

# 基于 HMM 和微粒群优化算法的表情识别

陈燕龙<sup>1</sup>, 钟碧良<sup>2</sup>

(1. 广州航海高等专科学校基础部, 广州 510725; 2. 广州航海高等专科学校计算机与信息工程系, 广州 510725)

**摘要:** 提出基于微粒群优化算法(PSO)的隐马尔科夫模型(HMM)训练算法, 分别用 PSO 和量子微粒群优化算法进行 HMM 的参数估计, 以提高 HMM 的性能。将改进的 HMM 算法应用于人脸表情识别, 采用离散余弦变换提取表情特征向量。实验结果表明, 该算法能有效提高表情识别率, 解决 HMM 的参数估计问题。

**关键词:** 隐马尔科夫模型; 微粒群优化算法; 离散余弦变换; 表情识别

## Facial Expression Recognition Based on HMM and PSO

CHEN Yan-long<sup>1</sup>, ZHONG Bi-liang<sup>2</sup>

(1. Department of Basic Courses, Guangzhou Maritime College, Guangzhou 510725;

2. Department of Computer and Information Technology, Guangzhou Maritime College, Guangzhou 510725)

**【Abstract】** This paper presents an approach for Hidden Markov Model(HMM) training based on Particle Swarm Optimization(PSO). HMM parameter adjustment rules are induced by using PSO and Quantum-behaved Particle Swarm Optimization(QPSO) separately. Discriminative information contained in the training data is used to improve the performance on HMM effectively and the method is used in facial expression recognition. Facial expression feature vectors are extracted by using Discrete Cosine Transform(DCT). Experimental results show that the new method provides satisfactory recognition performance and is powerful for HMM parameter estimation.

**【Key words】** Hidden Markov Model(HMM); Particle Swarm Optimization(PSO); Discrete Cosine Transform(DCT); facial expression recognition

### 1 概述

隐马尔科夫模型(Hidden Markov Model, HMM)是一个统计模型, 具有十分丰富健壮的数学结构, 适用于动态过程时间序列建模并且具有强大的时序模式分类能力, 理论上可以处理任意长度的时序, 应用范围非常广泛。目前, HMM 已经成功应用于语音识别等领域。

通常采用 Baum-Welch 算法训练 HMM, 该算法实际上是一种极大似然参数估计法, 因为单个模型独立进化, 逐渐收敛, 所以通常只能得到局部最优, 找到最优 HMM 的概率不高。目前国内外的研究集中在使用遗传算法和神经网络等方法进行优化, 此外, 人们还提出了其他一些训练算法, 如最大互信息法<sup>[1]</sup>、最大模型距离法<sup>[2]</sup>, 这些方法各有优缺点。本文结合 HMM 的结构特点, 提出了一种结合多个次优 HMM 的量子微粒群优化算法 (Quantum-behaved Particle Swarm Optimization, QPSO), 利用 QPSO 优秀的全局收敛性提高 HMM 寻优的概率, 并成功地将其应用于人脸表情识别中。

### 2 HMM

HMM 是一种用参数表示的、用于描述随机过程统计特性的概率模型。它是一个双重随机过程: 一个随机过程是马尔科夫链, 由初始状态分布概率和转移概率矩阵描述, 是一个隐藏层, 输出为状态序列; 另一个随机过程描述状态和观察值之间的统计关系, 是一个可观察层, 输出为观察序列, 但不一定是马尔科夫过程。

设有观察序列  $O = (o_1, o_2, \dots, o_T)$  和状态集  $S = (s_1, s_2, \dots, s_N)$ , HMM 可以表示为  $\lambda = (N, M, A, B, \pi)$ , 或简写为三元组  $\lambda = (A, B, \pi)$ , 其中,  $A = (a_{ij}) (1 \leq i, j \leq N)$  为状态转移概率

矩阵;  $B = [b_j(k)] (1 \leq j \leq N, 1 \leq k \leq M)$  为观察符号概率分布;  $\pi = (\pi_i) (1 \leq i \leq N)$  为初始状态概率向量;  $N$  为模型中 Markov 链状态数;  $M$  为每个状态对应的可能的观察数。人脸表情识别系统中的状态  $S = (s_1, s_2, \dots, s_N)$  是表情。人脸表情除中性外可以归纳为 6 种, 分别是高兴(happy)、生气(angry)、惊讶(surprise)、厌恶(disgust)、恐惧(fear)和悲哀(sad)。为每种表情设计一个 HMM, 每个 HMM 是一个带返回的从左到右模型, 6 个这样的 HMM 构成一个人脸表情分类器。

### 3 微粒群优化算法与量子微粒群优化算法

#### 3.1 标准微粒群优化算法

微粒群优化算法(Particle Swarm Optimization, PSO)是一种进化计算(evolutionary computation)技术, 由文献[3]首次提出。与其他进化类算法类似, PSO 算法也采用群体和进化的概念, 根据个体的适应值大小进行操作。不同的是, PSO 算法将每个个体看作是  $n$  维搜索空间中一个没有重量和体积的微粒, 并在搜索空间中以一定的速度飞行。该速度根据个体飞行经验和群体飞行经验进行动态调整, 最终使所有微粒趋近全局最好位置。

在标准 PSO 算法中, 在一个包含  $M$  个粒子的群体中, 每个微粒的维数为  $n$ , 其中,  $X_i = (x_{i1}, x_{i2}, \dots, x_{in})$  为微粒  $i$  的当前位置;  $V_i = (v_{i1}, v_{i2}, \dots, v_{in})$  为微粒  $i$  的当前速度向量;

**基金项目:** 广东省自然科学基金资助项目(04300614)

**作者简介:** 陈燕龙(1962-), 男, 讲师、硕士, 主研方向: 模式识别, 图像处理; 钟碧良, 教授、博士

**收稿日期:** 2007-12-15 **E-mail:** cylgz@126.com

$P_i = (P_{i1}, P_{i2}, \dots, P_{in})$  是微粒  $i$  的个体最好位置； $P_g$  为所有微粒的当前最好位置。标准PSO算法的进化公式如下：

$$V_i(t) = \omega(t)V_i(t-1) + \varphi_1\mu_1(P_i(t-1) - X_i(t-1)) + \varphi_2\mu_2(P_g(t-1) - X_i(t-1)) \quad (1)$$

$$X_i(t+1) = X_i(t) + V_i(t) \quad (2)$$

其中， $\omega(t)$  是惯性因子，一般可以设为随进化而线性减少； $t$  为当前进化代数； $\varphi_i$  是加速因子，通常在 0~2 间选取； $\mu_i \sim U(0,1)$  是 2 个相互独立的随机数。为了减少在进化过程中微粒离开搜索空间的可能性，限定  $V_i \in [V_{\min}, V_{\max}]$ 。

为了保证算法的收敛性，每个微粒必须收敛于各自的  $P_i$  点，这是由微粒的追从性和微粒群的集聚性决定的。从动力学的角度看，PSO算法中微粒的收敛过程是以  $P_i$  点为吸引子，随着速度的减小而不断接近  $P_i$  点。因此在整个过程中， $P_i$  点处实际存在某种形式的吸引势能场吸引着微粒，这正是微粒群能保持集聚性的原因。但由于在标准PSO微粒群系统中，微粒的收敛以轨道的形式实现，且微粒的速度总是有限的，因此在搜索过程中，微粒的搜索空间是一个有限的区域，不能覆盖整个可行的空间。所以，标准PSO算法最大的缺点是不能保证以概率 1 收敛到全局最优解，甚至无法收敛到局部最优解，许多学者为此对其进行了改进。

### 3.2 量子微粒群优化算法

QPSO<sup>[4-5]</sup> 是从量子力学的角度对标准PSO算法进行继承和推广的。在QPSO中，微粒群按如下公式移动：

$$mbest = \frac{1}{M} \sum_{i=1}^M P_i = \left( \frac{1}{M} \sum_{i=1}^M P_{i1}, \frac{1}{M} \sum_{i=1}^M P_{i2}, \dots, \frac{1}{M} \sum_{i=1}^M P_{id} \right) \quad (3)$$

$$P_{id} = \varphi P_{id} + (1-\varphi) P_{gd} \quad (4)$$

$$x_{id} = P_{id} \pm \alpha |mbest - x_{id}| \ln \frac{1}{u} \quad (5)$$

其中， $\varphi, u = rand$ ； $mbest$  是微粒群的平均最优位置； $p_{id}$  为  $P_{id}$  和  $P_{gd}$  之间的随机点； $\alpha$  为QPSO的收缩扩张系数，它是QPSO算法收敛性的重要参数：

$$\alpha = 0.5 \times (MAXITER - T) / MAXITER + 0.5 \quad (6)$$

其中， $MAXITER$  是迭代的最大次数； $\alpha$  的取值视具体情况而定； $T$  为当前迭代次数。

PSO 和 QPSO 的区别在于进化方法不同，即更新微粒位置的方法不同。在 PSO 算法中，微粒必须在一个有限的搜索范围内，以确保微粒群的集聚性使算法收敛于一个最优点或局部最优点。而在 QPSO 算法中，微粒能以某一确定的概率出现在整个可行搜索空间中的任意一个位置。而这个位置可能比当前群体中的个体极值具有更好的适应值，所以，QPSO 能较好地克服 PSO 在收敛性能方面的不足。

## 4 基于改进 HMM 算法的人脸表情识别

### 4.1 特征提取

特征提取是关键的一环，其目的是获取一组少而精、分类错误概率小的分类特征。离散余弦变换(Discrete Cosine Transform, DCT)的变换系数具有较好的变换系数特征。DCT 的变换核为实数的余弦函数，其计算速度比变换核为复指数的 DFT 快得多。而且 DCT 是 K-L 变换的很好的近似，而 K-L 变换是基于不同性能准则的一个最优变换<sup>[6]</sup>。数据压缩是 DCT 的一个主要特性；又因为 DCT 与离散傅里叶变换很相近，所以可以进行有效的计算。基于这 2 个特性，可用它作为人脸表情的特征。

选用 DCT 变换系数作为 HMM 的观察向量，HMM 观察

序列的获取采用遍历方法。在人脸图像平面上用一个像素为  $X \times L$  的采样窗，从上向下滑动，相邻窗口在垂直方向有  $M$  行。如果直接取采样窗内图像块的像素点灰度值构造观察向量，则存在下列问题：(1)观察向量维数( $X \times L$ )太大，计算量太大；(2)像素灰度值对光照条件变化、视点变迁以及其他噪声干扰太敏感。由于二维 DCT 的结果是能量分布向低频成分集中，因此取每一个图像块二维 DCT 左上角的低频系数组成观察向量。由采样图像块的二维 DCT 低频系数构造观察向量可以降低对噪声和光照变化的敏感度以及观察向量的维数。本文取  $L=16, M=12$ ，取每个采样图像块的 6 个二维 DCT 系数组成观察向量。

### 4.2 基于 QPSO 的 HMM 参数估计算法

首先初始化 HMM 参数，采用经典的 Baum-Welch 算法估计 HMM 参数，并把估计出的每个次优 HMM 作为 QPSO 算法的一个微粒。QPSO 的目标函数可依据最优模型的度量指标来选择。设  $\mathcal{R}$  为 HMM 模型集， $\mathcal{R} = (\lambda_1, \lambda_2, \dots, \lambda_v)$ ， $V$  为 HMM 个数，则训练样本集为

$$O = (O_1^1, O_2^1, \dots, O_{K_1}^1, O_1^2, O_2^2, \dots, O_{K_2}^2, \dots, O_1^v, O_2^v, \dots, O_{K_v}^v)$$

其中， $O_k^v$  为模型  $\lambda_v$  的第  $k$  个训练样本； $K_v$  为模型  $\lambda_v$  的训练样本数，且有  $O_k^v = (O_{k1}^v, O_{k2}^v, \dots, O_{kT_k}^v)$ ； $T_k^v$  为观察向量序列  $O_k^v$  的长度。本文的目标函数采用最大互信息准则<sup>[11]</sup>：

$$M(\mathcal{R}) = \ln P(\mathcal{R} | O) = \sum_{v=1}^V \sum_{k=1}^{K_v} \ln P(\lambda_v | O_k^v) = \sum_{v=1}^V \sum_{k=1}^{K_v} \ln \left( \frac{P(\lambda_v) P(O_k^v | \lambda_v)}{\sum_{u=1}^V P(\lambda_u) P(O_k^v | \lambda_u)} \right) \quad (7)$$

QPSO 算法流程如下：(1)迭代次数  $t=0$ ，对种群中每个粒子的位置向量进行初始化；(2)根据目标函数计算每个粒子的目标函数值；(3)更新每个粒子的局部最优位置  $P_{ij}$ ；(4)更新全局最优位置  $P_{gj}$ ；(5)根据式(3)计算  $mbest$ ；(6)根据式(4)计算每个粒子随机点  $p_{id}$ ；(7)根据式(5)更新每个粒子的位置；(8)判断参数估计是否达到预定迭代步数或预定精度，如条件满足，终止；否则， $t=t+1$ ，返回(2)。

## 5 实验

本文采用日本 Kyushu 大学的 JAFFE 表情库<sup>[7]</sup>进行实验，该库由 10 人的 210 幅图像组成，每人展示 7 种表情。每个人的每种表情都由 3 幅  $256 \times 256$  像素的图像组成，选取 60 幅作为训练样本，剩下的作为测试样本。实验中的部分表情图像如图 1 所示。

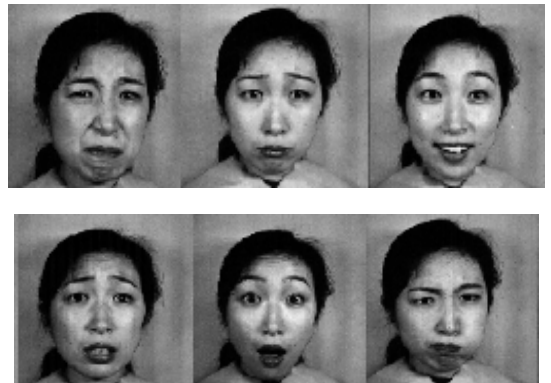


图 1 JAFFE 表情库部分表情图像

为了便于比较，分别采用 HMM, HMM+PSO 和 HMM+QPSO 3 种方法进行识别，最大迭代次数取为 100，QPSO 的参数  $\alpha = 0.5 \times (100 - T) / 100 + 0.5$ 。仿真结果见表 1。

表 1 3 种方法的表情识别率 (%)

识别方法	表情识别率						平均识别率
	悲伤	生气	高兴	惊讶	恐惧	厌恶	
HMM	68.2	84.3	69.2	93.1	62.3	89.4	77.6
HMM+PSO	74.3	86.9	73.1	95.2	71.4	90.7	82.0
HMM+QPSO	77.0	88.4	80.7	95.8	73.9	91.2	84.5

由表 1 可以看出,与采用 Baum-Welch 算法训练的 HMM 相比,采用微粒群优化算法进行训练可以有效提高参数的寻优效率和识别率;QPSO 的识别率高于 PSO,说明 QPSO 算法的寻优效果优于标准的 PSO 算法。

识别时间与取得最优值的迭代次数见表 2。实验结果表明,基于 QPSO 的 HMM 算法识别时间最短,几乎是传统 HMM 算法的一半;QPSO 寻优的迭代次数少、效率高、稳定性好。实验中还发现,识别时间与人脸图在训练库中的位置无关,但与训练库的大小有直接的关系:训练库越大,耗时越多。

表 2 识别时间与取得最优值的迭代次数

识别方法	识别时间/s	平均迭代次数	最大迭代次数
HMM	7.5	12.7	38
HMM+PSO	5.2	9.2	23
HMM+QPSO	4.1	5.2	8

图 2 是本文方法与其他表情识别分类器<sup>[8]</sup>在 JAFFE 表情库上的识别率比较结果。其中,RBFNN 为径向基函数神经网络分类器;SVM 为支持向量机分类器。从图 2 中可见,本文方法的识别率略高于上述方法。

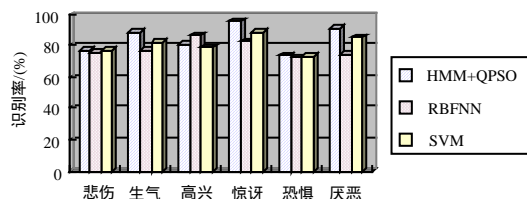


图 2 几种不同分类方法的识别率

## 6 结束语

实验结果表明,基于微粒群的 HMM 训练算法具有优良

的性能,识别率、鲁棒性和实时性均优于传统方法。用于训练 HMM 的 QPSO 算法的参数设置少(仅一个)、操作简单、收敛速度快。相比之下,传统 PSO 算法的效果取决于参数  $\omega(t), \varphi_i, V_{\min}, V_{\max}$  的取值,算法的运行具有一定的盲目性,但与 Baum-Welch 算法相比具有明显的优势。与其他几种分类器的比较显示本文算法具有较高的识别率。进一步改进算法,增强其自适应性,并将其应用在大规模样本的训练中,是下一步研究的方向。

## 参考文献

- [1] Assaf B Y, Brushtein D. A Discriminative Training Algorithm for Hidden Markov Models[J]. IEEE Trans. on Speech Audio Processing, 2004, 12(3): 204-217.
- [2] Kass M, Witkinm A, Terzopoulos D. Snakes: Active Contour Models[C]//Proc. of the 1st International Conference on Computer Vision. London, U.K.: [s. n.], 1987: 259-269.
- [3] Kennedy J, Eberhart R C. Particle Swarm Optimization[C]//Proc. of IEEE Int'l Conf. on Neural Network. Perth, Australia: IEEE Press, 1995: 1942-1948.
- [4] Sun Jun, Xu Wenbo. A Global Search Strategy of Quantum-behaved Particle Swarm Optimization[C]//Proc. of IEEE Conference on Cybernetics and Intelligent Systems. Singapore: [s. n.], 2004.
- [5] Sun Jun, Feng Bin, Xu Wenbo. Particle Swarm Optimization with Particles Having Quantum Behavior[C]//Proc. of Congress on Evolutionary Computation. Portland, USA: [s. n.], 2004: 325-331.
- [6] Pan Z, Adams R, Bolouri H. Dimensionality Reduction of Face Images Using Discrete Cosine Transforms for Recognition[C]//Proc. of IEEE Conf. on Computer Vision and Pattern Recognition. Hilton Head, USA: [s. n.], 2000.
- [7] Michael J L, Budynek J, Akamatsu S. Automatic Classification of Single Facial Images[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1999, 21(12): 1357-1362.
- [8] 王志良, 刘芳, 王莉. 基于计算机视觉的表情识别技术综述[J]. 计算机工程, 2006, 32(11): 231-233.

(上接第 176 页)

## 5 结束语

实验中发现蚁群等仿生算法对平衡无线传感器网络的能量消耗有明显改善。下一步将继续深入研究各种仿生算法在无线传感器网络中的应用,更好地改进协议的性能。

## 参考文献

- [1] Su Hang, Zhang Xi. Optimal Transmission Range for Cluster-based Wireless Sensor Networks with Mixed Communication Modes[C]//Proceedings of the 2006 International Symposium on a World of Wireless, Mobile and Multimedia Networks. [S. l.]: IEEE Press, 2006.
- [2] Heinzelman W, Chandrakasan A, Balakrishnan H. Energy-efficient Communication Protocol for Wireless Microsensor Networks[C]//Proc. of the 33rd Annual Hawaii Int'l Conf. on System Sciences.

[S. l.]: IEEE Computer Society, 2000.

- [3] Younis O, Fahmy S. Heed: A Hybrid, Energy-efficient, Distributed Clustering Approach for Ad-hoc Sensor Networks[J]. IEEE Trans. on Mobile Computing, 2004, 3(4): 660-669.
- [4] Heinzelman W. Application-specific Protocol Architectures for Wireless Networks[D]. Boston, USA: Massachusetts Institute of Technology, 2000.
- [5] 沈波, 张世永, 钟亦平. 无线传感器网络分簇路由协议[J]. 软件学报, 2006, 17(7): 1588-1600.
- [6] Dorigo M, Maniezzo V, Colomi A. The Ant System: Optimization by a Colony of Cooperating Agents[J]. IEEE Transactions on Systems, Man & Cybernetics B, 1996, 26(2): 29-41.
- [7] 卿利, 朱清新, 王明文. 系统异构传感器网络的分布式能量有效成簇算法[J]. 软件学报, 2006, 17(3): 481-489.