

基于边缘梯度方向直方图的静态手语识别

孙丽娟, 张立材

(西安建筑科技大学 信息与控制工程学院, 陕西 西安 710055)

摘 要: 文中采用边缘梯度方向直方图作为手势的特征矢量进行手语识别, 建立归一化的边缘梯度直方图, 使用欧氏距离模板匹配法进行手势的特征匹配, 手势特征矢量之间的识别速度较快. 实验表明: 该方法对图像亮度、缩放、平移具有不变性, 该方法计算简单、快速, 可以用于手语识别系统.

关键词: 手语识别; 梯度; 边缘梯度直方图

中图分类号: TP391

文献标识码: A

文章编号: 1000-7180(2010)03-0148-03

Static Sign Language Recognition Based on Edge Gradient Direction Histogram

SUN Li-juan, ZHANG Li-cai

(School of Info. and Control Engineering, Xi'an University of Architecture & Technology, Xi'an 710055, China)

Abstract: In this paper, we use the edge gradient direction histogram as a feature vector for sign language recognition, set up normalized edge gradient direction histogram. The compare of gesture feature vector is rapid used Euclidean distance template matching method. The experiments result shows that the algorithm has the robustness to scene illumination changes, position translation and scale. This method is simple and fast, so we can apply this method to the hand gesture recognition system.

Key words: sign language recognition; gradient; edge gradient direction histogram

1 引言

随着多模式人机接口技术的发展, 手语识别作为智能人机交互领域中的重要组成部分得到了广泛的关注. 手语作为一种自然语言, 是具有听力障碍者之间进行交流的重要工具.

手语识别的方法很多, 比较有代表性有美国的 Cui 和 Weng^[1,2]研究了最具区分性特征(MDF)和最具表达性特征(MEF)的不同, 采用模板匹配的方法, 实现了一个可以识别 26 个手势的系统, 识别率为 93%. 德国的 Kirsti Grobel 等^[3]利用 HMM 方法实现了可识别 262 个孤立手语词的识别系统, 识别率为 91.3%. 文献[4]采用 DGMM(Dynamic Gaussian Mixture Model)方法, 孤立词的识别率为 90%

左右, 句子的识别率为 95.2%. 这些算法都比较复杂, 计算量大, 执行时间较长. 文中使用边缘梯度方向直方图^[5]作为手势的特征矢量进行静态手语识别. 实验表明: 该算法计算简单、快速, 手势特征矢量之间的识别速度较快, 能够满足实时的手语识别的要求.

2 边缘梯度直方图

图像中手势和背景具有不同的灰度值, 相邻区域之间边缘的灰度值是不连续的, 因此边缘上每一像素都对应此处一个边缘梯度方向. 边缘可以看作由具有特定方向的边缘像素点构成, 由边缘像素点的梯度方向建立的直方图能够表示物体的形状. 对于图像 $f(x, y)$ 在 (x, y) 处的梯度定义为:

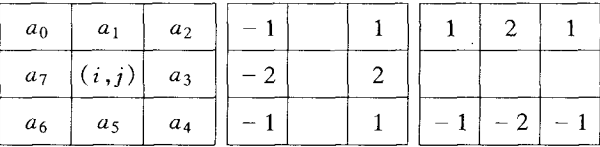
$$\nabla f(x,y) = [G_x, G_y]^T = \left[\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \right]^T \quad (1)$$

梯度是一个矢量,其大小和方向为

$$G[f(x,y)] = |\nabla f(x,y)| = \sqrt{G_x^2 + G_y^2}$$
$$= \sqrt{\left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2} \quad (2)$$

$$\theta(x,y) = \arctan[G_y/G_x] \quad (3)$$

其中, $\theta(x,y)$ 为 $f(x,y)$ 在该点灰度变化最大的方向. 在数字图像中用小区域模板卷积来近似计算每个像素的梯度值,文中选用如图1所示的 Sobel^[6] 梯度算子. 在 3×3 邻域计算梯度,对 G_x 和 G_y 用图1所示模板计算其梯度图.



(a) 3×3 邻域 (b) 水平方向模板 (c) 垂直方向模板

图1 Sobel 算子模板

偏导数用下式计算:

$$G_x = (a_2 + 2a_3 + a_4) - (a_0 + 2a_7 + a_6),$$

$$G_y = (a_4 + 2a_5 + a_6) - (a_0 + 2a_1 + a_2).$$

使用模板卷积计算出水平方向的梯度图 G_x 和垂直方向的梯度图 G_y ,利用 Sobel 算子计算手势的边缘,如图2所示是字母 y 的 G_x 梯度图和 G_y 的梯度图及其手势边缘.

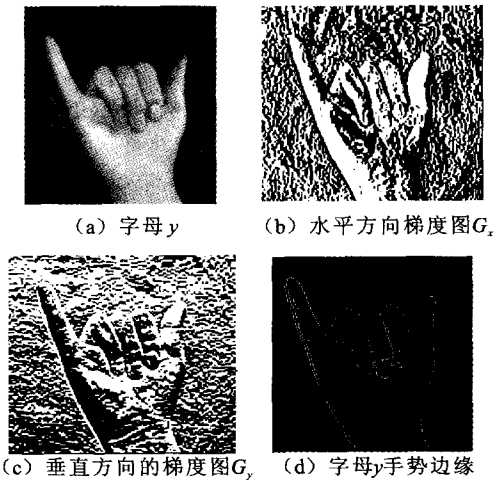


图2 字母 y 的梯度及其边缘

按照式(3) 计算出手势边缘的方向角 $\theta(x,y) \in [0 \sim 2\pi]$,根据边缘像素点的方向角建立梯度方向直方图 $H(i)$. 具体方法为将梯度方向角 $\theta(x,y) \in [0 \sim 2\pi]$ 均匀地量化为 k 个区间,即 $\varphi(i)$ 的取值范围为 $0,1,\dots,k$,定义为:

$$\varphi(i) =$$

$$\sum_{x,y} \begin{cases} 1 & \text{if } \theta(x,y) \in [(i-1) \times \frac{360}{k}, i \times \frac{360}{k}] \\ 0 & \text{其他} \end{cases}$$
$$i = 0,1,\dots,k \quad (4)$$

用边缘像素的总数得到归一化的梯度方向直方图 $H(i)$. 在实验中分别取 $k = 20,24,36,72$. 考虑到既要包括足够的信息以区分不同的手语图像,又要使直方图特征矢量的数据量尽可能少以便实现快速匹配,最后选择 $k = 36$. 在边缘梯度直方图中,横坐标代表 36 个角度级,每级为 10° ,纵坐标代表对应角度级的像素个数的百分比.

如图3所示字母 y 及其边缘梯度方向直方图. 图中可见,边缘梯度方向直方图从统计意义上表征图像边缘取向的分布特性,不同的图像对应不同的边缘分布,即具有边缘取向分布特性.

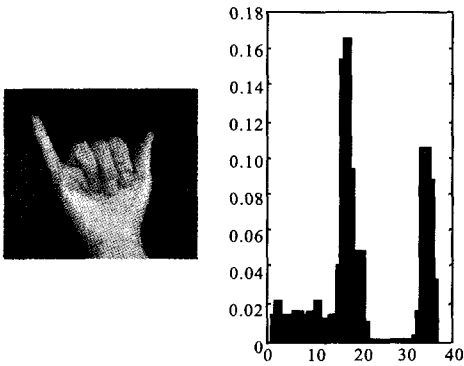


图3 字母 y 及边缘梯度直方图

字母 w 在不同光照下的边缘梯度直方图如图4所示. 从图中可见,边缘梯度方向直方图对图像的光照变化不敏感. 当图像放大、缩小时,手势图像的边缘像素数目发生变化,采用归一化的梯度方向直方图使其具有尺度不变性. 另外,同一手势图像平移后其直方图不变,但对角度旋转敏感. 当手势图像的角度发生变化时,所有边缘像素的梯度方向将增加或减少角度,相应的边缘梯度直方图会发生移位.

在实验中将同一手势边缘方向直方图进行平均作为该手势的特征矢量,将所有要识别的手语手势的边缘梯度直方图建立一个数据库. 识别时使用欧式距离法将测试手势图像的边缘梯度方向直方图与边缘直方图数据库中的特征矢量进行匹配.

边缘梯度直方图识别算法描述如下:

- (1) 对图像进行滤波,得到图像的水平方向梯度 dx 、垂直方向的梯度 dy .
- (2) 用边缘检测得到手势图像的边缘像素点,利用水平和垂直方向的梯度图像得到边缘像素的梯度值.

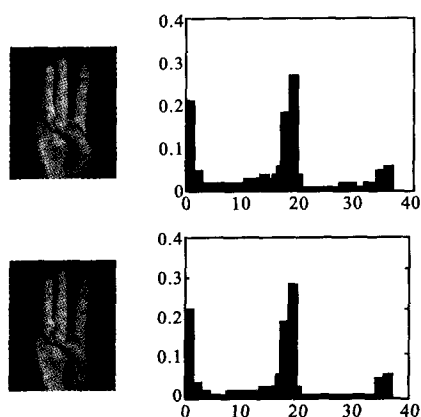


图4 字母 *w* 在不同光照下的边缘梯度直方图

(3) 计算边缘像素点的梯度方向角 $\theta(x, y)$, 并将其转换到 $0 \sim 2\pi$ 上, 同时对梯度方向角进行分类, 将梯度方向角分为 k 类 ($k = 36$), 统计每一类的像素个数, 用边缘像素的总数得到归一化的梯度方向直方图。

(4) 利用欧式距离的模板匹配法进行手势的特征匹配。欧式距离计算公式如下

$$D = \min_{k=1}^n (\theta - \theta_k)^2$$

其中, θ 是待识别图像的边缘方向直方图矢量, θ_k 是边缘梯度直方图数据库中的特征矢量。

3 实验仿真

根据上面所述方法, 文中采用如图5所示的手语识别系统进行实验。实验以黑色为背景, 用数码相机采集手语的手势图片, 采用不同的光照, 拍摄了30种不同手势, 每种手势10个样本, 共300个手势样本, 将待识别的手语图片建立了一个小型的静态手语库, 图像大小归一化为 225×216 。其中一部分作为训练使用, 首先对训练图像计算边缘梯度方向角, 建立归一化的边缘梯度直方图数据库作为手势识别的特征矢量。将测试的手语图像转化为灰度图像, 进行滤波, 计算图像在水平方向和垂直方向的梯度, 求得手势边缘的梯度方向角, 统计方向直方图, 将测试样本的方向直方图矢量与建立的边缘梯度直方图数据库中训练样本的方向直方图矢量用欧式距离进行比较, 欧式距离最小认为该测试样本与训练样本的手势为同一手势。

为验证梯度直方图对图像大小变化的不变性, 将同一种手势采用不同尺寸进行实验, 如图6所示。可见它们的直方图是一样的, 即边缘梯度直方图对

整幅图像大小的比例变化是不敏感的。对手语库中的近300张手势进行识别, 识别率达95%以上。

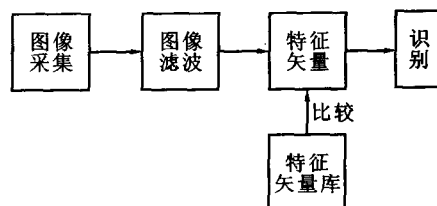


图5 手语识别系统

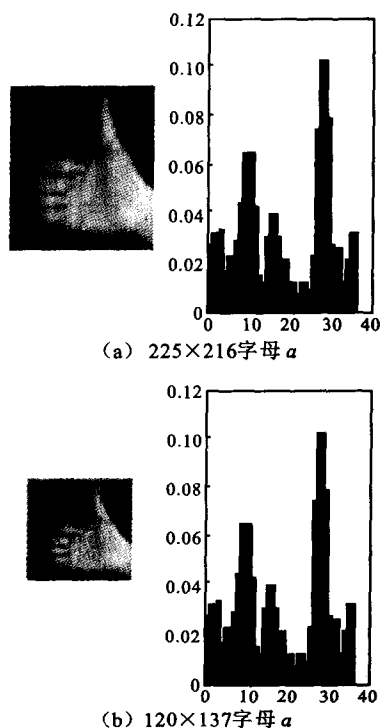


图6 尺寸不同的字母手势 *a* 及其直方图

4 结束语

文中的创新点是采用归一化边缘梯度方向直方图进行静态手语识别。该方法的特点是, 只是统计图像中手势在每个方向上出现的像素个数的百分比, 因而对光照变化不敏感, 而且对手势的位置变化具有不变性, 对整幅图像的比例大小具有尺度不变性。该方法计算简单、快速、识别率高, 实验中识别率达95%以上。本实验只是对部分静态手语进行测试, 没考虑人脸因素, 手语的含义和面部表情是紧密相连的。目前人脸识别技术也日益完善^[7], 因此下一步的工作是将此方法与人脸识别技术结合用于动态手语的识别中。

(下转第153页)

(11) 付网关得到企业传送的消息后,通过使用与企业相同的检验方法,支付网关解密出客户的支付信息以及企业的支付请求信息。

(12) 通过 CA 信用认证核实客户身份、支付信息和账户信息以后,支付网关向企业返回确认信息。

(13) 消费者发卡行将相应的交易金额从客户账号转出到中介行。

(14) 企业继而向客户返回响应信息,同时向配送服务发出运货信息,配送服务向企业、支付网关发出货物成功递送的确认信息及其摘要的数字签名,同时支付网关也将同一份确认信息和自己对该信息摘要的签名转发给消费者。

(15) 客户确认,中介行则将暂存的交易金额转送到企业的收单行,否则,将交易金额转回发卡行,结束交易。

3 结束语

文中讨论了如何通过 WS-Security 规范和 XML 技术实现电子商务交易中的安全支付。XML 安全组件集成了包括 XML 数字签名服务和元素级加密服务在内的多项重要功能,并且实现了交易的不可否认性。元素级加密非常灵活地对传输的信息进行选择性的加密,使订单信息和支付信息同一个过程中进行不同的处理,从而保证了在消息层上任意敏感

信息的安全性。WS-Security 保证了 SOAP 消息安全传递的增强,提供了一种构建安全模型和加密技术的灵活机制。

参考文献:

- [1] 张爱菊. 电子商务安全技术[M]. 北京:清华大学出版社,2008.
- [2] 黄伟峰. Web Services 的安全性[J]. 计算机工程与设计,2003,24(6):52-55.
- [3] 陈桦,麻风梅,韩艳艳. 基于 XML 的异构数据集成模式的研究[J]. 微电子学与计算机,2009,26(1):137-139.
- [4] 李慧盈,张长海,李德昌. 基于 SOAP 协议的 Web Services 安全性扩展实现[J]. 计算机应用研究,2006,23(1):106-107.
- [5] 刘义春. 电子商务中的信任机制研究[J]. 微电子学与计算机,2008,25(3):121-125.
- [6] 谢琳,卢建军. 电子商务中第三方电子支付平台分析[J]. 计算机应用研究,2003,20(12):149-151.
- [7] 郭瑞颖,毛明,吴明玉,等. 一种基于 Web 的 B2B 安全电子支付模型的探讨[J]. 小型微型计算机系统,2001,22(10):1274-1276.

作者简介:

郭悦红 女,(1977-),博士研究生,讲师.研究方向为信息管理。

(上接第 150 页)

参考文献:

- [1] Cui Y, Weng J J. Hand segmentation using learning-based prediction and verification for hand sign language [C]// Proc of ICCV'96. USA:San Francisco,1996:88-93.
- [2] Cui Y, Swets D L, Weng J J. Learning based hand sign language using SHOSLIF-M[C]// Proc of ICCV'96. USA:San Francisco,1996:61-636.
- [3] Grobel K, Assan M. Isolated sign language recognition using hidden markov models[C]// Proc. of the International Conference of System, Man and Cybernetics. USA:Orlando,1996:162-167.
- [4] Wen Gao, Ma Jiyong, Wu Jiangqin, et al. Lange vocabulary sign language recognition based onHMM/ANN/DP

[J]. International Journal of Pattern Recognition and Artificial Intelligence,2000,14(5):587-602.

- [5] Woods Richard E, Eddins Steven L. 数字图像处理[M]. 阮秋琦,译.北京:电子工业出版社,2003.
- [6] 章毓晋. 图像工程:上册——图像处理和分析[M]. 北京:清华大学出版社,2001.
- [7] 刘志镜,夏勇,李夏忠. 基于正交视图的多姿态人脸识别算法[J]. 微电子学与计算机,2004,21(3):11-15.

作者简介:

孙丽娟 女,(1976-),硕士研究生.研究方向为图像信息处理。

张立材 男,副教授,硕士生导师.研究方向为信号处理。