

一种快速精确的人脸三维形状重构新方法

袁友伟

湛含辉

(株洲工学院计算机系, 株洲 412008)

(株洲工学院环保研究所, 株洲 412008)

摘要 针对传统 SFS(Shape from Shading)的不足,提出了一种新的基于 BP 神经网络的明暗恢复形状的方法,该方法是基于兰伯特(Lambertian)反射模型的改进算法,利用了 BP 神经网络强的非线性映射能力,将 Lambertian 表面反射模型与光滑表面模型相结合,然后再利用一些已知条件,构成 SFS 问题的正则化模型;变换不同的照明条件,将模型平移或旋转获得多幅图象,以增加约束条件;计算出误差补偿参数去修正邻域内的三维误差。由于考虑了邻域的平均值,使算法的稳定性和精确性都得到了加强。实例表明,该算法较传统的算法更快和更精确。

关键词 计算机图象处理(520·6040) 从明暗恢复形状 兰伯特(Lambertian) 反射模型 BP 神经网络 误差补偿
中图分类号: TP391.41 **文献标识码**: A **文章编号**: 1006-8961(2003)08-0907-05

An Approach for Fast Human Face Three-dimensional Shape Reconstruction

YUAN You-wei

(Department of science and technology, Zhuzhou Institute of technology, Zhuzhou 412008)

ZHAN Han-hui

(Research Institute of Environment, Zhuzhou Institute of technology, Zhuzhou 412008)

Abstract Shape from shading (SFS) is one of the critical techniques to recover three-dimensional shape in computer vision, which obtains 3D of the visible surface of an object from only one images of its using the shading knowledge in the given picture. The former methods are ill-posed. In this paper, the algorithm of shape from shading based on BP neural networks is proposed, which is based on Lambertian reflection model. The suggest method has error compensation. We demonstrate the robustness of our approach to strong illumination variations and with significant pose variations. The recovered face shape is then shown along with the original surface. In comparison with the traditional methods, examples show that the algorithm is verified to be accurate and applicable.

Keywords Computer image processing, Shape from shading, Lambertian reflection, Back-propagation neural network, Error compensation

0 引言

面部是人们在相互交往中最关注的部位,人或动物面部的计算机三维造型是近几年来图形学领域里一个十分活跃的研究方向^[1]。从明暗恢复形状(shape from shading,简称 SFS)是计算机视觉中三维形状恢复(3D shape recovery)问题的关键技术之一,其任务是利用单幅图象中物体表面的明暗变化

来恢复三维形状^[2]。

现已出现了多种 SFS 算法^[3],其中,最重要的有最小值方法、演化方法、局部方法和线性化方法 4 类。传统的 SFS 方法作如下假设:(1)照射的光线为平行光;(2)反射模型为兰伯特(Lambertian)反射模型;(3)投影为正投影^[3]。

物体表面图象光亮度 $I(x, y)$ 仅由入射光 i 及相应的物体表面法向量 $n(x, y)$ 决定,它们的关系为^[4]

$$I(x, y) = R(i, n(x, y)) \quad (1)$$

基金项目:国家自然科学基金项目(50274080);湖南省教育厅科研项目(02C643)

收稿日期:2002-05-20;改回日期:2003-02-20

其中, $R()$ 为反射函数. 一般情况下, 仅由该模型所确定的 SFS 问题是病态的, 在求解过程中, 还存在多解的问题, 且计算量大, 误差较大, 因此该方法存在一定的局限性.

神经网络是近年来发展较快的一个交叉研究学科. 人工神经网络模型具备人脑思维的一些特征, 如自组织、联想记忆、非线性、大规模并行连接等, 并且具有强大的学习能力. 1988 年, 美国南加州大学 Zhou 等人首次将 Hopfield 神经网络方法用于二值图象的恢复, 并建立了一种新的图象恢复方法^[5]. 这种方法既解决了正约束问题, 也无需对图象作出广义平稳过程假设, 因而在一定程度上克服了传统图象恢复方法的主要缺陷, 但这种神经网络采用的神经元数巨大, 故硬件难以实现.

由于 BP 神经网络有强的非线性映射能力, 大规模并行处理、分布式存储以及网络的计算时间几乎为零, 且对退化的连续图象进行恢复能取得很好的效果. 更主要的是, BP 网络不需要巨大复杂的预处理就可以应用于真实图象的识别, 通过对样本的学习, 相似的样本在输出端相似, 另外, 它的收敛速度快, 只要有足够的内存单元来储存它的学习权值, 它就可以逼近任意的非线性关系. 因此 BP 神经网络运行速度快, 易满足图象实时处理的要求.

针对 SFS 方法中的不足, 本文提出了一种新的基于 BP 神经网络的明暗恢复形状的方法, 它是基于兰伯特(Lambertian)反射模型的改进算法, 提出运用不同照明情况下, 变换模型(平移或旋转)获得多幅图象, 增加约束条件. 先假设所研究的物体表面是光滑的, 即认为物体表面高度函数是连续的; 将上述表面反射模型与光滑表面模型相结合, 再利用一些已知条件(如关于问题表面的初边值条件), 就构成 SFS 问题的正则化模型^[6]. 此模型能改变问题的病态性, 且具有自动误差补偿的功能. 其中将考虑到镜头的投影、噪声的混入、数字化抽样等因素. 这一方法已成功地应用于人脸的三维形状恢复中, 且当光源方向和人脸模型角度变化时, 均能很好地恢复其三维形状, 能得到很好的精度.

1 兰伯特(Lambertian)反射模型

数字成像过程是指空间中三维物体信息到二维数字图象的变换过程; 视觉过程则是从二维数字图象信号中提取出被观测对象的三维信息的过程.

建立如图 1 所示的坐标系^[7].

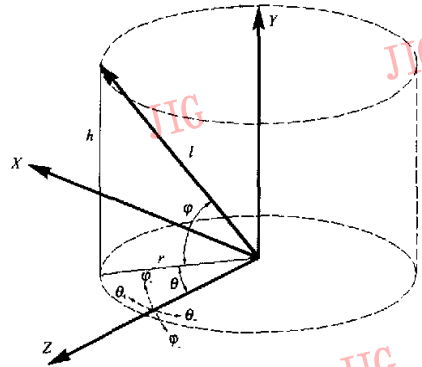


图 1 投影坐标系

(θ, φ 的下角表示以 Z 轴为界, 头部模型向右为正“+”向左为负“-”)

图 1 表示了人脸模型在半径为 r 高度为 h 的圆柱形内, θ 表示模型平移的角度, φ 表示偏角. 模型表面法方向上的点 $S(h, \theta)$ 的计算如下^[2]

$$S(h, \theta) \equiv [x(h, \theta), y(h, \theta), z(h, \theta)] \\ = [x_0 + r(h, \theta)\cos\theta, y_0 + h, z_0 + r(h, \theta)\sin\theta] \quad (2)$$

其中, (x_0, y_0, z_0) 表示坐标系的原点, 脸部平移的角度可直接引用其实际测量得到的数据, $S(h, \theta)$ 由以下的交叉积得到

$$n_s(h, \theta) \propto \frac{\partial \vec{S}}{\partial h} \times \frac{\partial \vec{S}}{\partial \theta} \quad (3)$$

其中, $n_s(h, \theta)$ 为产生表面图象光亮度 $I(x, y)$ 的正则化模型. 其算法如下

$$n_s(h, \theta) = \frac{1}{\sqrt{r^2 + \left(\frac{\partial r}{\partial \theta}\right)^2 + \left(\frac{\partial r}{\partial h}\right)^2 + r^2 \left(\frac{\partial r}{\partial \theta}\right)^2 - 2 \frac{\partial r}{\partial \theta} \frac{\partial r}{\partial h}}} \cdot \begin{bmatrix} \frac{\partial r}{\partial \theta} \cos\theta - \frac{\partial r}{\partial h} \cos\theta + r \sin\theta \\ -r \frac{\partial r}{\partial \theta} \\ -\frac{\partial r}{\partial h} \sin\theta + \frac{\partial r}{\partial \theta} \sin\theta + r \cos\theta \end{bmatrix} \quad (4)$$

这样, 兰伯特(Lambertian)反射的反射光为漫反射, 其模型可以表示为

$$I(h, \theta) = \eta(h, \theta) n_s(h, \theta) \cdot i \quad (5)$$

其中, $i = (i_x, i_y, i_z)$ 是表示光源方向的单位向量, $\eta(h, \theta)$ 是反射率.

在实验中, 人脸模型的运动范围为: 平移角度 $\theta \in [-15^\circ, +15^\circ]$, 偏角 $\varphi \in [-5^\circ, +5^\circ]$, 光源的方向在 $\theta \in [-30^\circ, +30^\circ]$ 和 $\varphi \in [-30^\circ, +30^\circ]$ 的范围内变化.

2 BP 神经网络的三维图象恢复算法及误差补偿

由于光照射到一个物体表面,其反射光是环境反射、漫反射和镜面反射 3 个分量的结合,而 Lambertian 反射仅考虑光的漫反射,是不切实际的,因此 Lambertian 反射模型也存在一定的局限性和病态性.

2.1 BP 神经网络

反传网络 BP 是一种映射网络^[8],它能根据训练样本对 (x, y) 来近似模型函数 $f(x)$,其中在样本点上 $y=f(x)$. 函数 $y=f(x)$ 是非线性的,但反传网络 (BPN) 能够模拟非线性. 反传网络共有 3 层:输入层、隐层和输出层. 输入层中的神经元存储输入值. 神经元的输出 O_i 是 S_i 的西哥玛 (Sigma) 函数.

在训练的网络数据中,其输入的图象光亮度 and 输出图象像素的高度已标准化,中心像素点设为 0,则输入的图象光亮度和输出的图象的像素高度的范围按比例缩减为 $[-1, 1]$ 之间. 用于恢复人脸三维形状的神经网络如图 2 所示.

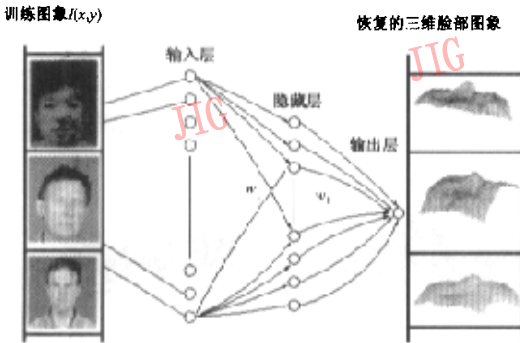


图 2 用于恢复人脸三维形状的神经网络

2.2 基于 BP 神经网络的三维图象恢复算法

人脸是一个很特殊的生物模式,在人脸的照片中,人们可以发现人脸具有明显的灰度特征,如眼睛的灰度、鼻孔的灰度、嘴巴中间的灰度值都比周围低;同时,人脸还具有很多的梯度信息,如眼睛、嘴巴、鼻子等部位的边缘信息比人脸的其他部位更丰富. 当原始特征样本的维数很大 (比如人脸图象矢量为 $92 \times 112 = 10\ 304$ 维) 时,在如此高维的原始特征空间内,用传统的扰动法和补空间法直接求解最优鉴别矢

量集几乎是不可行的. Karhunen-Loeve 提出了主成分分析法 (Principal component analysis 简称 PCA), 也称为特征脸方法 (Eigenface). 该方法将人脸图象按行 (列) 展开所形成的一个高维向量看成是一个随机向量. 因此可以采用 K-L 变换获得其正交 K-L 基底等^[9,10], 对应于其中较大特征的基底具有与人脸相似的形状, 故称其为特征脸. 利用相对较小的 Eigenface 集描述人脸, 这样每幅人脸图象就对应于一个维数较低的权向量. 同时, 主分量分析中提取的特征向量可使三维重建图象对应的特征向量的均方差最小. 提取的这些特征向量 PC 是最好的描述特征, 如眼睛、嘴巴、鼻子等. 在以下的算法中具有相同描述特征的为同一类, 如鼻子为一类, 眼睛为一类.

$I_P(x, y)$ 表示每个训练样本的主分量的输入图象光亮度数据, 变换不同的照明条件, 将同一模型平移或旋转获得 N 幅图象的光亮度数据. 同类的尺寸大小用 $S_x \times S_y$ 表示.

$$I_P(x, y) \approx \bar{I}_P(x, y) + \sum_{t=1}^T a^{(t)} \varphi_t^{(t)}(x, y) \quad (6)$$

其中, $\bar{I}_P(x, y) = \frac{1}{N} \sum_{t=1}^N I_P^{(t)}(x, y)$ 是该处图象光亮度的平均值, 平均值可以减少误差, $\varphi_t^{(t)}(x, y)$ 是数据集的主分量, PC (Principal component) 是数据集的特征向量, $t=1, 2, \dots, N, T$ 表示在已经得到的 N 幅图象的光亮度数据中抽取的第 T 个主分量特征向量, t 的大小一定小于或等于 N , 因为 N 幅图象的光亮度数据中可能有些数据不能用来作计算数据 (如眼睛照斜了, 鼻子照歪了). 在训练中抽取的特征向量 PC 的系数 $a^{(t)}$ 为

$$a^{(t)} = \sum_x \sum_y (I_P(x, y) - \bar{I}_P(x, y)) \varphi_t^{(t)}(x, y) \quad (7)$$

$$G_T \varphi_T = \lambda K_T \quad (8)$$

其中, G_T 是训练集中提取的 N 个图象光亮度的自动协方差 (auto-covariance), λ 是特征值, K_T 是主分量的特征变量, 因此基于 BP 网络恢复的三维图象的平均方差 e 为

$$e = \sum_{x=1}^{S_x} \sum_{y=1}^{S_y} \lambda \quad (9)$$

因此由二维图象恢复的三维图象的数据 $r_P(x, y)$ 为

$$r_P(x, y) \approx \bar{r}_P(x, y) + \sum_{t=1}^T e \varphi_t^{(t)}(x, y) \quad (10)$$

其中, $\bar{r}_P(x, y)$ 是由二维图象恢复的三维图象的平均数据范围.

一般情况下,由于拍摄的照片通常是含有噪声的图象,而上述算法也是一种近似解法,所以求得的 $r_P(x,y)$ 不可避免地带有误差。

2.3 误差补偿

信号与噪声总是结伴而行,三维图象的恢复还应考虑到数字成像过程中信号变换的各种因素,如镜头的几何投影、镜头的畸变与信号失真、图象信号的光电转换等因素而产生的误差,计算出误差补偿参数去修正邻域内点的三维数据,利用一些已知条件,构成了 SFS 问题的正则化模型,再利用不同照明情况下,模型的平移或旋转来获得多幅图象,增加约束条件,其具体方法是:

利用上述算法得出三维图象数据范围 $r_P(x,y)$, 同时测得该处真实的数据为 $r_P^{orig}(x,y)$, 其像素平均绝对百分比误差 *AAPEPP* (the average absolute percentage error per pixel) 为

$$AAPEPP = \frac{1}{N} \sum_{x,y} \frac{\|r_P(x,y) - r_P^{orig}(x,y)\|}{\|r_P^{orig}(x,y)\|} \quad (11)$$

这样,由于考虑了邻域的平均值,使算法的稳定性和精确性都得到了加强,为了提高精度,可进一步的细化区间,取得更多的点,但计算量也会增加。

3 仿真实验

根据实际情况,采集了 1000 多个训练样本,建立了人脸图象库,其中有 120 个不同的人,每个人均有 9 幅不同的人脸图象,形成了共 120 人 1080 幅人脸的图象库。这些人脸图象分别在不同时间、不同光照、不同头部角度、不同人脸表情(睁/闭眼,笑/严肃)和不同人脸细节(有无眼镜)条件下摄制得到。原始图象为 256 级灰度图,尺寸为 92×112 (宽 \times 高) 像素大小,人脸图象用双线性插值方法归一化到 24×28 像素大小,以减少计算处理量。

训练完成后用于对原二维图象图 3(a)、图 3(e) 分别用模型的真实值、Hopfield 神经网络和本文的 BP 神经网络进行三维图象的恢复与重建。图 3(a), 图 3(e) 为同一人脸模型得到的图象,表 1 为实验结果的量化值。图 3(a) 人脸模型在图 1 坐标系的平移角度 θ 和偏角 φ 分别为 $(-5.6^\circ, -1.8^\circ)$, 光源方向分别为 $(-20.8^\circ, -18.5^\circ)$, 图 3(e) 人脸模型的角度为 $(5.8^\circ, -0.8^\circ)$, 光源方向为 $(8.8^\circ, 13.8^\circ)$ 。图 3(b)、图 3(c)、图 3(d) 为图 3(a) 分别用模型的真

表 1 Hopfield 与 BP 神经网络恢复人脸形状运算结果

人脸角度 (θ, φ)	光源方向 (θ, φ)	Hopfield 神经网络				本文 BP 神经网络			
		CPU 时间 (s)	高度方向绝对误差(mm)			CPU 时间 (s)	高度方向绝对误差(mm)		
			Mean	Max	Variance		Mean	Max	Variance
$(-5.6^\circ, -1.8^\circ)$	$(-20.8^\circ, -18.5^\circ)$	235	0.058	0.318	0.382×10^{-2}	28	0.030	0.182	0.256×10^{-2}
$(5.8^\circ, -0.8^\circ)$	$(8.8^\circ, 13.8^\circ)$	270	0.108	0.935	0.416×10^{-2}	33	0.056	0.397	0.285×10^{-2}

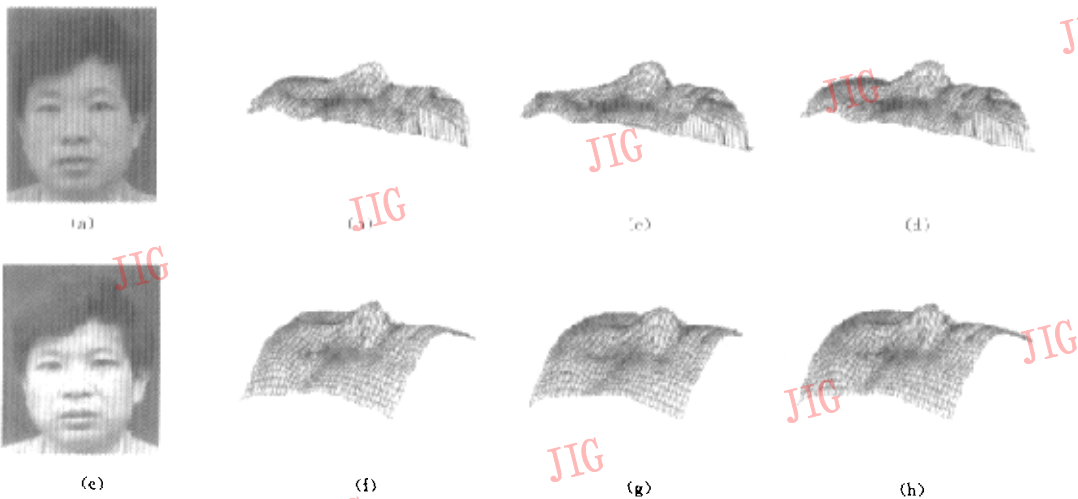


图 3 三维图象的恢复比较效果图

实值、Hopfield 神经网络和本文 BP 神经网络恢复得到的三维图象;图 3(f)、图 3(g)、图 3(h)为图 3(e)分别用模型的真实值、Hopfield 神经网络和本文 BP 神经网络恢复得到的三维图象。

从表 1 和图 3 可以看出,图 3(d)与图 3(b)几乎相同,图 3(h)与图 3(f)几乎相同,即用 BP 神经网络恢复的人脸三维图象与真实值几乎相同;而图 3(c)与图 3(b)相差较远,图 3(g)与图 3(f)相差较远,即用 Hopfield 神经网络恢复的三维图象与真实值相差较远(在放大图中可以看得更清楚)。

以上实验结果表明:用 BP 神经网络恢复得到的三维图象比 Hopfield 神经网络恢复的三维图象更接近真实值,且恢复的速度更快。

4 结 论

针对传统的 SFS 的不足,提出了一种新的基于 BP 神经网络的明暗恢复形状的方法,它是基于兰伯特(Lambertian)反射模型的改进算法,利用了 BP 神经网络强的非线性映射能力,提出运用不同照明情况下,先将 Lambertian 表面反射模型与光滑表面模型相结合,然后利用一些已知条件,构成 SFS 问题的正则化模型;再利用不同照明情况下,模型的平移或旋转获得多幅图象,以增加约束条件;计算出误差补偿参数去修正邻域内的三维误差。由于考虑了邻域的平均值,使算法的稳定性和精确性都得到了加强。

从图 3 的实例中可以看出,运用本文方法恢复的三维图象与真实图象相比,几乎相同,而用 Hopfield 神经网络方法恢复的三维图象与真实三维图象却相差较大。基于 BP 神经网络的明暗恢复形状的方法,当光源方向和人脸模型角度变化时,均能很好地恢复其三维形状,该算法具有快速性和精确性的特点,值得推广。

参 考 文 献

- 1 YUAN You-wei, YAN La-mei. Active networks for efficient and distributed network management based on an artificial neural networks [A]. In: DCABES 2001 Proceedings [C]. Wuhan, China, Wuhan University of technology press, 2001: 172~175.
- 2 Chao Siu-Yeung, Tommy W S Chow. A neural-learning based reflectance model for 3-D shape reconstruction [J]. IEEE Transactions on Industrial Electronics, 2000, 47(12): 1346~1350.
- 3 赵保军,史彩成,沈胜宏等.一种基于 CMAC 的图象恢复算法[J].中国图象图形学报,2001,6A(8):780~783.
- 4 廖熠,赵荣椿.从明暗恢复形状(SFS)的几类典型算法分析与评价[J].中国图象图形学报,2001,6A(10):953~956.
- 5 Zhou Y T, Chellappa R C. Artificial neural networks for computer vision[M]. New York: Springer-Verlag, 1992.
- 6 陈刚,戚飞虎.实用人脸系统的本征脸法实现[J].计算机研究与发展,2001,38(2):170~175.
- 7 Dibyendu Nandy, Jezekiel Ben-Aric. Shape form recognition: A novel approach for face shape from shading [J]. IEEE Transactions on Image Processing, 2001, 10(2): 206~207.
- 8 Lisboa P G J, 刑春颖, 阳影. 现代神经网络应用[M]. 北京: 电子工业出版社, 1996: 54~58.
- 9 Samuel Pierre, Hassance Said, Wilfried G Probst. An artificial neural network approach for routing in distributed computer networks[J]. Engineering Application of Artificial Intelligence, 2001, 14(4): 51~60.
- 10 Turk M, Pentland A. Eigenface for recognition[J]. Journal of Cognitive Neuroscience, 1991, 3(1): 71~86.



袁友伟 1966年生,株洲工学院计算机系副教授,副主任。主要从事神经网络、数字图象处理的研究,发表论文50余篇。



湛含辉 1961年生,1989年获中国矿业大学矿物工程博士学位,1991年中南工业大学博士后出站,现为株洲工学院环保研究所所长,研究员、博导。主要从事图象处理和二次流基础理论等方面的研究。