

⑦

基于视觉生理特性的视差求取算法

李峰 周源华

(上海交通大学图像通信与信息处理研究所, 上海, 200030)

R318.18

摘要 在 Freeman 等人关于眼睛的视觉生理特性的研究基础上, 进一步从数学上推导出了“复合型”双目视觉细胞的模型, 并且将其用于视差求取, 实现了一种基于视觉生理特性的视差求取算法. 对随机点图和五边形像对的实验表明本算法能够快速高效地计算视差.

关键词 简单型双目细胞, 复合型双目细胞, 视差.

视觉生理

引言

视差是从立体像对中恢复三维景物的基础, 这正是通过研究动物(包括人)的视觉细胞的立体成像原理而得出的结论. Poggio 等人发现^[1]: 左、右眼中一部分对应的视觉细胞所组成的细胞对(称为双目细胞)具有视差敏感性, 即双目细胞的视网膜调谐于某个误差 D_0 , 当一对输入信号之间的视差 D 等于其谐视差 D_0 时即达到“谐振”, 双目细胞的输出最大, 而 $D \neq D_0$ 时, 输出减小, 从而揭示了双眼立体成像的基本原理. 之后, Freeman 等人通过实验又发现双目细胞可以分为简单型和复合型, 前者的输出特性可以用 2 个 Gabor 函数(分别对应于左右细胞视网膜的接受区域)来描述, 其输出与视差和信号位置有关; 后者的输出则可以通过综合 2 个对偶的简单型的双目细胞的输出而得到, 且其输出仅受视差的影响^[2-4].

但是, Freeman 等人仅仅是通过实验得出了上述结论, 并未给出严格的数学推导和证明. 本文则从数学上进行推导, 得出了描述复合型双目细胞的数学模型, 使通过模拟复合型双目细胞来求取视差成为可能.

1 复合型双目细胞模型和视差求取

Freeman 和 Ohzawa 等人的实验表明一对典型的“简单型”双目细胞可以用两个 Gabor 函数来刻画^[2], 即

$$f_1(x) = \exp(-x^2/(2\sigma^2))\cos(\omega x + \varphi_1), \quad (1)$$

$$f_2(x) = \exp(-x^2/(2\sigma^2))\cos(\omega x + \varphi_2). \quad (2)$$

其中, σ 和 ω 决定了输出响应的范围和工作频率, φ_1, φ_2 是初始相位. 如果输出响应范围 σ 远大于视差 D (一般地有 $3\sigma \gg D$), 则 Gabor 函数中 Gaussian 函数的窗口效应可以忽略.

设 $I_1(x)$ 和 $I_2(x)$ 分别是“简单型”双目细胞的左右视网膜接收到的信号, 根据 Freeman 等人的实验结果^[2,4]简单型双目细胞的输出可以写成

$$r_s = \int [f_l(x)I_l(x) + f_r(x)I_r(x)]dx. \quad (3)$$

如果 $I_l(x), I_r(x)$ 之间具有视差 D , 则

$$I_r(x) = I_l(x - D). \quad (4)$$

当 Gabor 函数的窗口宽度较大时 ($6\sigma \gg D$), Gabor 函数可看作一个窄带滤波器, 则式(3)的卷积相当于用窄滤波器对 $I_l(x)$ 和 $I_r(x)$ 进行滤波, 其输出近似为频率 ω 单频正弦波

$$r_s \approx \rho [\cos(\theta + \varphi_l) + \cos(\theta + \varphi_r - \omega D)], \quad (5)$$

或等价地有

$$r_s \approx 2\rho \cos(\theta + (\varphi_l + \varphi_r)/2 - \omega D/2) \cos((\varphi_l - \varphi_r)/2 + \omega D/2). \quad (6)$$

其中, ρ 和 $\theta = \omega t$ 是信号在工作频率 ω 处的幅度和相位。

式(5)、(6)表明简单型的双目细胞的输出不但与差视 D 有关, 且受到输入信号相位的影响, 图1表明了这种情况。构造一个“简单型”的双目细胞模型, 其 $\sigma = 4$ pixels, $\omega/2\pi = 0.125r/\text{pixel}$, $\varphi_l = \varphi_r = 0$, $I_l(x) = \delta(x)$, $I_r(x) = \delta(x + D)$, 图中, 纵轴 R 是输出, 横轴 d 是视差 (disparity)。当 $I_l(x)$ 位于左眼视网膜中央时 ($\theta = 0$), $I_r(x)$ 以右眼视网膜中央为基准左右移动, 双目细胞的输出为图1中的实线; 如果 $I_l(x)$ 不在左眼视网膜的中央, 而向左或向右偏移2个单位 ($\theta < 0$ 或 $\theta > 0$), 则输出如2条虚线所示。图1的结果与式(5)、(6)一致。所以, 简单型的双目细胞不能用于视差估算。根据 Freeman 等人的实验结果, 综合2个简单型的双目细胞模型而得到复合型的双目细胞模型, 就能消除输入信号相位对输出的影响。但是 Freeman 等人并未从数学上证明如何由2个简单型双目细胞综合成复合型双目细胞模型。

显然, 从式(6)可以看出, 综合2个简单型的双目细胞合成1个复合型双目细胞, 其目的就是消去式(6)中含 θ 项。下面就由式(6)所表示的简单型双目细胞来构造复合型双目细胞的数学模型。

设有2个简单型双目细胞

$$r_{is} \approx 2\rho \cos[\theta + (\varphi_{i1} + \varphi_{i2})/2 - \omega D/2] \cos[(\varphi_{i1} - \varphi_{i2})/2 + \omega D/2], \quad (i = 1, 2) \quad (7)$$

若令 $(\varphi_{11} + \varphi_{12})/2 = (\varphi_{21} + \varphi_{22})/2 = \Delta\varphi/2$, $|(\varphi_{11} + \varphi_{12})/2 - (\varphi_{21} + \varphi_{22})/2| = \pi/2$, 则称这2个双目细胞为对偶双目细胞。则定义复合型双目细胞模型为

$$r_c = r_{1s}^2 + r_{2s}^2 = 4\rho^2 \cos^2(\Delta\varphi/2 + \omega D/2). \quad (8)$$

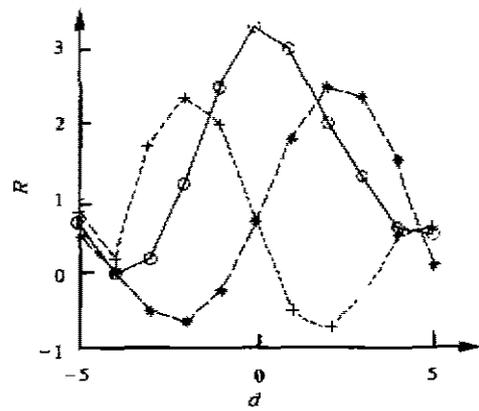


图1 “简单型”双目细胞的输出特性
Fig. 1 The output character of simple binocular cell

式(8)就描述了复合型双目细胞的特性. 显然在 $\Delta\varphi$ 和 ω 选定之后, 它只是视差 D 的函数, 其调谐视差为

$$D_{\text{opt}} = -\Delta\varphi/\omega. \quad (9)$$

调谐宽度(输出特性曲线下降到一半时的宽度)为 $\Delta D = \pi/\omega$, 因为 r_i 是周期为 π 的周期函数, 为避免卷绕问题, 要求 $|\omega D/2| \leq \pi$ 即 $|D| \leq 2\pi/\omega$.

2 基于复合型双目细胞模型的视差求取算法

首先, 构造一族(N 个)复合型双目细胞模型为

$$r_k = 4\rho^2 \cos^2(\Delta\varphi_j/2 + \omega D/2), \quad (j = 1, 2, \dots, N) \quad (10)$$

其中 $\Delta\varphi_j$ 均匀分布于 $[-\pi, \pi]$. 这样的一族细胞即有 N 个不同的调谐视差, 即

$$D_j = -\Delta\varphi_j/\omega \quad (11)$$

与 N 个复合型细胞相对应有 $2N$ 个简单型细胞 r_k ($k=1, 2, \dots, 2N$), 设输出信号为 $I_l(x)$ 和 $I_r(x)$ (本算法只适用于一维视差求取), 计算 r_k 在位置 x 处的输出为

$$r_k(x) = \int [f_{s_l}(x)I_l(x) + f_{s_r}(x)I_r(x)] dx. \quad (12)$$

然后利用式(8)将 N 对对偶双目细胞的输出组合成族中 N 个复合型细胞的输出 $r_k(x)$, 最后计算 $\max(r_k)$, 其中 $j=1, 2, \dots, N$, 找出具有最大响应的细胞

$$r_{m_c}(x) = 4\rho^2 \cos^2(\Delta\varphi_m/2 + \omega D(x)/2), \quad (13)$$

则 $D(x) = -\Delta\varphi_m/\omega$ 即为输入信号在位置 x 处的视差.

如果 x 处实际的视差值 D_{real} 落在 N 个调谐视差 $D_j(x)$ 当中且不等于任何一个 $D_j(x)$ 时, 可以由 $r_k(x)$ 等来拟合一条曲线, 求出其极大值处所对应的视差值, 这一改进可以进一步提高视差的精度.

在实际应用中, 结合从粗至精的多尺度匹配策略, 可以大大扩展视差的估算范围, 通过 FFT 来计算卷积则能加快算法的速度.

3 实验结果

分别以随机点图和 pentagon 像对对算法进行了验证, 如图 2~7 所示. 随机点图大小为 512×512 pixels, 点密度为 50%, 点大小为 1 pixel. 只包括水平视差, 边缘的视差为 0 pixel, 中心的视差为 11 pixels, 图 3 是视差的位图形式, 图 4 是视差所反映的蕴藏于点图中的景物的形状. pentagon 像对的大小也为 512×512 pixels, 图 6 给出了在灰度形式显示的视差. 其视差值从 $-9 \sim +11$ pixels 变化. 图 7 是利用当前较为流行的相位匹配^[3]的方法算出的视差, 可见本算法的效果要好于相位匹配.

算法中采用的参数为: $N=8$, $\sigma=12$ pixels, $\omega=\pi/2$ (rad/s). 以速度较慢的 MATLAB 5.0 为平台, 得到图 6 的结果需要 260s 左右, 其 C 语言可执行程序版本可望在几十秒内完成.

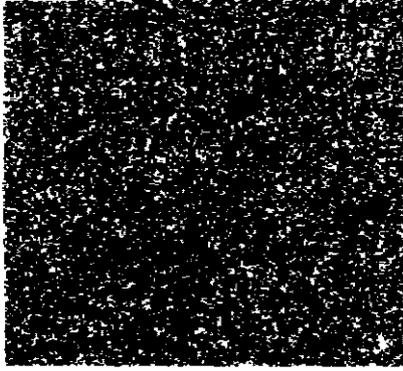


图 2 随机点图
Fig. 2 Random dots diagram

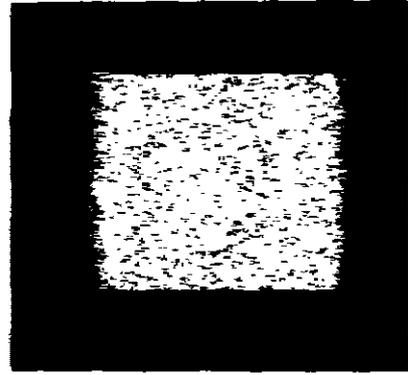


图 3 以灰度显示的视差
Fig. 3 Disparities shown in the form of intensity

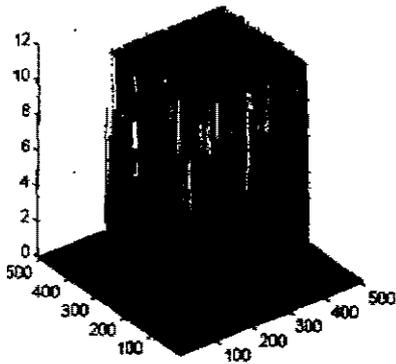


图 4 视差所反映的三维场景图
Fig. 4 Three-dimensional scene reflected by the disparities

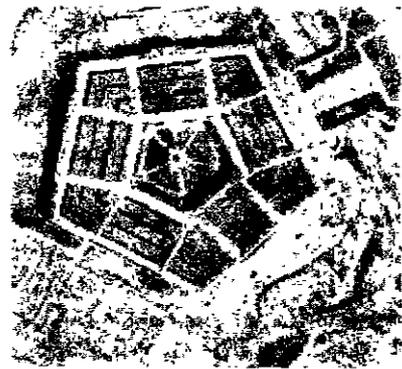


图 5 五边形立体像对
Fig. 5 Pentagon stereopair



图 6 以灰度显示的视差
Fig. 6 Disparities shown in the form of intensity



图 7 用文献[6]的方法得到的视差
Fig. 7 Disparities obtained by the method of Ref. [6]

4 结语

本文推导出了复合型双目细胞的数学模型并成功地将其用于视差估算, 此模型严格描述了眼睛的立体视觉特性, 表明了复合型双目细胞的输出和视差之间的解析关系. 显而易见, 随着视觉生理和心理学的不断深入, 我们应尽可能地对其最新成果从数学上加以证明并用于计算机立体视觉的研究中, 本文的模型就是基于这一考虑而提出来的, 通过随机点图和 pentagon 像对的实验, 表明了本算法是高效而可靠的.

REFERENCES

- 1 Poggio G F, Motter B C, Squatrito S, *et al.* Responses of neurons in visual cortex(v1 and v2) of the Alert Macaque to dynamic random-dot stereograms, *Vision*, 1985, **25**:397-406
- 2 Freeman R D, Ohzawa I. On the neurophysiological organization of binocular vision, *Vision*, 1990, **30**: 1661-1676
- 3 Ohzawa I, Freeman R D, De Aangelis G C. Stereoscopic depth discrimination in the visual cortex: neurons ideally suited as disparity detectors, *Science*, 1990, **249**, 1037-1041
- 4 Nomura M, Matsumoto G, Fujiwara S A. binocular model for the simple cell, *Biol. Cybern.* 1990, **63**, 237-242
- 5 Fleet D J. Disparity from local-weighted phase-correlation, *IEEE Inter. Confer. Systems, Man and Cybernetics, San Antonio*, 1994, 48-56

AN ALGORITHM FOR DISPARITY ESTIMATION BASED ENTIRELY ON THE VISUAL PHYSIOLOGICAL CHARACTER

LI Feng ZHOU Yuan-Hua

(Institute of Image Communication & Information Processing, Shanghai Jiaotong University, Shanghai 200030, China)

Abstract Based on the research work of Freeman *et al.* about the visual physiological character, a mathematical model of complex binocular cell was deduced and thereby an algorithm of disparity computing was realized by using this model. Experiments on random dots stereopair and pentagon stereopair show that this method can be used to estimate disparities rapidly and efficiently.

Key words simple binocular cell, complex binocular cell, disparity.