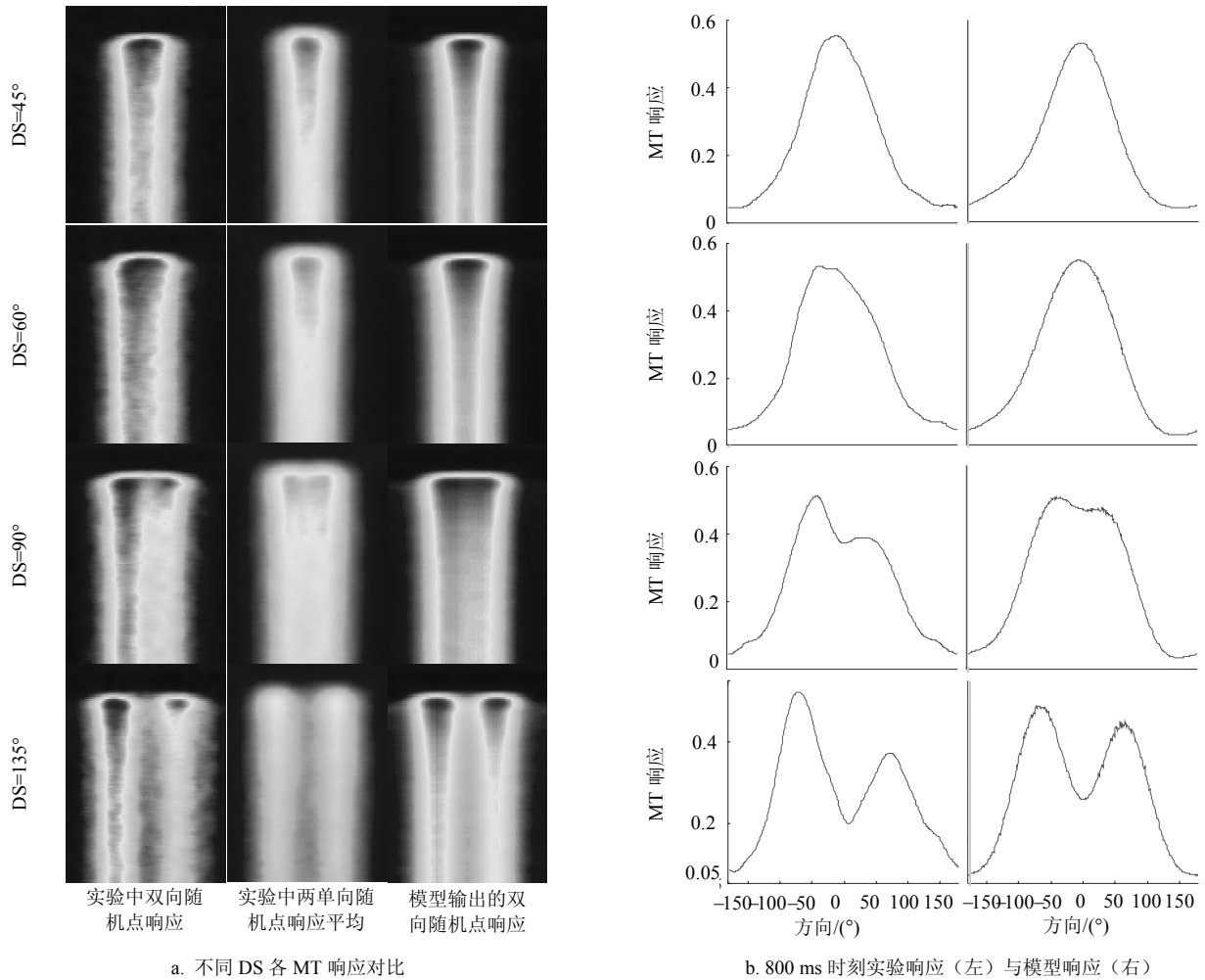


图8 各情况下MT细胞响应

### 3.2.2 误差评估

为了实现对模型有效性的量化评估, 本文以均方差来衡量模型的输出响应与生物实验数据的差异。以图8b中800 ms时刻生物实验数据与模型输出为例, 将两条曲线对应的值相减取均值, 求出此刻的平均绝对误差。以此类推, 求出所有时刻的平均绝对误差, 通过运算可获得均方差, 如图9所示。

从图9可以看出, 在前165 ms, 模型的均方差较小, 在0.01以内; 在165~210 ms期间, 均方差增加, 直至达到峰值0.054, 这与视皮层神经细胞在感受突发刺激前后暂态的复杂的随机不稳定性有关; 210 ms以后均方差逐渐减小, 且在1 000 ms时已小于0.033, 并继续呈减小趋势。结合图8和图9, 可得本文的模型具有鲁棒性, 且误差持续性减小, 能较好地吻合生物实验数据, 反映了视皮层功能特性, 描述了生物行为, 可对目标的运动特征进行持续性提取。

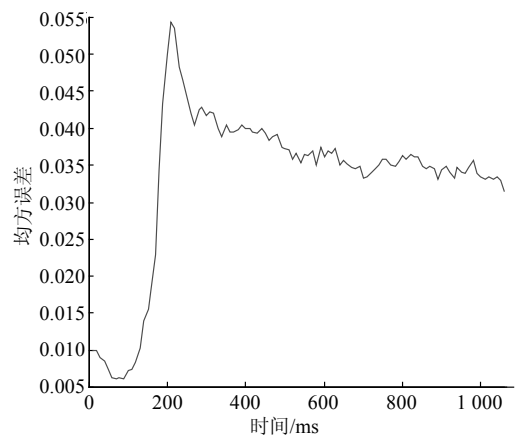


图9 本文模型的均方差

## 4 结束语

为准确地模拟出部分MT细胞对不同夹角运动随机点视频序列的响应侧偏特征, 本文深入探索生物视觉系统分层分区处理运动信息的机制, 提出了一种仿视皮层机理的多方向随机点视频序列运动模

型, 该模型基于生物实验数据, 可有效提取视皮层中部分MT细胞对多方向随机点的运动感知特性, 具有一定的开创性。

然而, 当夹角较大时, 本文模型输出响应的侧偏程度较生物实验数据略显不足, 这与人脑视皮层复杂的神经结构有关, 需要对大脑信息处理机制进行进一步探索, 特别是视皮层对复杂视频序列的运动特征提取机理, 以建立更加完善的仿生视觉运动感知模型。

### 参 考 文 献

- [1] POO M, DU J, IP N Y, et al. China brain project: basic neuroscience, brain diseases, and brain-inspired computing[J]. *Neuron*, 2016, 92(3): 591-596.
- [2] DANUSER G. Computer vision in cell biology[J]. *Cell*, 2011, 147(5): 973-978.
- [3] 寿天德. 视觉的神经基础[J]. *自然杂志*, 2015, 37(1): 17-25.  
SHOU Tian-de. Neuronal basis of vision[J]. *Chinese Journal of Nature*, 2015, 37(1): 17-25.
- [4] SIMONCELLI E P, HEEGER D J. A model of neuronal responses in visual area MT[J]. *Vision research*, 1998, 38(5): 743-761.
- [5] JOUKES J, HARTMANN T S, KREKELBERG B, et al. Motion detection based on recurrent network dynamics[J]. *Frontiers in Systems Neuroscience*, 2014, 8: 239-239.
- [6] SOLARI F, CHESSA M, MEDATHATI N V K, et al. What can we expect from a V1-MT feedforward architecture for optical flow estimation?[J]. *Signal Processing: Image Communication*, 2015, 39: 342-354.
- [7] CHESSA M, SABATINI S P, SOLARI F, et al. A systematic analysis of a V1-MT neural model for motion estimation[J]. *Neurocomputing*, 2016, 173: 1811-1823.
- [8] AN X, GONG H, MCLOUGHLIN N, et al. The mechanism for processing random-dot motion at various speeds in early visual cortices[J]. *PLOS ONE*, 2014, 9(3): 1-22.
- [9] 邹洪中, 许悦雷, 马时平, 等. 基于视皮层 V1 模型的随机点视频序列运动特征提取[J]. *计算机应用*, 2016, 36(6): 1677-1681.
- ZOU Hong-zhong, XU Yue-lei, MA Shi-ping, et al. Motion feature extraction of random-dot video sequences based on V1 model of visual cortex[J]. *Journal of Computer Applications*, 2016, 36(6): 1677-1681.
- [10] MAHMOODI S, SABA N. Nonlinear model for complex neurons in biological visual visions[C]//International Conference on Bio-Inspired Systems and Signal Processing. Rome: Biomedical engineering systems and technologies. [S.l.]: [s.n.], 2016: 162-167.
- [11] LIU L, PACK C. Training alters the causal contribution of area MT to visual motion perception[J]. *Journal of Vision*, 2016, 16(12): 1133.
- [12] LUI L L, ROSA M G. Structure and function of the middle temporal visual area (MT) in the marmoset: Comparisons with the macaque monkey[J]. *Neuroscience Research*, 2015, 93: 62-71.
- [13] XIAO J, HUANG X. Distributed and dynamic neural encoding of multiple motion directions of transparently moving stimuli in cortical area MT[J]. *The Journal of Neuroscience*, 2015, 35(49): 16180-16198.
- [14] JONES E G, POWELL T P S. Connexions of the somatic sensory cortex of the rhesus monkey[J]. *Brain*, 2016, 93(1): 37-56.
- [15] XIAO J, NIU Y, WIESNER S, et al. Normalization of neuronal responses in cortical area MT across signal strengths and motion directions[J]. *Journal of Neurophysiology*, 2014, 112(6): 1291-1306.
- [16] JENISON R L, FISSELL K. A comparison of the von Mises and Gaussian basis functions for approximating spherical acoustic scatter[J]. *IEEE Transactions on Neural Networks*, 1995, 6(5): 1284-1287.
- [17] NAVALPAKKAM V, ITTI L. Attentional modulation of tuning width, preferred features and gains during visual search[J]. *Journal of Vision*, 2007, 7(9): 948-948.
- [18] HARTUNG L, KLIMOVSKY A. The glassy phase of the complex branching Brownian motion energy model[J]. *Electronic Communications in Probability*, 2015, 20(0): 1-15.

编辑 税红