

基于组合不变矩和神经网络的三维物体识别

徐 胜, 彭启琮

XU Sheng, PENG Qi-cong

电子科技大学 通信与信息工程学院, 成都 610054

School of Communication and Information Engineering, University of Electronic Science & Technology of China, Chengdu 610054, China

E-mail: xs@uestc.edu.cn

XU Sheng, PENG Qi-cong. Three Dimensional object recognition based combined moment invariants and neural network. Computer Engineering and Applications, 2008, 44(31): 78-80.

Abstract: In the 3D object recognition system, this paper presents to combine the lower order of Hu's moment invariants and affine moment invariants together as features of 3D objects, then these features are presented to the modified BP neural network for 3D object recognition. The theoretical and experimental analyses prove that using the combination of Hu's moment invariants and affine moment invariants as features to classify 3D objects can achieve better recognition performance than only using Hu's moment invariants. If the combination of Hu's moment invariants and affine moment invariants is further processed by principal components analysis, system recognition performance can be improved greatly and network training time can be reduced.

Key words: 3-D object recognition; Hu's moment invariants; affine moment invariants; BP neural network; principal components analysis

摘 要: 在三维物体识别系统中, 提出将三维物体的 Hu 不变矩和仿射不变矩两者的低阶矩组合作为三维物体的特征, 结合改进的 BP 神经网络应用于三维物体的分类识别。理论分析和仿真实验表明组合这两种矩特征进行物体识别, 性能优于单独使用 Hu 不变矩, 如果进一步对这两种组合的矩特征进行主成分分析处理, 可显著提高系统识别性能, 并减少网络的训练时间。

关键词: 三维物体识别; Hu 不变矩; 仿射不变矩; BP 神经网络; 主成分分析

DOI: 10.3778/j.issn.1002-8331.2008.31.022 **文章编号:** 1002-8331(2008)31-0078-03 **文献标识码:** A **中图分类号:** TP391

1 引言

计算机视觉研究的目的是使计算机可通过二维图像认知三维物体及周边环境信息的能力, 包括识别三维物体的几何形状、位置、运动、姿态等。出于军事侦察、城市规划、交通监控、图像检索及工业自动化等各个领域的大量现实应用的需要, 目前三维物体识别已成为一个较活跃的研究领域, 具有很大的实用价值和重要意义^[1]。

三维物体识别是模式识别的一个重要分支, 一般来说, 物体识别的方法分为三大类, 即基于特征的识别方法, 基于模型的识别方法和基于知识的识别方法。主要研究基于特征的三维物体识别方法, 步骤可分为: 图像预处理, 特征的提取和分类识别, 如图 1 所示。

主要指物体的几何尺寸, 包括物体的长、宽、高和面积等, 代数特征一般需要对物体进行数学描述后并计算, 是一些数学特性, 如代数不变性等^[2]。为了有效的进行物体识别, 所提取的三维物体特征必须要能够抵抗图像的各种变形所带来的差异, 即提取所谓的图像不变性特征。矩特征正是一种具有不变性的全局不变量, 以图像分布的各阶矩来描述图像灰度的统计特性, 对噪声不敏感, 并且不管物体是否封闭, 都能较好的识别物体^[3]。

在文献[4]的工作基础之上, 提出将三维物体的 Hu 不变矩和仿射不变矩两者的低阶矩组合作为三维物体的特征, 送入改进的 BP 神经网络进行训练, 将训练完成的 BP 神经网络用于三维物体识别, 取得了比较满意的测试结果。本文进一步提出对这两种组合的不变矩特征向量进行主成分分析处理, 显著地提高了训练后的 BP 神经网络的识别性能, 并减少了网络的训练时间。

2 矩的性质及不变矩特征

矩具有唯一性的特征, 若 $f(x, y)$ 分段连续, 且 (x, y) 在平面的有限部分中有非零值, 则所有各阶矩均存在, 并且矩序列 M_{pq} 唯一的被 $f(x, y)$ 所确定; 反之 M_{pq} 也唯一确定 $f(x, y)$ 。几何矩是矩

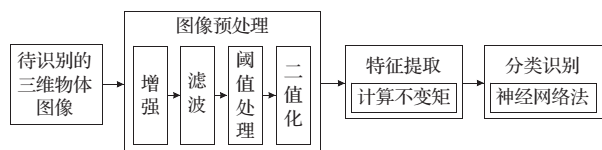


图 1 基于特征的三维物体识别系统框图

三维物体的特征分几何特征和代数特征两大类。几何特征

作者简介: 徐胜(1978-), 男, 博士生, 主要研究领域为图像识别、DSP 技术工程应用; 彭启琮(1946-), 男, 教授, 博士生导师, 主要研究领域为高速信号处理与实时信号处理、通信与信息系统中的信号处理、DSP 技术方面的研究。

收稿日期: 2007-12-03

修回日期: 2008-02-18

函数中最简单的一种, 其核为像素坐标的幂, 相对于含有复数核的矩, 几何矩的计算较容易完成。对于图像平面变换具有不变性的几何矩函数, 在目标鉴别和目标方位估计上有很多应用^[5]。

几何矩定义于基本集 $\{x^p y^q\}$ 。设 $f(x, y)$ 是一幅图像的二维连续函数, 则它的 $(p+q)$ 阶几何矩 m_{pq} 定义为:

$$m_{pq} = \int \int_{\xi} x^p y^q f(x, y) dx dy, p, q=0, 1, 2, \dots \quad (1)$$

式中, ξ 是图像函数 $f(x, y)$ 定义的像素空间区域。上式将函数 $f(x, y)$ 投影到单项式 $x^p y^q$ 上, 整个基本集 $\{x^p y^q\}$ 完备但不正交。

对于离散化为 $M \times N$ 的数字图像 $f(x, y)$, 其 $(p+q)$ 阶几何矩和中心矩为:

$$m_{pq} = \sum_x \sum_y f(x, y) x^p y^q, p, q=0, 1, 2, \dots$$

$$\mu_{pq} = \sum_x \sum_y f(x, y) (x-\bar{x})^p (y-\bar{y})^q, p, q=0, 1, 2, \dots \quad (2)$$

其中, $\bar{x}=m_{10}/m_{00}$, $\bar{y}=m_{01}/m_{00}$ 为图像重心。对于灰度图像, m_{00} 为图像质量, 对于二值图像, m_{00} 为图像面积。图像的几何矩和中心矩可用于描述图像的形状, 中心矩与图像的平移无关。

2.1 Hu 不变矩组

用零阶中心矩 μ_{00} 对其余各阶中心矩进行归一化可得到图像的归一化中心矩:

$$\eta_{pq} = \frac{\mu_{pq}}{\mu_{00}}, r=(p+q+2)/2, p+q=2, 3, 4, \dots \quad (3)$$

Hu M K^[6]利用二阶和三阶归一化中心矩的线形组合, 推导并证明了 7 个矩构成的矩组, 对于图像的平移、旋转和大小尺度变化均保持不变, 用于区域形状识别, 称为 Hu 矩组或不变矩组:

$$M_1 = \eta_{20} + \eta_{02}$$

$$M_2 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2$$

$$M_3 = (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2$$

$$M_4 = (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2$$

$$M_5 = (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + (\eta_{03} - 3\eta_{21})(\eta_{03} + \eta_{21})[(\eta_{03} + \eta_{21})^2 - 3(\eta_{12} + \eta_{30})^2] \quad (4)$$

$$M_6 = (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03})$$

$$M_7 = (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] - (3\eta_{12} - \eta_{30})(\eta_{03} + \eta_{21})[(\eta_{03} + \eta_{21})^2 - 3(\eta_{12} + \eta_{30})^2]$$

因为矩值较大, 一般利用取对数的方法进行数据压缩, 同时因为矩值可能为负, 所以实际采用的矩值需按照式(5)修正。经修正后的 Hu 不变矩值仍具有平移、旋转、大小尺度变化的不变性。

$$M_k = \log |M_k|, k=1, 2, 3, 4, 5, 6, 7 \quad (5)$$

2.2 仿射不变矩组

在实际的场景中, 距离、天气、照相设备以及视觉角度将影响物体的成像, 如果三维物体图像发生了变形(即仿射变换), 仅用 Hu 不变矩提取的目标特征将会不完全准确。Jan Flusser 等提取了仿射不变矩作为目标特征^[7], 在物体图像发生仿射变化下仍具有不变性。

对于二维空间中的任意一条曲线 $[x, y]$, 若 $[x', y']$ 为其仿射变换后的曲线, 则仿射变换为:

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = A[x, y]^T + B \quad (6)$$

仿射不变矩的一般形式是中心矩 μ_{pq} 的多项式除以 μ_{00}^r , 其

中 r 是某一合适的指数。Jan Flusser 等构造了 3 阶仿射不变矩, 应用在了字符识别和景像匹配等方面:

$$\begin{cases} I_1 = (\mu_{20}\mu_{02} - \mu_{11}^2) / \mu_{00}^4 \\ I_2 = (\mu_{30}^2\mu_{03}^2 - 6\mu_{30}\mu_{21}\mu_{12}\mu_{03} + 4\mu_{30}\mu_{12}^3 + 4\mu_{21}^3\mu_{03} - 3\mu_{21}^2\mu_{12}^2) / \mu_{00}^{10} \\ I_3 = (\mu_{20}(\mu_{21}\mu_{03} - \mu_{12}^2) - \mu_{11}(\mu_{30}\mu_{03} - \mu_{21}\mu_{12}) + \mu_{02}(\mu_{30}\mu_{12} - \mu_{21}^2)) / \mu_{00}^7 \end{cases} \quad (7)$$

3 神经网络分类识别方法

根据特征进行三维物体的识别从模式识别角度看是一个分类问题, 从而使神经网络方法应用于三维物体特征识别领域成为可能。与传统的方法相比, 神经网络方法具有可并行运算、处理速度高和鲁棒性好等优点。Irie 等已证明不超过 3 层的前馈神经网络能以任何精度逼近任何连续函数, 解决许多分类问题^[8], 所以本文采用 3 层的 BP 神经网络进行分类识别。

对于确定网络的隐含层和输出层的阈值及权值的初值, 本文选用 Nguyen-Widrow 初始化算法, 该算法可使各层神经元的有效区域基本均匀的分布到输入空间。相对于传统的取 $(-1, +1)$ 之间随机值的方法, 该算法浪费的神经元更少, 且输入空间的每个区域基本都有神经元存在, 从而使网络训练更快^[9]。

通常选用梯度下降算法进行网络训练, 该算法在最初几步下降较快, 但随着接近最优值, 由于梯度趋于零, 致使均方误差函数下降缓慢; 而牛顿法则可在最优值附近产生一个理想的搜索方向, 所以本文选用梯度下降法和牛顿算法相结合的 Levenberg-Marquardt 算法改进训练过程, 且该算法无需计算 Hessian 矩阵, 减少了计算复杂性, 比一般的梯度下降算法缩短训练时间 1~2 个数量级^[10]。

4 计算机仿真实验

将上述两种不变矩理论和神经网络方法融合在一起, 用于三维物体的分类识别。使用 COIL(Columbia Object Image Library)三维物体图像数据库, 测试了该方法的识别性能。COIL 三维物体图像数据库包含 100 个物体, 每个物体有 72 个不同视角的真彩图像(分辨率为 128×128), 共计 7 200 幅。该数据库的生成是将物体置于一个旋转台面, 每 5°拍摄一个视角的图像。本文任选择了 5 个三维物体, 共计 360 幅图像进行实验, 见图 2(a)~(e)。图 3(a)~(e)是各物体旋转了 45°视角的图像, 其余视角图像略过。

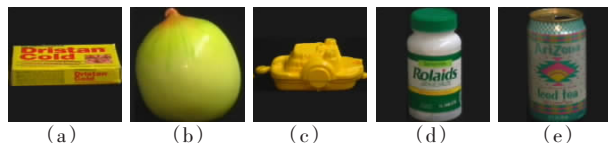


图 2 本文选取的 5 个三维物体

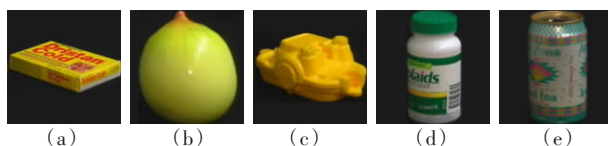


图 3 旋转 45°视角的图像

实验选择了上述 5 个三维物体每隔 10°视角的图像作为训练样本, 即 0°, 10°, ..., 350°共 180 幅; 为了全面测试本方法的识别性能, 将其余 180 幅视角为 5°, 15°, ..., 355°的图像和训练样本都作为测试图像。实验按照图 1 所示的基于特征的三维

物体识别系统步骤进行:

(1) 首先将 180 幅训练样本图像进行预处理, 包括图像增强、滤波、阈值处理、二值化, 便于后续计算两种不变矩值。这里选用的阈值处理方法为最大方差阈值。预处理后得到的二值化图像仍以 45° 视角为例, 如图 4 所示, 其余经过预处理的视角图像略过;

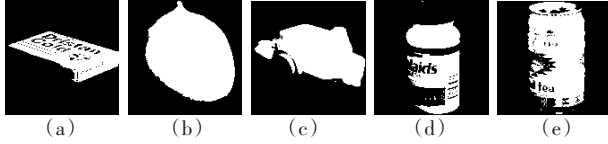


图 4 旋转 45° 视角的二值化图像

(2) 根据式(2)计算出预处理后的 180 幅训练样本图像的中心矩, 考虑到高阶矩计算量过大, 并且经过实验发现高阶矩对识别性能贡献也较小, 所以仅选用前 4 个低阶的修正 Hu 不变矩 $M_1 \sim M_4$ 和第 1 个低阶的仿射不变矩 I_1 作为物体的特征向量, $M_1 \sim M_4$ 可根据步骤(3)~(5)计算得到, I_1 根据式(7)计算得到;

(3) 组合 $M_1 \sim M_4$ 和 I_1 为 5 个分量的特征向量, 作为 BP 神经网络的输入。因此, BP 网络的输入层节点选择为 5; 要求 BP 网络识别并输出这 5 个物体, 因此输出层节点选择为 5;

(4) 为了对比考察组合特征向量的性能, 只选用前 5 个修正 Hu 不变矩 $M_1 \sim M_5$, 构成另外一类同样为 5 个分量的特征向量作为 BP 网络的输入;

(5) 经过多次实验确定, BP 网络的隐层节点数选择为 9, 可以提高识别率, 并且降低计算的复杂性;

(6) 为了便于 Levenberg-Marquardt 训练算法的工作, BP 网络的隐层传输函数选择 $TF_1(x) = 1/(1+e^{-x})$; 考虑到 BP 网络输出为 0 或 1 的判决, 范围为, 所以输出层传输函数选择 $TF_2(x) = 2/(1+e^{-2x}) - 1$;

(7) 设定 BP 网络的训练在实际输出值和期望输出值的均方误差小于 $1e-6$ 时结束, 使用 Nguyen-Widrow 初始化算法对 BP 网络的的权和阈值初始化后, 开始 BP 网络训练, 采用 Levenberg-Marquardt 算法改进权和阈值, 每次训练让 BP 网络迭代 300 次;

(8) BP 网络训练完毕后, 对其输入 360 个测试图像 (包括了训练样本图像), 考察其识别性能, 实验结束。

BP 神经网络训练容易陷入局部极值, 使得网络识别性能下降。为使实验结果更具全面性, 本文重复步骤(1)~(8)完成了 100 次独立实验, 得到了由组合的特征向量 $M_1 \sim M_4$ 和 I_1 训练的 100 个 BP 网络, 测试并统计了 BP 网络的平均误识率为 0.88%, 且得到了 3 个识别率为 100% 的网络, 见图 5 误识率曲线 A; 同时对比经过步骤(4)得到的另外 100 个只通过前 5 个修正 Hu 不变矩 $M_1 \sim M_5$ 训练的 BP 网络, 测试并统计了其网络的平均误识率为 1.37%, 最低误识率为 0.27%, 没有得到 100% 识别率的 BP 网络, 见图 5 中误识率曲线 B。对比误识率曲线 A 和 B, 可见仅添加 1 个低阶仿射不变矩作为特征向量训练 BP 网络, 就可以得到更优的识别性能。其中, 曲线 A: 修正 Hu 不变矩 $M_1 \sim M_4$ 和仿射不变矩 I_1 的组合作为特征; 曲线 B: 修正 Hu 不变矩 $M_1 \sim M_5$ 作为特征; 曲线 C: 修正 Hu 不变矩 $M_1 \sim M_4$ 和仿射不变矩 I_1 的组合 (经过 PCA 处理后) 作为特征。

本文进一步对 180 幅训练样本图像的组合特征向量 $M_1 \sim M_4$ 和 I_1 进行主成分分析 (PCA) 处理, 去除对整体方差贡献度

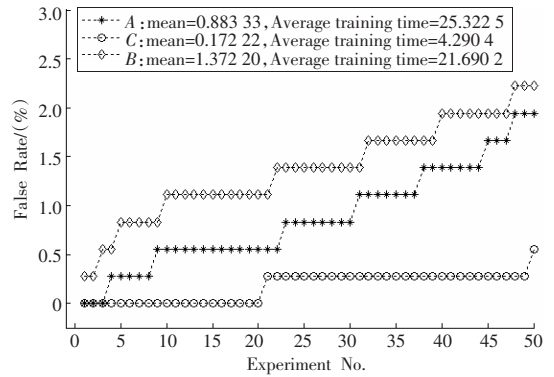


图 5 3 类特征的误识率比较

低于 1% 的分量, 作为 BP 网络新的输入特征向量。再次按照步骤(5)~(8)重新进行 100 次独立实验, 得到并测试了 100 个新的 BP 神经网络, 结果为平均误识率降为 0.17%, 并且得到了 20 个识别率为 100% 的 BP 网络, 见图 5 中误识率曲线 C。对比没有经过 PCA 处理的曲线 A, 平均误识率下降约 4/5, 平均训练时间减少约 5/6。

5 结语

通过理论分析和仿真实验表明, 提取三维物体的低阶修正 Hu 不变矩和仿射不变矩, 组合作为物体特征向量, 训练改进的 BP 神经网络, 可以达到比较满意的识别性能, 优于仅使用修正 Hu 不变矩作为特征向量; 若进一步采用主成分分析方法将组合的不变矩特征向量进行正交化和数据压缩, 作为 BP 神经网络新的输入特征向量进行训练, 更能显著提高系统的识别性能, 并大为减少 BP 网络训练时间, 证明本文方法应用在三维物体的识别领域是可行的。基于本文的工作, 可进一步研究: (1) 寻找可表达三维物体局部特征的方法来弥补矩方法不能检测物体的局部特征的缺点; (2) 进一步提取物体的颜色作为特征向量训练神经网络, 可以期望得到更高识别性能的网络。

参考文献:

- [1] 李庆, 周曼丽, 柳健. 三维物体识别研究进展[J]. 中国图象图形学报, 2000, 5(12): 985-993.
- [2] 周振环, 李言俊. 三维目标识别原理及关键技术[J]. 河南师范大学学报: 自然科学版, 2001, 29(2): 22-25.
- [3] 朱志刚. 数字图像处理[M]. 北京: 电子工业出版社, 2002.
- [4] 徐胜, 彭启琮. 基于不变矩和神经网络的三维物体识别[J]. 西安电子科技大学学报, 2007, 34(增刊): 189-192.
- [5] 王耀明. 图像的矩函数[M]. 上海: 华东理工大学出版社, 2002.
- [6] Hu M K. Visual pattern recognition by moment invariants[J]. IEEE Transactions on Information Theory, 1962, 8(2): 179-187.
- [7] Flusser J, Suk T. Pattern recognition by affine moment invariants[J]. Pattern Recognition, 1993, 26(1): 167-174.
- [8] Irie B, Miyake S. Capabilities of three-layered perceptrons[C]//Proceedings of the IEEE International Conference on Neural Networks, 1988: 641-648.
- [9] Nguyen D, Widrow B. Improving the learning speed of 2-layer neural networks by choosing initial values of the adaptive weights[C]//International Joint Conference on Neural Networks, 1990: 21-26.
- [10] Hagan M T, Menhaj M B. Training feedforward networks with the marquardt algorithm[J]. IEEE Transactions on Neural Networks, 1994, 5(6): 989-993.