



未知光源位置环境中物体形状恢复的 神经网络方法研究

田英利 马颂德

(中国科学院自动化研究所模式识别国家重点实验室 北京 100080)

徐孔达

(香港中文大学电子工程系)

摘要 用神经网络方法解决未知光源位置环境中物体三维形状恢复的问题. 对漫反射表面, 用神经网络方法由已知表面形状物体及其对应图像的灰度值进行学习, 所得权值可视为环境光源参数. 由此可恢复同样光源环境中其它物体的三维形状. 实验证明, 神经网络方法可以解决未知光源位置环境(包括多个光源)中漫反射表面物体的三维形状恢复问题.

关键词 神经网络, 形状恢复, 漫反射表面.

SHAPE RECOVERY IN UNKNOWN ENVIRONMENT BY NEURAL NETWORKS

TIAN Yingli MA Songde

(National Laboratory of Pattern Recognition

Institute of Automation, Chinese Academy of Sciences, Beijing 100080)

H. T. Tsui

(Dept. of Electronic Engineering, the Chinese University of HongKong)

Abstract In this paper, we propose a new method based on neural networks to recover the 3D shape of an object with Lambertian surface. The method can be applied under multiple light sources of unknown distances and sizes. A multi-layer neural network is used to learn the mapping between the image intensities and the corresponding surface normals. A sphere is used as the calibration object in training the neural network by a backpropagation algorithm. The weights of the network can be regarded as the environment parameters. The efficiency of the method is verified by simulated and real experiments.

Key words Neural network, shape recovery, Lambertian surface.

1 引言

对漫反射物体的形状恢复问题,人们提出了各种各样的解决方法,如由灰度恢复物体形状方法(shape-from-shading),光度立体法(photometric stereo)等.作者也提出在几种扩展光源下的物体形状恢复方法.但所有这些方法中光源位置或方向必须已知.如果能解决未知光源参数在实际环境中的形状恢复问题,无疑可以把物体三维形状恢复方法的研究向实际应用推进了一大步.近年来,神经网络在许多领域得到广泛的应用. Lehky 等人^[1]首次把神经网络引入由灰度恢复物体形状方法(shape-from-shading)中.他们用多层前向网络通过上千幅已知物体表面形状的图像进行学习,然后得到未知表面物体的形状. Iwahori 等人^[2,3]用神经网络方法解决由镜面反射恢复物体形状及求解光度立体法(photometric stereo)问题中的本征向量.这些研究表明,神经网络方法可以有效地应用在物体形状恢复方面.

本文提出了一种只由一幅图像恢复未知光源参数环境中漫反射表面物体形状的新方法.作者没有直接计算物体的表面形状,而是用多层前向网估计环境的光源参数.对一种环境,只用一已知表面形状的物体学习输入(图像灰度值)和输出(物体表面法线)之间的关系.为了消除图像噪声的影响及引入表面连续性约束,用每一 3×3 窗口的九个像素的灰度值对应其中心像素点的物体表面法线.实验证明,神经网络方法可以解决未知光源位置环境中漫反射表面物体的三维形状恢复问题.

2 形状估计

2.1 实际环境中物体三维形状恢复问题

实际环境中的光源,大部分是球扩展光源、矩形光源、圆柱形光源或它们的组合的光.漫反射物体上点 P 在球扩展光源,矩形光源或圆柱型光源照射下,其亮度 B' 与此点法线方向 $n(\theta_n, \varphi_n)$ 的关系为

$$B' = A \sin \theta_n \cos \varphi_n + B \sin \theta_n \sin \varphi_n + C \cos \theta_n,$$

其中 $A = (A_1 + A_2 + \dots + A_i)$, $B = (B_1 + B_2 + \dots + B_i)$, $C = (C_1 + C_2 + \dots + C_i)$ 为光源参数, i 为光源个数.

如果光源参数 A, B, C 已知,可以用其它形状恢复方法得到物体形状.这里作者用神经网络方法恢复未知光源位置环境中物体三维形状.

2.2 用前向网络描述物体三维形状

许多实际环境中的光源参数是未知的.本文用神经网络方法对光源参数进行估计.再用一已知表面形状物体(本文称其为校准物体)的法线矢量作为输出,借此物体在未知光源参数环境中的图像作为输入对神经网络进行训练,从而得到表示环境参数的网络权值.然后,便可用此网络恢复同样环境中其他物体的表面形状.应用三层前向网络,三个输出元 n_x, n_y, n_z 表示物体法线矢量 n

$$n_x(u, v) = \sum_{j=1}^N w_{j1}(2) \phi(f_j(u, v)) + w_1(2),$$

$$n_y(u, v) = \sum_{j=1}^N w_{j2}(2) \phi(f_j(u, v)) + w_2(2),$$

$$n_z(u, v) = \sum_{j=1}^N w_{j3}(2) \phi(f_j(u, v)) + w_3(2),$$

其中 N 为隐层节点数; $w_{j1}(2), w_{j2}(2), w_{j3}(2)$ 为隐层节点到输出节点的网络权值; $w_1(2), w_2(2), w_3(2)$ 为输出节点 1, 2, 3 的阈值; $\phi(x) = 1/(1+e^{-x})$ 为 sigmoid 函数; $f_j(u, v)$ 为第 j 个隐层节点的输入, 且有

$$f_j(u, v) = \sum_{i=1}^9 w_{ij}(1) I_i + w_j(1) = w_{1j}(1) I(u-1, v-1) + w_{2j}(1) I(u-1, v) + w_{3j}(1) I(u-1, v+1) + w_{4j}(1) I(u, v-1) + w_{5j}(1) I(u, v) + w_{6j}(1) I(u, v+1) + w_{7j}(1) I(u+1, v-1) + w_{8j}(1) I(u+1, v) + w_{9j}(1) I(u+1, v+1) + w_j(1),$$

其中 $I_i, i=1, 2, \dots, 9$ 为输入图像上以像素 (u, v) 为中心的 3×3 窗口的灰度值; $w_{ij}(1)$ 为第 i 个输入节点到第 j 个隐层节点的权值; $w_j(1)$ 为第 j 个隐层节点的阈值.

作者用一个球作为校准物体训练神经网络. 网络权值的最小二乘解可以通过求所有表面法线误差和 E 的最小值而获得, 其中

$$E = \sum_{(u,v) \in D_n} E_{uv} = \frac{1}{2} \sum_{(u,v) \in D_n} \{ [n_x(u, v) - \hat{n}_x(u, v)]^2 + [n_y(u, v) - \hat{n}_y(u, v)]^2 + [n_z(u, v) - \hat{n}_z(u, v)]^2 \},$$

式中 D_n 为图像域; $n_x(u, v), n_y(u, v), n_z(u, v)$ 为校准球上对应像素点 (u, v) 的法线方向实际值; $\hat{n}_x(u, v), \hat{n}_y(u, v), \hat{n}_z(u, v)$ 为由神经网络计算所得此点法线方向. 为使上式中 E 最小, 本文引用 BP 算法. 对从节点 k 到节点 l 的第 m 个训练样本, 其广义 δ 规则可描述为

$$w_{kl}^{(m+1)} = w_{kl}^{(m)} + \Delta w_{kl}^{(m)},$$

$$\Delta w_{kl}^{(m)} = \eta \frac{\partial E_{uv}}{\partial w_{kl}^{(m)}} = \eta \frac{\partial E_{uv}}{\partial \text{net}_l} \frac{\partial \text{net}_l}{\partial w_{kl}^{(m)}} = \eta \frac{\partial E_{uv}}{\partial \text{net}_l} o_k = \eta \delta_l o_k,$$

式中 η 为学习常数; net_l 为节点 l 的输入; o_k 为节点 k 的输出. 如果节点 l 为输出节点, 且其输出为 $o_l (o_1 = \hat{n}_x, o_2 = \hat{n}_y, o_3 = \hat{n}_z)$. 所以

$$\delta_l = \delta_p = \begin{cases} \frac{\partial E_{uv}}{\partial \hat{n}_x} \frac{\partial \hat{n}_x}{\partial \text{net}_p} = - (n_x - \hat{n}_x) \phi'(\text{net}_p), & p = 1, \\ \frac{\partial E_{uv}}{\partial \hat{n}_y} \frac{\partial \hat{n}_y}{\partial \text{net}_p} = - (n_y - \hat{n}_y) \phi'(\text{net}_p), & p = 2, \\ \frac{\partial E_{uv}}{\partial \hat{n}_z} \frac{\partial \hat{n}_z}{\partial \text{net}_p} = - (n_z - \hat{n}_z) \phi'(\text{net}_p), & p = 3. \end{cases}$$

如果节点 l 为隐层节点, 则有

$$\delta_l = \phi'(\text{net}_l) \sum_{p=1}^3 \delta_p w_{lp}.$$

3 神经网络方法恢复未知光源位置环境中物体三维形状的应用

应用三层神经网络(输入层、输出层和隐层)实现物体形状恢复. 隐层节点数为 1×12 . 输入图像尺寸为 128×128 , 输出物体表面法线尺寸为 126×126 . 为了同时应用图像灰度信

息及物体表面连续性约束,每一个 3×3 窗口的九个像素的灰度值作为输入,对应此窗口的中心像素点的物体表面法线矢量作为输出对网络进行训练.此输入窗口的移动步长为一个像素.

为了得到正确的网络权值,必须对输入与输出进行规一化.对输入图像,首先找出整个图像中最大灰度值,然后使每一像素点的灰度值除以图像最大灰度值.这样规一化后的输入图像的灰度值便介于0—1之间.

$$I_{\text{normalized}} = \frac{I_{\text{measured}}}{I_{\text{max}}}$$

对输出的表面法线 n_x, n_y, n_z , 其中 n_x, n_y 的取值范围为 $[-1, 1]$, n_z 的取值范围为 $[0, 1]$. 所以用下式把 n_x, n_y 的取值范围规一到 $[0, 1]$:

$$n'_x = \frac{n_x + 1}{2}, \quad n'_y = \frac{n_y + 1}{2}.$$

4 实验结果

4.1 模拟图像实验结果

作者用光线追踪法在单个光源和多个光源下分别产生不同颜色和不同反射率的不同物体上验证此方法.一个已知形状的校准球及其深度图用来训练神经网络.实验表明,对未知光源位置环境(包括多个光源)中漫反射表面物体,神经网络方法可以有效的恢复多光源环境下的不同颜色、不同大小和不同表面反射率的物体表面形状.由于篇幅限制,只给出实际图像实验结果.

4.2 实际图像实验结果

实验中,用了两个矩形光源和一个球光源.用一个具有16mm镜头的CCD摄像机获取图像.实验结果如图1所示.这里给出了(a),(b)两组实验结果.可以看出,神经网络方法对未知光源环境中实际物体也可以很好地恢复其形状.

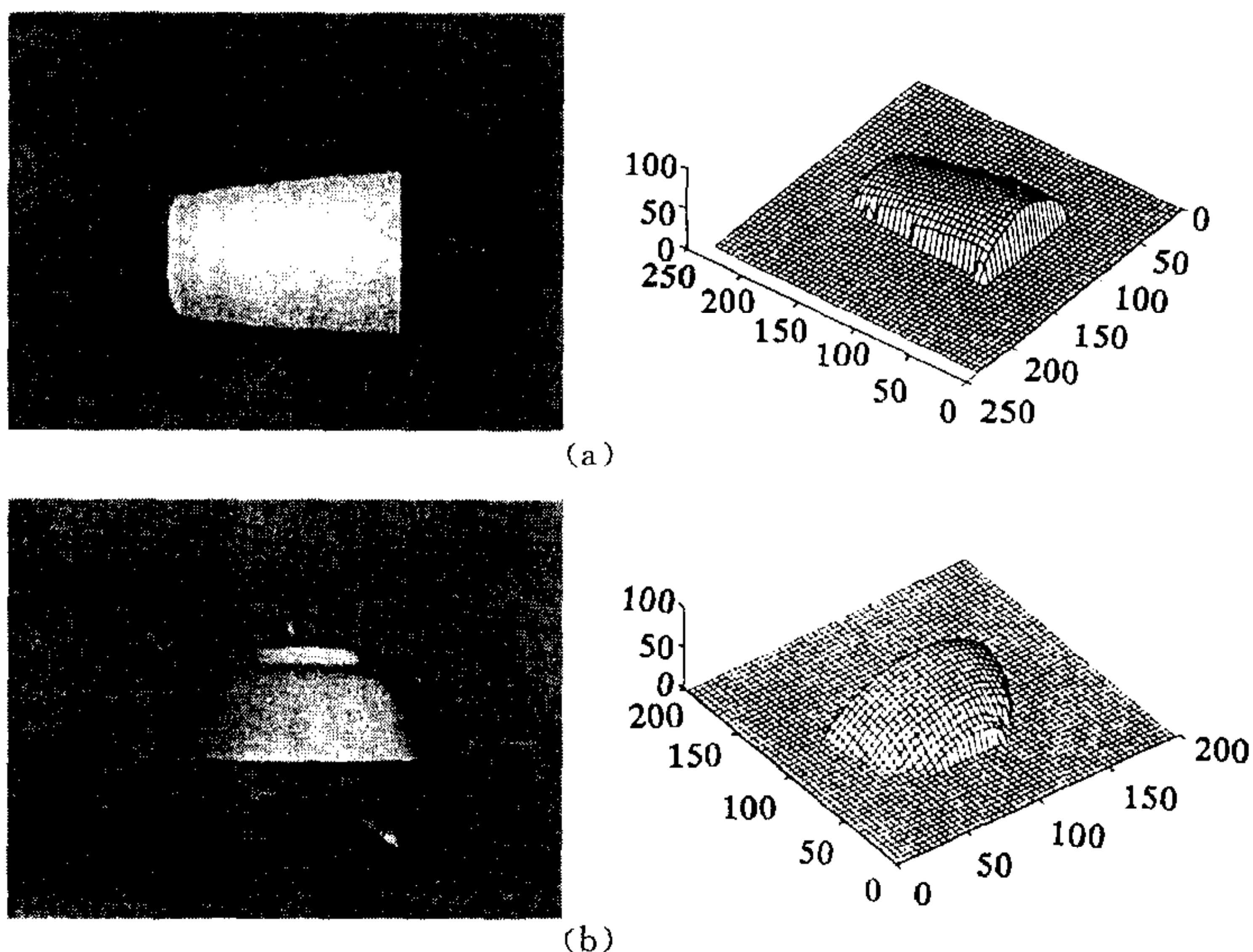


图1 用神经网络方法恢复的多光源实际环境下物体深度图

5 结论和讨论

本文提出了一种有效的用神经网络方法恢复未知光源位置环境(包括多个光源)中漫反射表面物体的三维形状的方法. 对漫反射表面, 由已知表面形状物体及其对应图像的灰度值对神经网络进行学习, 所得权值可视为环境光源参数. 然后恢复此光源环境中的物体三维形状. 实验证明, 本方法可以有效地恢复漫反射物体的表面形状. 对本文工作下一步的研究包括: 1) 用球这样具有简单形状的物体作为校准物体训练神经网络, 对复杂物体的形状恢复是否足够? 2) 在我们的实验中, 3×3 的输入窗口是适合的. 对更复杂的物体, 输入窗口大小的影响需进一步研究.

参 考 文 献

- 1 Lehky S R, Sejnowski T J. Network model of shape-from-shading neural function arises from both receptive and projective fields. *NATURE*, 1988, **333-2**:452—454
- 2 Iwahori Y, Tanaka H, Woodham R, Ishii N. Photometric stereo for specular surface shape based on neural network. *IEICE Trans. INF. & SYST.*, 1994, **E77-D(4)**:498—506
- 3 Iwahori Y, Woodham R, Bagheri A. Principal Components Analysis and Neural Network of Implementation Photometric Stereo. In: *proc. of the Workshop on Physics-based Modeling in Computer Vision*, 1995, 117—125

田英利 女, 1966年生. 1996年于香港中文大学获博士学位. 现在中国科学院自动化研究所模式识别国家重点实验室工作. 研究方向为计算机视觉、图形学和图像处理.

马颂德 1946年生于上海. 1968年毕业于清华大学自动化系, 1983年获巴黎第六大学博士学位, 1986年获巴黎第六大学 Science d'Etat 博士学位. 1983年至1984年为 University of Maryland 的客座研究员, 1984年至1986年为法国 INRIA 的 Robot Vision Group 的客座研究员, 现为中国科学院自动化研究所所长. 主要研究领域为计算机视觉、神经计算和基于传感器的机器人控制.