

相位特征在三维物体识别中的应用^{*}

申金媛^{1)†} 李现国¹⁾ 常胜江²⁾ 张延²⁾

¹⁾ 郑州大学信息工程学院, 郑州 450052)

²⁾ 南开大学现代光学研究所, 光电信息技术科学教育部重点实验室, 天津 300071)

(2004 年 12 月 7 日收到, 2005 年 3 月 21 日收到修改稿)

提出利用物体的相位特征联合神经网络的方法对透明半透明三维物体进行识别. 首先利用波长扫描数字全息技术和数字再现技术提取物体的相位特征, 然后将物体的这些相位特征作为学习模式训练一个 BP 神经网络, 最后利用训练好的网络对三维物体进行识别. 实验表明, 对于具有小尺度变化的透明半透明三维物体识别, 该方法的正确识别率为 100%.

关键词: 相位特征, 波长扫描技术, 数字全息, BP 神经网络

PACC: 4230S, 4240, 4240F

1. 引 言

三维物体识别是当今一个非常活跃的研究领域, 半个世纪以来受到各国研究者的广泛关注, 并提出了众多的理论和方法^[1-6], 但几乎所有的方法都是基于物体的几何特征或者它的颜色、纹理等特征进行识别的, 这在很多情况下是切实可行的, 但是, 对于一些几乎是透明的物体, 比如说一些生物薄膜、细胞等, 这些方法显然是行不通的. 这样的物体其几何特征、颜色或纹理可能没有多少变化, 然而经过它们的光却会改变其相位, 据此本文提出一种基于相位特征的三维(3D)物体识别方法.

利用全息术不仅可以记录三维物体的振幅, 而且还可以记录物体的相位, 相当多的人都致力于全息技术的实现及其应用方面的研究^[7,8]. 本文提出了一种提取三维物体的相位特征的方法——利用波长扫描数字全息技术^[9]提取物体的相位特征. 两束光干涉的强度变化可用余弦平方表示, 如果用一系列等间隔波长记录和再现全息, 那么同一个再现位置的不同波长的再现像叠加所得到的像的强度就由一系列等间隔波长的余弦叠加, 因此叠加像的强度变化变得更加尖锐和细窄. 其中余弦中的角与相位有

关, 每一个再现面上的强度变化对应着三维物体中具有相同相位的点, 即物体中具有相同相位的点聚焦于同一再现面, 不同相位的点聚焦于不同的再现面. 因而, 一个三维物体的相位信息可以用一系列不同位置的再现像来表征. 同传统的再现方法相比, 数字再现波长可以和记录时的波长完全相同, 不会产生任何由于波长不匹配造成的误差, 数字全息避免了全息像的多次曝光和化学处理过程, 因而有高的衍射效率.

得到物体的相位特征后, 如何对物体进行识别呢? 本实验选择被广泛应用在预测、识别、控制等领域的具有的高度并行性和自学习自适应能力的神经网络方法^[10-18]. 作为一个特殊的映射机理, 神经网络可以把不同的模式映射到一个或多个神经元中, 实现对这些模式的识别或者分类. 因此, 当一个神经网络模型用三维物体的相位特征训练后, 网络的互联权重就已经包含了该训练物体的相位信息, 依此可以用来对输入的物体进行识别.

2. 基本原理

2.1. 数字全息术

由于计算机技术和数字成像设备的发展, 使得

^{*} 国家自然科学基金(批准号: 60277022), 河南省杰出青年基金, 教育部留学回国人员科研启动基金, 博士点基金(批准号: 20030055022)资助的课题.

[†] E-mail: jyshen@zsu.edu.cn, jyshen@nankai.edu.cn

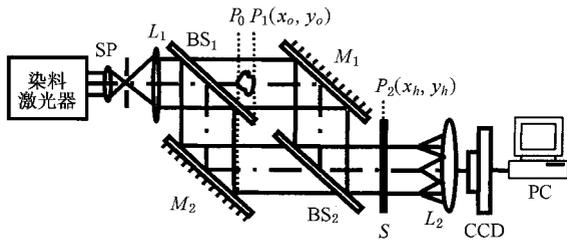


图 1 数字全息实验装置 SP 针孔滤波器; L_1, L_2 透镜; M_1, M_2 反射镜; BS_1, BS_2 分束器; P_0, P_1, P_2 平面; S 漫散屏

数字全息和数字再现技术更加行之有效^[19]. 因而传统的全息照相记录材料就可以用数字成像设备代替, 比如说数码相机、CCD (charged couple device) 等. 这样就避免了对全息照片繁琐的化学处理过程, 而且还可以通过减去由物光和参考光在全息面形成的直流项, 除去 0 级衍射像, 并可消除部分像差、噪声及记录过程中底片非线性等因素的影响, 整个过程简单, 有利于进行定量分析和测量.

采用如图 1 所示装置制作数字全息. 假设物体是半透明的, 所以透过率可写为 $C \exp(ikd_0(x_0, y_0))$. 其中 C 是常数, $d_0(x_0, y_0)$ 表示光透过物体的光程, 与物体的厚度和折射率有关, 它与相位变化成正比. 设光束进入物体前(平面 P_0) 场分布为 1, 那么经过物体后的光场分布 O_1 可表示为

$$O_1 = C \exp(ikd_0(x_0, y_0)). \quad (1)$$

设物体到全息面 P_2 的距离为 Z_0 , 在平面 P_2 处, 物体的光场分布 O_2 可表示为

$$O_2 = C_1 \exp(ikZ_0) \iint dx_0 dy_0 O_1 \exp(ik((x_0 - x_h)^2 + (y_0 - y_h)^2) / 2Z_0).$$

$$+ (y_0 - y_h)^2) / 2Z_0) \\ = C_2 \exp(ikZ_0) \iint dx_0 dy_0 \exp(ikd_0(x_0, y_0)) \\ \times \exp(ik((x_0 - x_h)^2 + (y_0 - y_h)^2) / 2Z_0). \quad (2)$$

参考光 R 正入射到全息面 P_2 处, 则干涉后所得到的全息图像可表示为

$$H_0 = |R|^2 + |O_2|^2 + R \times O_2^* + R^* \times O_2, \quad (3)$$

式中 $*$ 表示共轭. 利用实验装置(图 1), 借助于数码相机或 CCD 很容易得到 $H_0, |R|^2, |O_2|^2$.

参考光束 R 可以表示为

$$R = \exp(ikZ_0), \quad (4)$$

因而, 利用计算机可以得到消除直流噪声的数字全息 H :

$$H = R \times O_2^* + R^* \times O_2 \\ = C_2 \iint dx_0 dy_0 \exp(-ikd_0(x_0, y_0)) \\ \times \exp(-ik((x_0 - x_h)^2 + (y_0 - y_h)^2) / 2Z_0) \\ + C_2 \iint dx_0 dy_0 \exp(ikd_0(x_0, y_0)) \\ \times \exp(ik((x_0 - x_h)^2 + (y_0 - y_h)^2) / 2Z_0). \quad (5)$$

2.2. 数字再现

得到数字全息信息后, 用共轭参考光照射全息可以对原物体进行再现. 这里利用计算机数字再现的方法实现再现, 由菲涅尔衍射原理, 可以得到位于 Z_i 的再现像的场分布表达式为

$$A_i = C_3 \exp(ikZ_i) \iint dx_h dy_h R^* \times H \times \exp(ik((x_h - x)^2 + (y_h - y)^2) / 2Z_i) \\ = \begin{cases} C_4 \left\{ \exp(-ikd_0(x, y)) + C_5 \iint \exp\left(ik\left[d_0(x_0, y_0) + \frac{(x - x_0)^2 + (y - y_0)^2}{4Z_0}\right]\right) dx_0 dy_0 \right\}, & Z_i = Z_0, \\ C_4 \exp(ik(Z_i - Z_0)) \left\{ \iint \exp\left(-ik\left[d_0(x_0, y_0) + \frac{(x - x_0)^2 + (y - y_0)^2}{2(Z_0 - Z_i)}\right]\right) dx_0 dy_0 \right. \\ \quad \left. + C_5 \iint \exp\left(ik\left[d_0(x_0, y_0) + \frac{(x - x_0)^2 + (y + y_0)^2}{2(Z_0 + Z_i)}\right]\right) dx_0 dy_0 \right\}, & Z_i \neq Z_0, Z_i \neq -Z_0, \\ C_4 \exp(ik(-2Z_0)) \left\{ \exp(ikd_0(-x, -y)) + C_5 \iint \exp\left(-ik\left[d_0(x_0, y_0) \right. \right. \right. \\ \quad \left. \left. + \frac{(x - x_0)^2 + (y - y_0)^2}{4Z_0}\right]\right) dx_0 dy_0 \right\}, & Z_i = -Z_0, \end{cases} \quad (6)$$

其中 Z_i 表示全息面到重构像的距离, R^* 表示再现光是原参考光的共轭. 显然, 由于所用的数字全息中不存在直流项, 所以, 数字再现像没有 0 级衍射像.

2.3. 波长扫描

当使用波长扫描时, 在 Z_i 面形成的图形是多个波长所对应的再现像的叠加. (考虑 $Z_i \neq Z_0, Z_i \neq -Z_0$ 位置)

$$\begin{aligned}
 A_{wi} &= \int A_i dk \\
 &= C_1 \iint dx_0 dy_0 \delta(Z_i - Z_0 - d_0(x_0, y_0) - \frac{(x - x_0)^2 + (y - y_0)^2}{2Z_0 - Z_i}) \\
 &\quad + C_2 \iint dx_0 dy_0 \delta(Z_i - Z_0 + d_0(x_0, y_0) + \frac{(x - x_0)^2 + (y - y_0)^2}{2Z_0 + Z_i}), \quad (7)
 \end{aligned}$$

假设 $\Delta Z_i = Z_i - Z_0$, 则 A_{wi} 可以表示为

$$A_{wi} = \begin{cases} C_1 \iint dx_0 dy_0 \delta(\Delta Z_i - d_0(x_0, y_0) + x - x_0, y - y_0) + \text{noise} & \Delta Z_i \geq 0 \\ C_2 \iint dx_0 dy_0 \delta(\Delta Z_i + d_0(x_0, y_0) + x - x_0, y - y_0) + \text{noise} & \Delta Z_i < 0 \end{cases} \quad (8)$$

这里 C_1, C_2 为常数. 显然, 在某一个再现面 Z_i 其波长扫描后的再现像是由所有 $d_0(x_0, y_0)$ 等于 ΔZ_i ($\Delta Z_i \geq 0$ 时) 或者 $-\Delta Z_i$ ($\Delta Z_i < 0$ 时) 的点组成的轮廓线以及一些附加项(这里称之为噪声)组成, 而 $d_0(x_0, y_0)$ 与透明或半透明三维物体在光轴方面的相位是等价的, 因此不同位置的波长扫描的数字全息的再现像包含有物体的某一相位特征. 不同的 ΔZ_i 即不同的再现平面可表征一个三维物体不同的相位分布情况, 因此一个透明或半透明的三维物体可以由一系列不同位置的再现像来表征.

利用图 1 中的光学系统对波长扫描数字全息及数字再现提取三位物体的相位特征进行了验证. 图 2 中的(c)是所采用的物体 (c) 由透明片(a)和两片显微镜用基片(b)组成, 这样物体 A 部分的厚度可以忽略不计, 即光经过后相位几乎不变, B 和 C 两部分的厚度为 d , 光经过后相位变化 knd , k 是所用光的波矢, n 是基片的折射率, d 是基片的厚度, 在 (1)~(8) 式中, 将 nd 合在一起, 用 $d_0(x, y)$ 表示光经过后光程差的变化. 物体 D 部分的厚度为 $2d$, 光

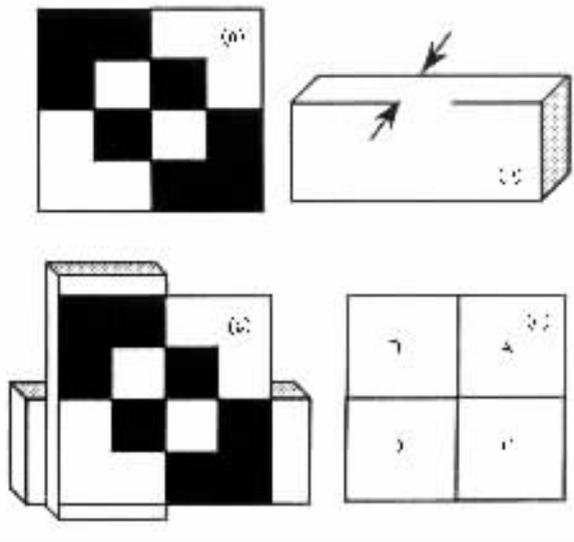


图 2 用于制作数字全息的物体

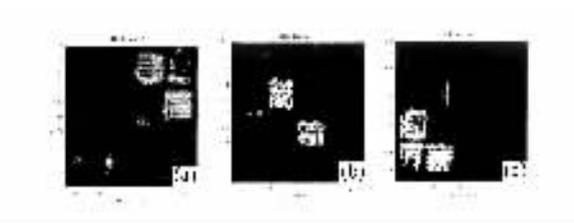


图 3 不同再现位置的数字再现像

经过后相位变化 $2knd$. 图 3 是再现距离分别为 $271.6\text{cm}, 272.1\text{cm}, 272.6\text{cm}$ 的数字再现像, 显然 271.6cm 的再现像对应于相位变化为 α (物体的 A 部分) 的部分, 272.1cm 的再现像对应于相位变化为 knd (物体的 B 部分) 的部分, 272.6cm 的再现像对应于相位变化为 $2knd$ (物体的 D 部分) 的部分. 系统所用激光器为 50mW 染料激光器, 起始扫描波长为 602.0nm , 光波长间隔为 0.15544nm .

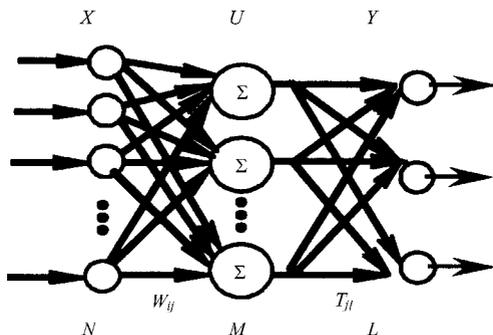


图 4 BP 神经网络模型结构图

2.4. 神经网络对物体的识别

利用这些包含有物体相位特征的数字再现像联合神经网络的方法,可以对这些三维物体进行识别.这里选用具有一个隐藏层的 BP 神经网络模型(如图 4),网络的输入是一个三维物体波长扫描数字全息再现像—— N 维矢量 $X, X = [x_1, \dots, x_i, \dots, x_N]$, 输出是对应于物体的编码一个 L 维矢量 $Y, Y = [y_1, \dots, y_i, \dots, y_L]$, 网络的隐层包括 M 个神经元, 其输出为 $U = [u_1, \dots, u_j, \dots, u_M]$. 在训练过程中, 神经元的输入输出值都是当前训练次数 t 与输入样本序列号 s 的函数. 第 i 个输入神经元至第 j 个隐层神经元之间的权重系数用 W_{ij} 表示, 第 j 个隐层神经元的阈值用 θ_j 表示. 第 j 个隐层神经元至第 l 个输出神经元的权重系数用 T_{jl} 表示, 第 l 个输出神经元的阈值用 φ_l 表示.

网络训练采用 BP 算法, 设输入样本序列号为 s , 对于第 t 次迭代

$$u_j^{(s)}(t) = f\left(\sum_{i=1}^N x_i^{(s)}(t)W_{ij}(t) - \theta_j(t)\right), \quad (9)$$

$$y_l^{(s)}(t) = f\left(\sum_{j=1}^M u_j^{(s)}(t)T_{jl}(t) - \varphi_l(t)\right), \quad (10)$$

其中 f 表示激活函数.

设样本 s 的期望值为 $d^{(s)} = (d_1^{(s)}, d_2^{(s)}, \dots, d_l^{(s)}, \dots, d_L^{(s)})$, 则误差函数 $e_l^{(s)}$ 为

$$e_l^{(s)}(t) = \frac{1}{2}(y_l^{(s)}(t) - d_l^{(s)})^2. \quad (11)$$

设包括阈值单元在内的网络权重参数的修改值为 $\Delta W_{ij}, \Delta T_{jl}$, 学习步长为 α , 则

$$\Delta W_{ij} = -\alpha \times \frac{\partial e^{(s)}}{\partial W_{ij}(t)} + \eta \times \Delta W_{ij}(t-1) \quad (12)$$

$$\Delta T_{jl} = -\alpha \times \frac{\partial e^{(s)}}{\partial T_{jl}(t)} + \eta \times \Delta T_{jl}(t-1), \quad (13)$$

其中 η 为动量因子 $0 \leq \eta < 1$, 动量项的引入能加快收敛速度, 使训练跳出局部极小. 根据上面 BP 算法中的误差传递计算方法, 可得到 $\Delta W_{ij}, \Delta T_{jl}$ 的具体表达式.

将物体的波长扫描数字全息再现像作为学习样本, 通过 (12) 和 (13) 式不断地修改网络的权重, 直到网络收敛. 训练好的 BP 网络就可以用来识别透明或半透明的 3D 物体.

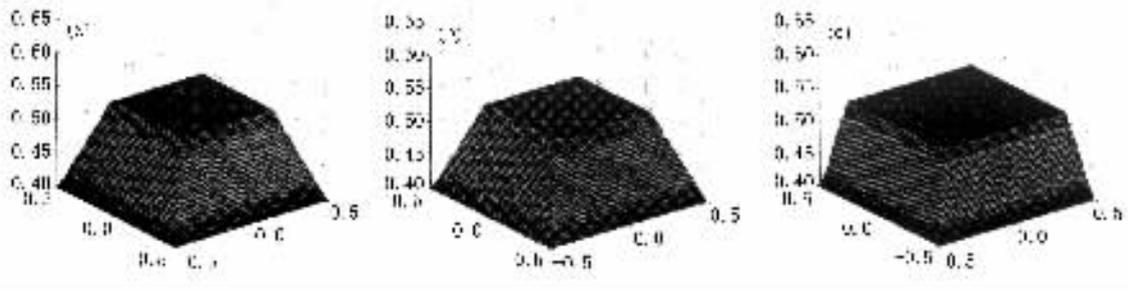


图 5 三个棱台的外观图

3. 计算机模拟

利用上述原理对三个物体(透明的棱台如图 5 所示)进行计算机模拟. 三个物体的上下底边长和高分别为 (a) 0.4mm, 1.0mm, 0.4mm; (b) 0.6mm, 1.0mm, 0.4mm; (c) 0.8mm, 1.0mm, 0.4mm. 图 6 为三个物体所对应的等光程线, 它们与物体的相位成正比. 取初始波长 $\lambda = 595\text{nm}$, 用 100 个间隔 $\Delta\lambda = 0.1\text{nm}$ 的不同波长记录和再现这三个物体. 选择 10

个再现位置, 图 7 是三个物体在其中 5 个不同位置的再现像. 里面白色的轮廓线与物体中的等光程线即等相位点对应, 可见不同的再现位置 Z_i 对应于物体不同的等相位点.

用这三个物体的再现像——相位特征图像训练第二部分中描述的 BP 神经网络, 这样得到的网络就包含了这三个物体的相位信息. 本实验中间层神经元的个数为 5 个, 输出层为 3 个, 选用双极的 sigmoid 函数作为激活函数. 用物体的相位特征图像作为输入, 用三位由 -1 或 1 组成的编码作为理想

输出(如-1 -1 1 代表物体 1,当训练物体的数量增加时
时可以扩展输出位数)训练网络,使得该网络能够

100% 的识别训练过的物体.

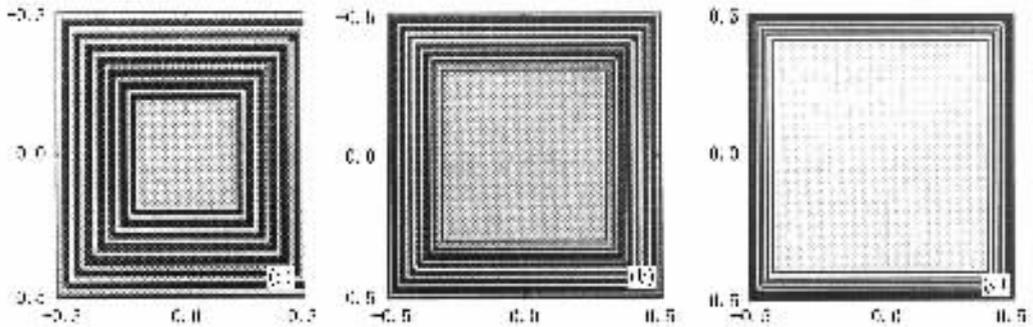


图 6 $\lambda = 595\text{nm}$ 时三个棱台的等光程线示意图

利用训练后的网络识别非学习训练物体,即棱台的底边边长或高度有一个小的变化量,模拟结果表明三维物体的尺度在一定的变化范围内,网络同

样可以得到 100% 的正确识别率.表明我们所训练的网络具有一定的推广性和容错能力,可以用来对透明或半透明三维物体进行小尺度不变性识别.

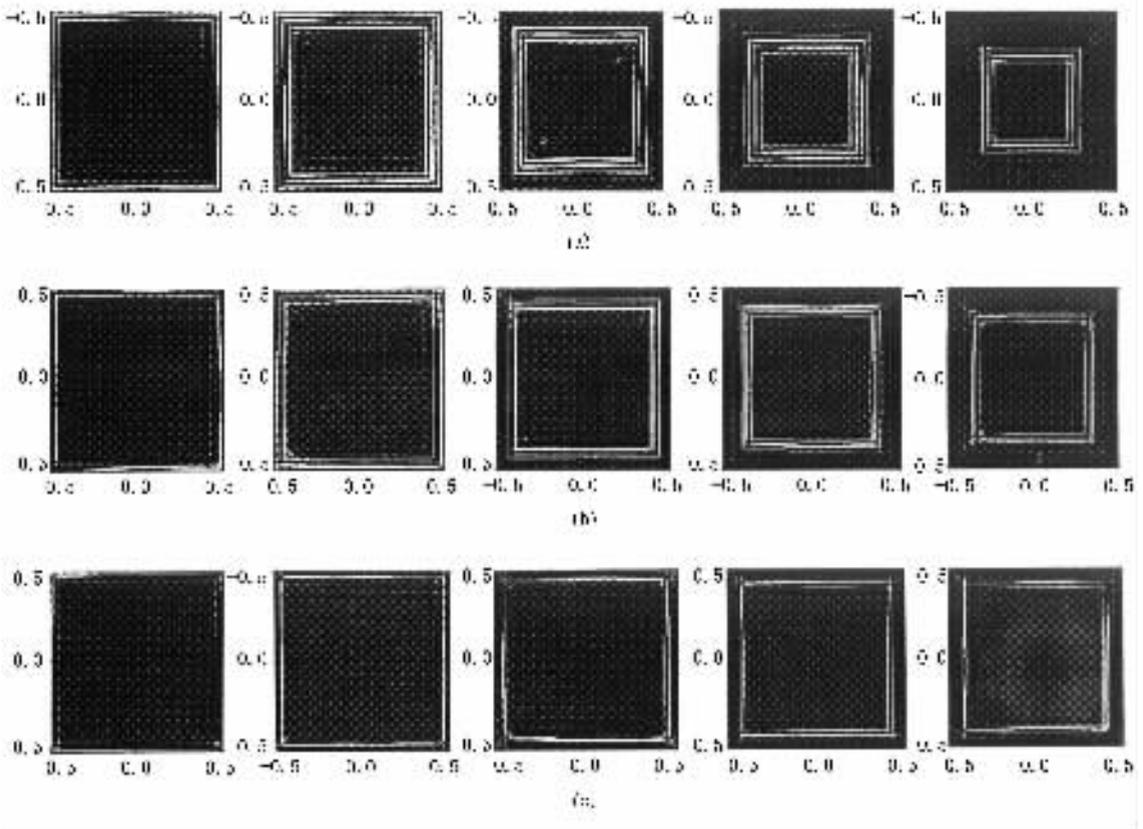


图 7 三个棱台在 5 个不同位置得到的再现图

表 1 所示是网络能够正确识别物体的畸变范围,以棱台 1 为例,当其他参数保持不变,只有上底边长变化时,上底边长可以在 0.396756—0.401336

间变化(即上底边长变化范围为 -0.811% 至 0.334%)网络仍然把它识别为棱台 1,超出这个范围则不会判断为棱台 1,以此类推.

表 1 可以正确识别的物体的取值范围

	只有上底边长变化	只有下底边长变化	只有高变化
棱台 1	[0.396756 0.401336]	[0.999170 1.000573]	[0.799756 0.800170]
棱台 2	[0.591396 0.601402]	[0.999332 1.000538]	[0.799690 0.800208]
棱台 3	[0.799397 0.800296]	[0.999632 1.000246]	[0.799746 0.800183]

4. 结论与展望

理论分析和计算机模拟表明,对于某些透明、半透明的 3D 物体,可以使用波长扫描数字全息技术和数字再现技术抽取其相位特征,然后利用这些相位特征训练神经网络模型,最后利用训练好的神经网络对 3D 物体进行识别.计算机模拟结果表明这个方法不仅是实用可行的,而且还具有一定的容错

性.这为透明或半透明三维物体进行小尺度不变性识别提供了一个新的方法.

同样利用波长扫描数字全息技术和数字再现技术可以得到一个三维物体不同角度(即旋转)的不同位置的再现像,将这些再现像作为神经网络(如 BP 模型)的学习样本训练网络,就可以使得网络具有一定的旋转不变性.因此,利用本文的方法可以实现具有一定'空间变'的三维物体的识别.

- [1] Paul J B and Ramesh C J 1985 *Computing Surveys* **17** 75
- [2] Dong J and Vijayant S 1997 *Computer-Aide Design* **29** 475
- [3] Shen J Y, Liu Y *et al* 1998 *Acta Phys. Sin.* **47** 1968 (in Chinese)
[申金媛、刘 等 1998 物理学报 **47** 1968]
- [4] Li C L and Hui K C 2000 *Computers & Graphics* **24** 569
- [5] Chen Z and Zhao R C 2003 *J. Image & Graphics* **8** A 993 (in Chinese)
[陈 柘、赵荣椿 2003 中国图象图形学报 **8A** 993]
- [6] Zeng J, Chen Z and Li M 2003 *Computer Application* **23** 9 (in Chinese)
[曾接贤、陈 震、黎 明 2003 计算机应用 **23** 9]
- [7] Wang Y L, Yao B L, Chen Y *et al* 2004 *Acta Phys. Sin.* **53** 66
(in Chinese)
[王英利、姚宝利、陈 懿等 2004 物理学报 **53** 66]
- [8] Zhang L, Li E P, Feng W *et al* 2005 *Acta Phys. Sin.* **54** 2038 (in Chinese)
[张 琳、李恩普、冯 伟等 2005 物理学报 **54** 2038]
- [9] Kim M K 1999 *Opt. Lett.* **24** 1693
- [10] Poggio T 1990 *Lett. Nat.* **343** 2673
- [11] Onwubolu G C 1999 *Journal of Intelligent Manufacturing* **10** 289
- [12] Onwubolu G C 1999 *Journal of Intelligent Manufacturing* **10** 251
- [13] Shen J Y, Liu Y and Zhang W W 2001 *Acta Optica Sinica* **21** 49 (in Chinese)
[申金媛、刘 、张文伟等 2001 光学学报 **21** 49]
- [14] Anghel M and Ben-Zion Y 2004 *Pure and Applied Geophysics* **161** 2023
- [15] Xiong T, Chang S J *et al* 2005 *Acta Phys. Sin.* **54** 1931 (in Chinese)
[熊 涛、常胜江等 2005 物理学报 **54** 1931]
- [16] Xiong T, Zhang B L, Chang S J *et al* 2005 *Acta Phys. Sin.* **54** 2435 (in Chinese)
[熊 涛、张便利、常胜江等 2005 物理学报 **54** 2435]
- [17] Zhang S, Liu H, Gao D and Du S 2003 *Chin. Phys.* **12** 594
- [18] Li K Gao Z and Chen T 2003 *Chin. Phys.* **12** 1213
- [19] Ge B Z, Luo W, Lu Q *et al* 2002 *J. Optoelectronics Laser* **13** 1289
(in Chinese)
[葛宝臻、罗文国、吕且妮等 2002 光电子·激光 **13** 1289]

Application of phase features in recognizing 3-D objects ^{*}

Shen Jin-Yuan^{1)†} Li Xian-Guo¹⁾ Chang Sheng-Jiang²⁾ Zhang Yan-Xin²⁾

¹⁾ Zhengzhou University, College of Information Engineering, Zhengzhou, 450052, China)

²⁾ Institute of Modern Optics, Nankai University, Tianjin 300071, China,

Key Laboratory of Opto-electronics Information Technical Science, CME)

(Received 7 December 2004 ; revised manuscript received 21 March 2005)

Abstract

A new approach based on phase features combined with neural network model is proposed for recognizing 3-D objects. The phase features of an object were extracted by wavelength-scanning digital holography and numerical reconstruction technique. A BP neural network with one hidden-layer trained by reconstructed images of three pyramids was used to recognize other pyramids with some variance, and the correct recognition rate of these pyramids is up to 100%. The simulation results demonstrate that the method is effective.

Keywords : phase feature, wavelength-scanning technique, digital holography, BP neural network model

PACC : 4230S, 4240, 4240F

^{*} Project supported by the National Natural Science Foundation of China(Grant No. 60277022), the Outstanding Youth Fund of Henan Province, The Project-sponsored by SRF for ROCS, SEM and The Project-sponsored by SRF for ROCS, SEM(Grant No.20030055022).

[†] E-mail jyshen@zzu.edu.cn, jyshen@nankai.edu.cn