

基于 SVM 信息融合的图像识别与并行实现

刘松

LIU Song

中国矿业大学 煤炭资源与安全开采国家重点实验室,北京 100083

National State Key Laboratory of Coal Resources and Safe Mining, China University of Mining and Technology, Beijing 100083, China

E-mail: bigvision2008@126.com

LIU Song. Image recognition based on SVM information fusion and DSP parallel realization. *Computer Engineering and Applications*, 2009, 45(33): 168-170.

Abstract: The paper proposes a multi-feature and multi-classifier fusion algorithm for image recognition called Support Vector Machine (SVM) information fusion. First, the paper applies three different algorithms to the feature extraction: PCA, SVM, LDA. The outputs of facial expression classifier based on three expression representations are input to SVM information fusion to get facial expression recognition. For hardware realization is few, this paper uses the DAVINCI DSP serial productions of TI company to realize the SVM information fusion algorithm scheme based on parallel structure. It discusses the process of algorithm realization by DSP. The actual results prove that SVM information fusion algorithm for expression recognition is effective and can be realized by DSPs.

Key words: Support Vector Machine (SVM); image fusion; Principal Component Analysis (PCA); Fisher Linear Discriminant Analysis (Fisher LDA); Digital Signal Processing (DSP)

摘要: 提出用支持向量机(SVM)融合三种基于不同特征表示的表情识别方法进行表情识别,即 PCA 表情表示、SVM 表情表示和 FLD 表情表示。在用 SVM 进行特征提取时,提出一种高效的方案选择投影轴。在提取各种特征表示后,对每一种表情特征用 1 阶最近邻分类器进行初步识别,最后用支持向量机融合这些分类结果进行表情的最终识别。并且针对目前还没有硬件实现情况,提出用 TI 公司的达芬奇系列的 DSP 芯片构建并行系统来实现 SVM 融合算法方案,讨论并优化 DSP 实现算法的过程,通过实验的结果表明,提出的方案是有效的。

关键词: 支持向量机; 图像融合; 主元分析; Fisher 鉴别分析; 数字信号处理(DSP)

DOI: 10.3778/j.issn.1002-8331.2009.33.055 **文章编号:** 1002-8331(2009)33-0168-03 **文献标识码:** A **中图分类号:** TP391

1 引言

复杂目标的图像识别,比如矿质主成分识别、弹道曲面识别、面部表情识别等,仅仅采用单一的图像特征提取和图像识别方法,很难到达很好的识别效果和鲁棒性,比如目前有很多表情识别方法被提出^[1]。然而每一种方法在不同的条件下有其优点和缺点,随着图像采集环境的不同,识别率和鲁棒性迅速下降。为了提高图像识别的系统整体识别率和鲁棒性,为此提出一种基于支持向量机信息融合的方案来融合三种基于不同特征表示的表情识别方法进行表情识别,以及设计一个多 DSP 的并行实现结构来实现算法。为了满足实时性,从 SVM 特征策略、多类 SVM 构造、软件优化等方面进行优化。

2 SVM 简介和系统结构

2.1 SVM 简介

支持向量机(SVM)是一种基于结构风险最小化原理的机器学习方法^[2],SVM 算法原理可简要表述如下:寻找一个在有

限样本情况下的最优分类面,即寻找最优判别函数,使分类间隔到达最远;而对于线性不可分的问题,特别是高维、非线性的图像识别问题,则通过核函数将低维空间的数据映射到高维空间,目的就是原低维不可分的问题转化为高维空间线性可分问题。

设 $\{x_i, y_i\}, i=1, 2, \dots, n$ 为训练样本, $x_i \in R^d$ 为 d 维训练样本, $y_i \in \{-1, 1\}$ 为样本 i 的类别信息。SVM 就是寻找最优的分类面: $x^T \cdot w + b = 0$ 。为了使待分样本尽可能分开,分类间距最大就是使得目标函数 $\frac{1}{2} \|w\|^2$ 最小。于是寻找最优分类面的问题就转化为下面的优化问题:

$$\min \frac{1}{2} \|w\|^2 \quad (1)$$

$$\text{s.t. } y_i \lfloor x_i^T \cdot w + b \rfloor - 1 = 0, i=1, 2, \dots, n \quad (2)$$

方程(1)是个优化问题,为了减少孤立点对分类的影响和提高训练和识别速度,常常将方程(1)的优化问题转化为下面

基金项目:教育部博士点基金资助项目(No.20050290010);北京市教育委员会共建经费研究生教育资助项目。

作者简介:刘松,男,博士,研究方向:图像处理与识别,机器视觉,嵌入式 DSP 系统,智能监控系统。

收稿日期:2008-07-01 修回日期:2008-08-06

其对偶问题:

$$\max T_D = \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i,j} \alpha_i \alpha_j y_i y_j x_i x_j \quad (3)$$

$$\text{s.t. } \sum_{i=1}^n \alpha_i y_i = 0 \quad (4)$$

$$0 \leq \alpha_i \leq C, i=1, 2, \dots, n \quad (5)$$

2.2 系统结构

提出的基于 SVM 信息融合技术的图像识别系统(SVMS)结构如图 1 所示。系统的基本思想是充分利用面部表情的低阶和高阶信息,通过三种不同的特征提出技术来提取面部表情特征的不同表示,即主元分析(PCA)、支持向量机(SVM)、Fisher 鉴别分析(FLD)。对于每一种特征表示分别用 1 阶最近邻分类器(KNN1)进行初步分类,最后用基于 SVM 的信息融合技术来融合这些初步分类结果进行面部表情的最终识别。其中主元分析(PCA)特征提取充分利用图像的低阶分量信息,Fisher 鉴别分析(FLD)特征提取利用了图像的高阶分量信息,支持向量机(SVM)特征提取充分利用了 SVM 算法内在的许多优秀性质,比如良好的未知样本的适应和推广能力,这对于提高系统鲁棒性很重要。

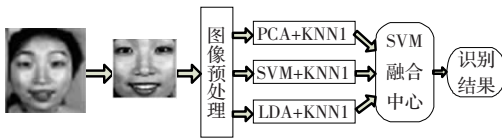


图 1 基于 SVM 信息融合系统图

3 面部表情的初次识别

3.1 主元分析

主元分析(PCA)是均方误差最小意义下的数据压缩的方法,一副 $n \times n$ 的人脸图像,将其按列相连构成一个 n^2 维的表情图像。设 N 为训练样本数, x_j 为第 j 副人脸图像形成的表情向量,则训练样本的总体散布矩阵为:

$$\Sigma = \sum_{j=1}^N (x_j - \mu)(x_j - \mu)^T \quad (6)$$

其中 μ 为训练样本的平均数: $\mu = \frac{1}{N} \sum_{j=1}^N x_j$ 令 $X = [x_1 - \mu, \dots, x_N - \mu]$,

$\Sigma = XX^T$ 令 $R = X^T X \in \Omega^{N \times N}$, 根据矩阵理论, R 的特征值 λ_j 及相应的正交归一化特征向量 v_j 与 Σ 的正交归一化特征向量 μ_j 的关系如下:

$$\mu_j = \frac{1}{\sqrt{\lambda_j}} X v_j, j=1, \dots, N \quad (7)$$

将特征值从大到小排列: $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_N$, 其对应的特征向量为 $\mu_1 \geq \mu_2 \geq \dots \geq \mu_N$ 。任何一副表情图像向量 x_j 投影到由 $\mu_1, \mu_2, \dots, \mu_N$ 张成的子空间 U^T , 都能得到一组表示图像在子空间位置的坐标系数,这些系数可作为代表人脸表情识别的依据。对于任意样本可通过向“表情脸”子空间投影求出其系数向量: $y = U^T(X - \mu)$ 。

3.2 支持向量机

支持向量机(SVM)由于表情图像识别是典型的多类非线性问题,首先要通过 $\Phi: x_i \rightarrow \Phi(x_i)$ 非线性映射将人脸表情映射到高维的特征空间,选用高斯径向基函数核函数 $k(x_i, x_j)$ 来实现两个向量 $\Phi(x_i), \Phi(x_j)$ 的内积,即有:

$$k(x_i, x_j) = \Phi(x_i) \cdot \Phi(x_j) = e^{-\frac{\|x_i - x_j\|^2}{2\sigma^2}} \quad (8)$$

于是优化问题式(3)转化为:

$$\max T_D = \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i,j} \alpha_i \alpha_j y_i y_j k(x_i, x_j) \quad (9)$$

$$\text{s.t. } \sum_{i=1}^n \alpha_i y_i = 0 \quad (10)$$

$$0 \leq \alpha_i \leq C, i=1, 2, \dots, n \quad (11)$$

这样就得到投影轴 w , 然后把任意样本 x 进行投影,从而得到任意样本 x 的投影坐标:

$$y = w \cdot x = \sum_{i=1}^n \alpha_i y_i k(x_i, x) \quad (12)$$

这样就把两类样本通过投影轴 w 很好地分开,对于多类样本,比如 c 样本,首先从 c 类样本中选择一类作为第一类,其余的 $c-1$ 类作为第二类,得到一个投影轴,对于其余的 $c-1$ 依次采取相同的策略后,得到 $c-1$ 个投影轴,然后把任意样本在 $c-1$ 个投影轴上投影就得到 $c-1$ 维表示的表情特征。

3.3 Fisher 鉴别分析

FLD 基本思想就是在 Fisher 鉴别函数取得极值的条件下,求得一个最佳的鉴别方向,然后再将高维模式空间表示人脸面部的特征向量投影到该最佳方向上,构成一维的鉴别特征空间,它试图最大化类间分布,而最小化类内分布,对于面部表情识别而言,FLD 移动同一类表情的人脸图像更近,而不同类表情的人脸图像移得越远。设训练样本集共有 N 个训练样本 $\{x_1, x_2, \dots, x_N\}$, 分为 c 类 $\{\chi_1, \chi_2, \dots, \chi_c\}$, 每一类的均值为 $\bar{\mu}_i =$

$$\frac{1}{N_i} \sum_{x_i \in \chi_i} x_i, \text{ 总均值为: } \bar{\mu} = \sum_{k=1}^{k=N} x_k, \text{ 第 } i \text{ 类样本的离散度矩阵为:}$$

$$S_i = \sum_{x_i \in \chi_i} (x_i - \bar{\mu}_i)(x_i - \bar{\mu}_i)^T \quad (13)$$

$$\text{总类内离散度矩阵: } S_w = \sum_{i=1}^i S_i, \text{ 类间离散度矩阵: } S_B =$$

$$\sum_{i=1}^c |\chi_i| (\bar{\mu}_i - \bar{\mu})(\bar{\mu}_i - \bar{\mu})^T, \text{ 离散度矩阵: } S_T = S_w + S_B, \text{ 求下列两两正交的矩阵:}$$

$$W_{opt} = \arg \max \frac{|W^T S_B W|}{|W^T S_w W|} = [w_1, w_2, \dots, w_m] \quad (14)$$

$$w_i \text{ 是 } S_B \text{ 与 } S_w \text{ 的广义特征值 } \lambda_i \text{ 相对应的广义特征向量, 即 } S_B w_i = \lambda_i S_w w_i, i=1, 2, \dots, m \quad (15)$$

由于最多只有 $c-1$ 个非零广义特征值, 因此 m 的最大值为 $c-1$ 。

为了解决小样本问题,提出 PCA+FLD 策略,通过 PCA 将特征空间降为 $K-c$ 维,然后通过标准的 FLD 将维数降为 $c-1$ 。通过 PCA 将人脸图像投影到 $K-c$ 维特征空间:

$$W_{pca} = \arg \max |W^T S_T W| \quad (16)$$

$$Z_k = W_{pca}^T x_k \quad (17)$$

通过标准的 FLD 将维数降为 c :

$$W_{fld} = \arg \max \frac{|W^T W_{pca}^T S_B W_{pca} W|}{|W^T W_{pca}^T S_w W_{pca} W|} \quad (18)$$

$$y_k = W_{fld} Z_k \quad (19)$$

4 基于支持向量机的信息融合

信息融合技术是将各种途径、时间、空间上获得的信息作为一个整体进行分析、处理和决策。其三个层次分别为:数据层次融合、特征层次融合、决策层次融合。采用支持向量机对多分类器进行融合,即决策层次的融合。假设有 n 个分类器,若已经取得 k 个决策样本值以及实际的决策值:

$$\{(z_1, y_1), (z_2, y_2), \dots, (z_k, y_k)\}$$

$z_i = [z_{i1}, z_{i2}, \dots, z_{in}]^T$ 表示对于第 i 个输入样本图像,这 n 个分类器作出的决策矢量, y_i 是对应第 i 输入样本图像的标签。要求能根据这组样本值,找到一个适当的融合函数 $f(z)$,使得对于确定的损失函数 L 具有最小的期望风险。

从信息变换的角度来看,多分类器起到把输入的人脸图像映射到决策空间的作用。为了很好地识别人脸表情,基于支持向量机的决策融合应作如下处理:首先,将多种分类器的决策 z 通过非线性映射 $\phi: z \rightarrow \phi(z)$ 映射到一个高维空间 Ω 中去,然后,在高维空间 Ω 中进行线性回归。综合上述两个过程,得到融合后的决策为:

$$f(z) = w \cdot \phi(z) + b \quad (20)$$

其中, \cdot 为普通的向量内积运算。类似于公式(12)得到融合函数为:

$$f(z) = \sum_{i=1}^k \alpha_i k(z, z_i) + b \quad (21)$$

并且选高斯径向基函数核函数为变换核变换函数。其多分类器的决策级融合过程如图2所示。

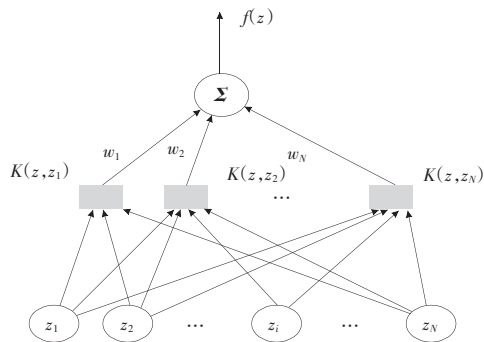


图2 基于 SVM 信息融合模型

为了满足实时性,采用简化结构构造多类 SVM 分类器。

5 并行实现方案

提出的基于 SVM 信息融合的图像识别方法中,由公式(11)和公式(20)可知,对输入对象的识别的主要计算量来自于新的输入特别或初次识别结果和支持向量机之间核函数计算,为了正确实时识别,需要大量乘累加(MAC),采用 TI 最新的 DAVINCI 系列高速 DSP 芯片,充分利用 DSP 的高指令和高效的并行结构补偿算法的复杂性和运算时间的高开销。

5.1 硬件实现

系统的硬件结构采取两级结构,如图3所示:第一级是由3片并行的 TMS320DM648 组成,分别实现初次识别的三个分类器,DM648 具有高性能的 TMS320C64PLUS+DSP(900 MHz)的体系结构^[4],第二级是有一片 TMS320DM6446 构成,DM6446 具有高性能的 TMS320C64+DSP(600 MHz)和精简指令集的 ARM926EJ-S(300 MHz)的双核体系结构^[5],双核的 ARM 核负责整个系统的控制和管理,DSP 核负责 SVM 融合算法,主处理器 DM6446 通过 EMIFA 总线连接三片从处理器的主机口(HPI)完成对从处理器的控制以及主从处理器之间命令和数据

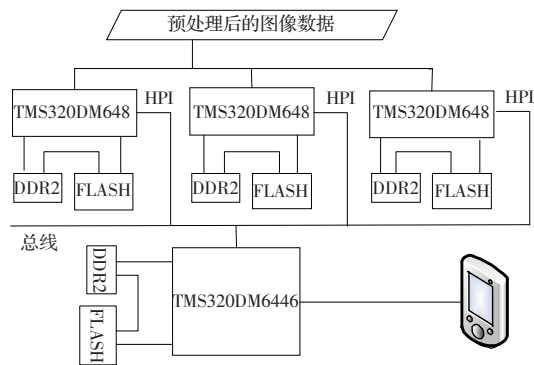


图3 硬件系统原理框图

的传输。这两种类型的芯片都支持高速的 DDR2 和 EDMA 传输方式,能够满足高速的数据传输和高速计算的要求。系统的每个芯片都有自己外扩的 DDR2 和 FLASH。保证了高速的并行性,DDR2 存放待 DSP 处理和分析的数据,FLASH 用于存放每个算法的代码和参数,对于跟 SVM 相关的芯片,对应的 FLASH 中还存放支持向量和系数表。

5.2 软件实现

为了实时实现图像识别,除了上面的 SVM 多类构造简化和硬件并行执行外,算法的软件实现过程要特别以下几点:

(1)首先用 MATLAB 进行算法仿真,尽量提高算法的效能,然后由仿真模块生成 TIC6000 下的 C 语言代码;

(2)对于 SVM 有关的算法,特别是对新的输入数据或初次识别结果和支持向量机之间核函数计算进行的线性汇编优化,以实现算法的并行流水执行;

(3)对每个芯片都采用4个线程的多线程执行:HWI,SWI1、SWI2、IDL, 优先级顺序有高到低,HWI 线程为了硬件事件,SWI1 为了识别算法的执行,SWI2 为了主从交互的数据更新,IDL 为了空闲执行。

6 结果与分析

实验采用日本女性表情数据库 JAFFE,该库有10人组成,每人包含愤怒、高兴、厌恶、悲伤、中性、恐惧、惊奇共7种表情,首先将图像预处理为 30×30 像素大小。采用每个表情10个样本共70个样本作为训练,每个表情20个样本共140个样本进行测试,表1是实验结果,从表1可以看出,提出的基于支持向量机信息融合的面部表情识别方法的识别明显高于基于单特征单分类器的面部表情识别方法,更重要的是该文基于信息融合的方法有更好的鲁棒性,即使融合方法中有一个、两个不能工作,识别系统照样能够正常工作,总之,基于信息融合的面部表情识别方法是更有效,无论在提高识别率,还是增加系统鲁棒性,从而也说明基于信息融合的方法是面部表情识别一个很值得探索的领域。

表1 表情识别率 (%)

PCA	SVM	FLD	SVM 融合
78.36	89.68	82.43	95.75

7 结论

为了提高复杂目标的图像识别问题,提出了基于 SVM 信息融合的识别方法,该方法不仅充分利用了图像的低维和高维信息,而且利用了 SVM 算法内在的许多优秀性质,比如良好的

(下转 182 页)