

# 多目标优化的静态手语识别算法研究

赵一丹, 肖秦琨, 郭 鹏

(西安工业大学 电子信息工程学院 陕西 西安 710021)

**摘 要:** 手势是人类与计算机交互的直观方式, 随着人工智能技术的发展, 以机器为核心的计算模式正朝着以人为中心的计算模式转变, 自然且符合人类习惯的人机交互(HCI)方式逐渐成为目前研究的热点。一个高效的人机交互系统应该以良好的识别精度和识别速度为目标。文中提出了一种基于深度信息的静态手势识别方法, 包含了手势大小、光照和旋转变化等因素的影响。从识别精度和速度两个方面, 综合比较了如 Hu 矩、Zernike 矩、伪 Zernike 矩、傅里叶描述符和 Gabor 特征等几种常见的图像特征描述符。手势识别采用多层感知器, 它具有结构灵活、识别速度快等特点。为了提高识别精度、减少计算量, 特征向量和神经网络均通过基于 NSGA-II 的多目标进化算法调整。在对手语识别的进一步探究中, 对所提出的方法的有效性进行验证。

**关键词:** 手势特征提取; 手语识别; 多目标优化; 遗传算法

中图分类号: TP301.6

文献标识码: A

文章编号: 1673-629X(2019)02-0054-06

doi: 10.3969/j.issn.1673-629X.2019.02.011

## Research on Static Sign Language Recognition Algorithm Based on Multi-objective Optimization

ZHAO Yi-dan, XIAO Qin-kun, GUO Peng

(School of Electronic Information Engineering, Xi'an Technological University, Xi'an 710021, China)

**Abstract:** Gesture is an intuitive way for human beings to interact with computers. With the development of artificial intelligence technology, the machine-centric computing mode is shifting toward the human-centered computing mode, and the human-computer interaction (HCI) methods that are natural and in line with human habits are gradually become the focus of current research. An efficient human-computer interaction system should aim for better recognition accuracy and recognition speed. For this, we present a static gesture recognition method based on depth information, including the influence of gesture size, lighting, and rotation changes. Several common image feature descriptors such as Hu moment, Zernike moment, pseudo Zernike moment, Fourier descriptor and Gabor feature are comprehensively compared from their respective recognition accuracy and speed. Gesture recognition uses a multi-layer sensor which has a flexible structure and fast recognition speed. In order to improve the recognition accuracy and reduce the computational complexity, the eigenvectors and the neural network are adjusted by the multi-objective evolutionary algorithm based on NSGA-II. In the further exploration of the recognition of the opponent language, the validity of the proposed method is verified.

**Key words:** gesture feature extraction; sign language recognition; multi-objective optimization; genetic algorithm

### 1 概 述

随着人工智能技术的发展, 以机器为核心的计算模式正朝着以人为中心的计算模式转变, 自然且符合人类习惯的人机交互(HCI)方式逐渐成为目前研究的热点。可以预知的是下一代的人机交互平台应该是以手势为主, 语音为辅, 而裸手操作则是最自然的人机交互方式。另外, 手语作为一种表达的工具在手语使用者与他人的交流之间产生着重大影响。对手语识别的

研究不仅可用于有效、自然的人机交互, 对聋哑人或听力受损人群与健全人之间的沟通也有重要的作用。

手势识别、步态识别、动作识别等众多行为分析技术已经逐步应用在社会生活当中。一个完整的手势识别系统包括手势数据的采集、手势分割、手势特征提取、手势特征表达以及手势识别分类<sup>[1-2]</sup>。在手势分类问题中关键在于如何提取一组好的手势特征并使用有效的分类器进行识别分类。Guo 等<sup>[3]</sup>提出了一种改

收稿日期: 2018-03-28

修回日期: 2018-08-02

网络出版时间: 2018-11-15

基金项目: 国家自然科学基金(61271362, 61572392, 61671362); 陕西省自然科学基金(2017JM6041)

作者简介: 赵一丹(1993-), 女, 硕士研究生, 研究方向为智能信息处理、手势识别; 肖秦琨, 博士, 教授, 硕导, 研究方向为人工智能与图像处理。

网络出版地址: <http://kns.cnki.net/kcms/detail/61.1450.TP.20181115.1050.070.html>

进的 Zernike 矩 26 个英文字母静态手势识别的精度和计算成本。他们认为大多数对静态手势识别的研究中,为了提高识别的准确率,大都着眼于对手势特征的提取以及分类的过程,很少有研究考虑降低识别过程中的计算成本。Wu 等<sup>[4]</sup>针对原始超限学习机在手势识别应用中欠缺良好的泛化性能和鲁棒性等问题,运用主成分分析(PCA)算法提取手势图像的主要特征,通过分层超限学习机的稀疏自动编码和分层训练,获得原始输入的多层稀疏表达,使自动编码后的输出近似原始输入,最大限度地减少重构误差,提高特征分类的精度。实验的识别精度达到 95.19%。Pu 等<sup>[5]</sup>针对 Kinect 深度信息下静态手势识别鲁棒性差及识别率低的问题,提出构造一种新的 Hu 不变矩,采用尺度归一法消除尺度变化对 Hu 矩的影响,对六种静态手势进行识别的准确率达到 98.9%。Chaudhary 等<sup>[6]</sup>比较了人工神经网络、模糊逻辑和遗传算法在基于计算机视觉的手势识别中的不同应用,成功地将进化算法应用于手势识别,但尚未用于特征选择和优化手势分类器。Hasan 等<sup>[7]</sup>提出了一种利用多层感知器和复数矩对 6 种不同的手势进行分类的手势识别系统,识别率为 86.38%。Aowal 等<sup>[8]</sup>测试了判别性 Zernike 矩(DZMs)提取图像特征用于识别,并将其与标准 Zernike 矩、主成分分析和傅立叶描述子(FD)进行比较。

虽然特征集对于任何分类任务都非常重要,但在实时的识别系统中计算成本对于提高实时系统的效率同样重要。例如,尽管 Zernike 矩成功地用于不同的图像识别问题,但仍具有缺点:与其他常用图像描述符相比,CPU 时间更长。在现有研究中,Gabor 滤波器所表示的特征描述符的计算速度最快。另外,神经网络的拓扑结构还会影响识别精度和速度。较大的神经网络会更好地识别某种模式,但它会带来更大的训练和分类成本。遗传算法已成功应用于神经网络的性能改进。Leung 等<sup>[9]</sup>提出了一种在其链路中具有开关的神经网络,使用遗传算法来调整其结构和参数。隐藏节点的数量是手动选择的,从一个小的数字开始并增加,直到达到很好的识别率。遗传调整的结果是一个较小的、部分连接的网络,从而减少了计算、硬件和内存成本。文中提出的遗传算法同时优化了隐含神经元的数量和神经网络中输入的配置。

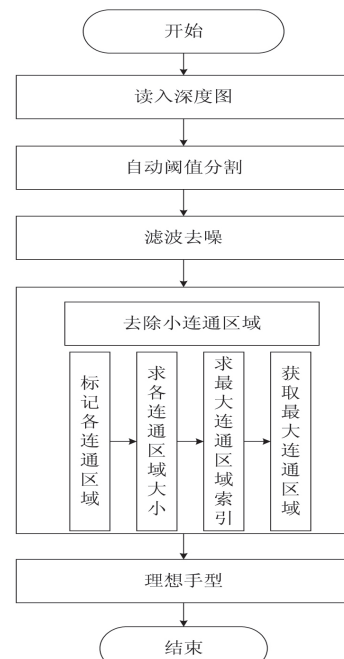
鉴于选择最佳特征集合来表示手势的重要性以及实时执行此任务的分类器的计算成本,提出比较了几种常用特征,即 Zernike 和伪 Zernike 矩、傅里叶描述符、Hu 特征和 Gabor 特征。基于识别准确度和计算成本将这些图像描述符的识别结果彼此进行比较。鉴于特征选择是一个多目标问题,提出使用多目标进化算法(NSGA-II)来选择特征的最佳组合并优化神经

网络的性能。人工神经网络和遗传算法已经在分类问题中应用,文中将两者结合起来,对测试不同特征描述符在准确性和识别速度方面进行比较。

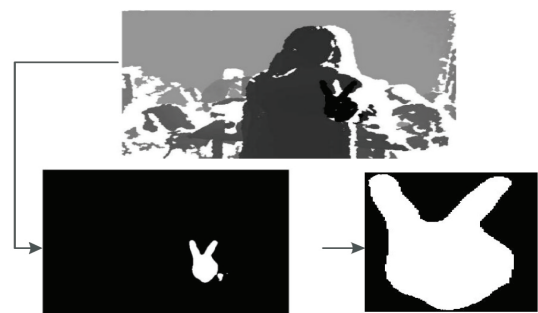
## 1 手势分割及特征提取

### 1.1 手势分割

在识别分类系统中,图像分割是为后续的特征提取和识别做准备,图像分割的好坏直接影响分类结果。文中主要是对 Kinect 2.0 体感外设获取的深度信息的手形分割,深度信息可以有效去除场景中光照、背景的干扰。首先规定手部位于图像最前端,即整幅图像中深度值最小的部分。其次确定深度阈值,将小于阈值的部分分割,从而得到手形部分。此方法主要用于深度图像,并且对背景和光照的限制极少,提高了实验的可行性和实用性。深度图像中,前景目标的深度值与背景各部分深度值不同,故可以利用阈值分割算法提取前景信息。整个手势分割算法流程如图 1 所示。



(a) 手势分割流程



(b) 手势分割结果

图 1 手势分割算法流程

### 1.2 手势特征提取

文中对手势特征的提取主要采用以下手势特征描

述符: Hu 矩、Zernike 矩、伪 Zernike 矩、傅里叶描述符和 Gabor 特征。同时,在进行手语识别时,将以上不同的手势特征描述符组合起来,对其分类结果的准确性进行对比。

### 1.2.1 Hu 矩

几何矩是由 M.K.Hu 在 1962 年提出的,并证明了 7 个不变矩对图像的平移、旋转、镜像和尺度变换具有不变性<sup>[10]</sup>,是最早建立的图像描述符之一。几何矩的七个分量定义为:

$$\begin{aligned} I_1 &= \eta_{20} + \eta_{02} \\ I_2 &= (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2 \\ I_3 &= (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2 \\ I_4 &= (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2 \\ I_5 &= (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \\ I_6 &= (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03}) \\ I_7 &= (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] - (\eta_{30} - 3\eta_{12})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \end{aligned} \quad (1)$$

### 1.2.2 Zernike 矩

Zernike 多项式是于 1934 年由 Frits Zernike 提出<sup>[11]</sup>。Zernike 提出了一组正交多项式  $\{V_{nm}(x, y)\}$ , 图像的 Zernike 矩是该图像在正交多项式上的投影,而这组多项式在单位圆内  $\{x^2 + y^2 \leq 1\}$  是正交的。多项式的形式是:

$$V_{nm}(x, y) = V_{nm}(\rho, \vartheta) = R_{nm}(\rho) \exp(jm\vartheta) \quad (2)$$

其中,  $n$  为正整数或零,  $m$  为正或负整数,必须满足  $n - |m| = \text{even}$ ,  $|m| \leq n$ ;  $\rho$  为圆点到  $(x, y)$  像素点长度的矢量;  $\vartheta$  为矢量  $\rho$  和轴  $x$  的夹角;  $R_{nm}(\rho)$  为径向多项式。

重复率为  $m$  的  $n$  阶 Zernike 矩定义为:

$$A_{nm} = \frac{n+1}{\pi} \iint_{x^2+y^2 \leq 1} f(x, y) V_{nm}^*(\rho, \vartheta) dx dy \quad (3)$$

对于一幅数字图像,积分用求和代替,即

$$A_{nm} = \frac{n+1}{\pi} \sum_x \sum_y f(x, y) V_{nm}^*(\rho, \vartheta) \quad x^2 + y^2 \leq 1 \quad (4)$$

为了计算一幅给定图像的 Zernike 矩,必须将图像的中心移到坐标圆点,将图像像素点映射到单位圆内。Zernike 矩具有旋转不变性,可以将  $|A_{nm}|$  作为目标的不变性特征。

由上式分析,使用  $V_{nm}^*(\rho, \vartheta)$  可提取图像的特征,

低频特征由  $n$  值小的  $V_{nm}^*(\rho, \vartheta)$  提取,高频特征由  $n$  值大的  $V_{nm}^*(\rho, \vartheta)$  提取。而提出的 7 个不变矩仅由低阶矩(二阶和三阶)构成,Zernike 矩可以任意构造高阶矩,所以从理论上分析,Zernike 矩的识别能力应该高于 Hu 矩。

### 1.2.3 伪 Zernike 矩

伪 Zernike 矩的缺点在于计算径向多项式时的计算量较大。文中使用 Chong 等<sup>[12]</sup>提出的一种改进的快速计算径向多项式  $R_{nm}(\rho)$  的方法。在  $N \times N$  图像上计算标准 PZMs 的计算复杂度为  $O(N^2 n^3)$ ,文中使用 Chong 等<sup>[12]</sup>提出的  $P$  阶递归方法的计算复杂度为  $O(N^2 n^2)$ 。为了进一步提高计算效率,当  $m=0$  时计算 PZM。因此,在该过程中,复数项  $e^{jm\vartheta}$  是恒定的并且等于 1。图 2 显示了通过计算包含 2 425 个图像的一组像素为灰度图像的 10 阶 Zernike 矩和伪 Zernike 矩 CPU 所花费的平均时间,可以观察出计算伪 Zernike 矩比 Zernike 矩要快得多,因此得到了广泛应用。

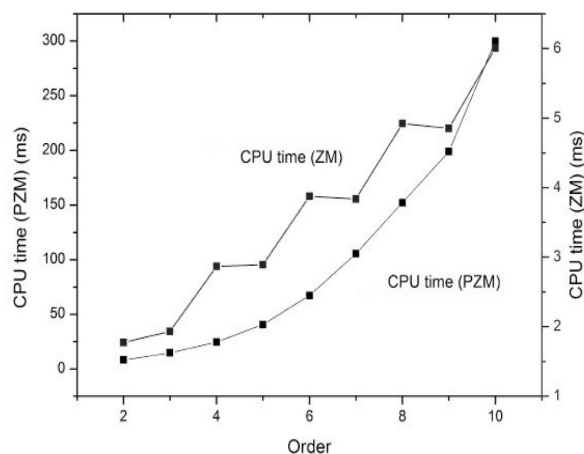


图 2 10 阶 Zernike 矩和伪 Zernike 矩的计算时间

### 1.2.4 傅里叶描述符

傅里叶特征是手语识别中应用最广的一种描述子,它是对手部形状轮廓特征进行提取和描述的一种算法。要得到一个目标图像的傅里叶描述符,首先要描绘出该目标图像的边缘,然后从图像边缘的一点开始,遍历图像边缘,将边缘每一点的坐标  $x(t)$  和  $y(t)$  记录下来。使用轮廓函数的两种不同表示法:复数坐标和质心距离函数<sup>[13]</sup>。

在第一个表示中,轮廓的每个坐标都转换为复函数。

$$z(t) = [x(t) - x_c] + i[y(t) - y_c] \quad (5)$$

其中,  $(x_c, y_c)$  是轮廓的质心坐标。

通过这一转换使  $z(t)$  具有平移不变性,然后对  $z(t)$  进行傅里叶转换。

第二种方法基于下面定义的质心距离函数。与  $z(t)$  一样,质心距离函数也使得具有平移不变性。

$$r(t) = \sqrt{[x(t) - x_c]^2 + [y(t) - y_c]^2} \quad (6)$$

将上述轮廓函数进行离散傅里叶变换(DFT):

$$u_n = \frac{1}{N} \sum_{t=0}^{N-1} s(t) \left( \cos \frac{2\pi nt}{N} - j \sin \frac{2\pi nt}{N} \right) \quad (7)$$

其中,  $N$  是轮廓点的数量;  $s(t)$  是轮廓函数, 即  $r(t)$  或  $z(t)$ 。通常  $u_n$  表示傅里叶描述符。

当  $s(t)$  是一个复函数  $z(t)$  时, 第一个分量  $u_0$  是直流分量, 因为它只依赖于形状的位置, 所以不被使用。特征向量  $f$  则包含  $N-2$  个描述符<sup>[13]</sup>。

$$f = \left[ \frac{|u_2|}{|u_1|}, \frac{|u_3|}{|u_1|}, \dots, \frac{|u_{N-1}|}{|u_1|} \right] \quad (8)$$

对于基于质心距离  $r(t)$  的傅里叶描述符, 式6仅使用实数, 使得傅里叶变换产生  $N/2$  个不同的值。因此, 特征向量大小被限制为  $N/2$ , 如下所定义:

$$f = \left[ \frac{|u_1|}{|u_0|}, \frac{|u_2|}{|u_0|}, \dots, \frac{|u_{N/2}|}{|u_0|} \right] \quad (9)$$

### 1.2.5 Gabor 特征

在图像处理中, Gabor 函数是一个用于边缘提取的线性滤波器<sup>[14]</sup>。Gabor 滤波器的频率和方向表达同人类视觉系统类似。Gabor 特征被广泛应用于通过将图像与滤波核函数进行卷积来识别纹理特征<sup>[15]</sup>。一个二维 Gabor 滤波器是一个由正弦平面波调制的高斯核函数。二维 Gabor 滤波器由下式给出:

$$G(x, y) = \exp\left(-\frac{1}{2}\left(\frac{x_\theta^2}{\sigma_x^2} + \frac{y_\theta^2}{\sigma_y^2}\right)\right) \cos\left(\frac{2\pi x_\theta}{\lambda} + \psi\right)$$

$$x_\theta = x \cos \theta + y \sin \theta$$

$$y_\theta = -x \sin \theta + y \cos \theta$$

$$\sigma_x = \sigma, \sigma_y = \frac{\sigma}{\gamma} \quad (10)$$

其中,  $\sigma$  为高斯包络的标准偏差;  $\theta$  为 Gabor 函数并行条纹的方向(0-360度);  $\lambda$  为波长;  $\gamma$  为空间纵横比;  $\psi$  为相位偏移;  $w$  为滤波窗口的大小。

该滤波器沿着  $\theta$  的方向执行低通滤波, 并执行与其  $\theta$  正交的带通滤波。因此, 通过选择上述参数, 可以增强图像的视觉属性, 例如空间频率和方向。

一个特征矢量的构造如下: 首先对仅包含手型的图像应用 Gabor 滤波器, 然后将得到的图像重新缩放到  $8 \times 8$ , 并且特征向量由相应的 64 个整数组成。得到的 64 个整数是 Gabor 滤波器的特征向量。

## 2 多目标优化算法

### 2.1 多目标优化

当一个最优决策是两个或多个相互冲突的目标之间权衡时, 就会出现多目标问题。特征选择本质上是一个多目标问题, 任何给定的解决方案都可以根据所

选择的特征的复杂性和准确性进行相应的评估。其最终的目标是最小化特征向量的复杂性, 并最大化特征向量的准确性<sup>[16]</sup>。在特征选择的问题上提出了一种解决方案, 如果符合以下标准, 则称解决方案 A 支配解决方案 B。

• A 特征向量的复杂性比 B 低, 或者 A 特征向量的准确性比 B 高。

• B 特征向量的复杂性和准确性都比 A 差。

如图3所示, 考虑三种不同的解决方案:  $S_1$ 、 $S_2$  和  $S_3$ 。假设对所有目标而言, 任意两个解  $S_1$  和  $S_2$ ,  $S_1$  均优于  $S_2$ , 则称  $S_1$  支配  $S_2$ , 若  $S_1$  的解没有被其他所支配, 则把  $S_1$  称为非支配解(不受支配解), 也称 Pareto 解。这些非支配解的集合即所谓的 Pareto front。所有坐落在 Pareto front 中的解皆不受 Pareto front 之外的解(以及 Pareto front 曲线以内的其他解)所支配, 因此这些非支配解较其他解而言拥有最少的目标冲突, 可为决策者提供一个较佳的选择空间。在这种情况下, 多目标优化设计用于搜索一组最接近 Pareto front 的解决方案。

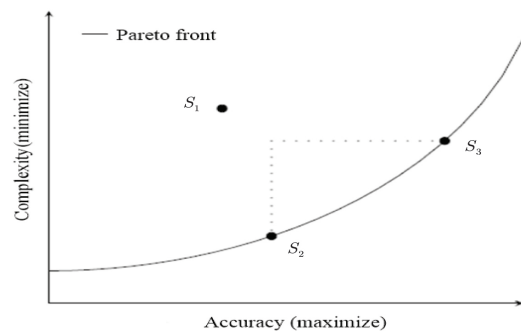


图3 多目标特征选择问题

### 2.2 优化算法

多目标进化算法 NSGA-II(非归类排序遗传算法)是 Deb 等<sup>[17]</sup>提出的, 近年来广泛用于各种多目标优化问题。文中将这一方法应用于静态手势识别。基于 NSGA-II 的多目标优化算法(算法1)的主要步骤如下:

Step1: 从获取的手语数据集中提取特征, 并随机将数据集划分为训练集和测试集;

Step2: 初始化大小为  $N$  的遗传群体  $P_0$ ,  $t=0$  且当  $t$  小于最大迭代次数时, 通过交叉和变异从  $P_t$  获得大小为  $N$  的种群  $Q_t$ ; 通过合并  $P_t$  和  $Q_t$  产生组合种群  $R_t = P_t \cup Q_t$ ;

Step3:  $R_t$  中的每个染色体生成编码输入 MLP; 使用交叉验证评估训练集;

Step4: 对  $R_t$  进行非支配排序, 并通过拥挤和精英保留策略选出  $N$  个个体, 组成新一代种群  $P_{t+1}$ ;

Step5: 评估  $P_t$  在测试子集上的最终 Pareto front;



返回 Pareto front 的解,以及训练和测试分类的准确性。

算法 NSGA-II 通过应用遗传算子交叉和变异获得新的种群  $Q_t$ 。对于种群  $P_t$  的每个染色体,以概率  $p_c$  进行均匀交叉。在均匀交叉中,每个新染色体都是来自  $P_t$  的两个随机选择的染色体(父代)的组合。如果不应用此运算符,则将染色体从  $P_t$  复制到  $Q_t$ ,然后以每个比特的概率  $p_m$  应用位翻转突变。

为了避免优化过程中的偏差,将特征数据库分为训练集(80%)和测试集(20%),如算法 1 所述。在优化循环的过程中,每个染色体代表一种可能的解决方案。通过训练集上的 5 倍交叉验证来评估此解决方案的准确性。其中,交叉验证是在优化算法的内部循环中执行的,以评估单个染色体的性能。在优化循环之后,最终的 Pareto front 在测试子集上进行测试。这被称为特征选择的方法,Kohavi 等<sup>[18]</sup>对此进行了更加详细的讨论。

在 NSGA-II 中,每个特征表示与其他二进制特征选择中使用的相同,其中 1 个特征以二进制字符串编码,“1”表示选择了特征,否则为“0”。

初始化二元染色体最常见的方法是基本上为每一位掷出一枚硬币。在特征选择的情况下,每个位代表一个特定的特征,因此通过翻转一枚硬币,每个特征被

选择的可能性为 50%。这种初始化技术对创建平均具有一半可用特征的染色体群体的结果不太理想。换句话说,像“000010000”这样具有很少特征的序列或者具有大多数可用特征(如“111101111”)的序列将不会出现在初始群体中,从而限制了算法的搜索能力。为了创建一个关于所选特征数量均匀分布的总体,执行统一的初始化,二元染色体群的统一随机初始化(算法 2)主要步骤如下:

Step1: 设  $n$  是染色体中的比特数, $m$  是每个染色体的最小特征数;

Step2: 初始化,将每一个染色体  $C_i$  初始化为  $n$  位,每一位为“0”的二进制字符串(创建了没有选定特征的干净染色体);

Step3:  $r = \max(m, i)$ ,用 1 到  $n$  之间的随机数字创建一个列表  $r$ ;

Step4: 对每一个列表中的每个数字,将  $C_i$  的相应位设置为“1”(结果只使用  $r$  个特征);

Step5: 结束。

图 4 给出了两种初始化方法的比较,其中左图为每个位翻转硬币初始化的 200 个染色体的群体的特征数的频率计数,右图为算法 2 中使用的统一初始化的频率计数。每种初始化方法应用于 200 个染色体的群体,编码 169 个特征。

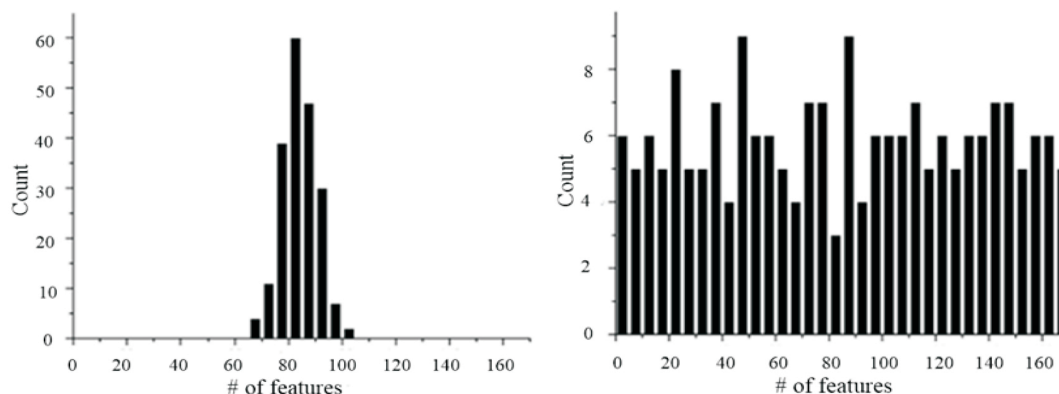


图 4 两种初始化方法的比较

### 3 静态手语分类

采用可配置的三层感知器网络对有限元特征矢量进行分类。图 5 所示为 1 个输入、隐含层和  $n$  个输出的神经网络。该网络有两个参数,可针对特定问题进行优化:

一组开关,将特征矢量  $f_k$  与输入层  $i_k$  连接起来。开关的排列定义了输入层中活动神经元的数量,即实际用于训练和分类的特征组合。

隐藏层中的神经元数量( $n$ ),该参数用于配置网络内的连接数量。

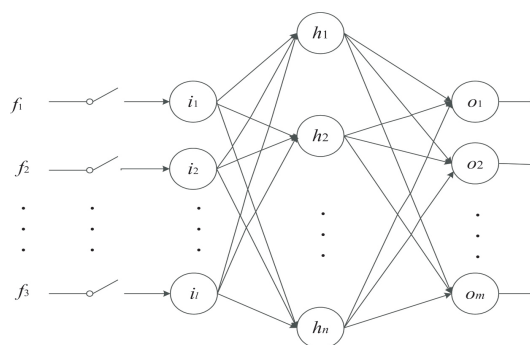


图 5 具有可变输入和隐藏层大小的 MLP

由于每个输出神经元对应于一个类,输出层的大小取决于类( $m$ )的个数。对网络的下列参数进行了

调整:

最大迭代次数: 设置为 300 ,范围介于 100~400。

允许误差: 设置为  $10^{-5}$  ,范围介于  $10^{-6} \sim 10^{-4}$ 。

梯度正常步长: 设置为 0.002 ,范围介于 0.001~0.

1。

反向传播动量值: 设置为 0.01 ,范围介于 0.05~0.

1。

为了评估上述特征向量的辨别能力,将在实验中采用 MLP 分类器测试每个特征向量以及不同特征向量的组合。对于每一组实验考虑具有 50,100,200,300,400 和 500 个神经元的隐藏层进行测试。通过分析实验结果验证该方法的有效性。并统计不同描述符的分类精度和 CPU 所花费的时间。实验设计有几种策略涉及多目标特征选择和结果呈现。为了确保识别系统的准确性,最终的识别结果在多次实验中取平均值。所采集手语数据集中的 80% 为样本数据集,20% 为测试数据集。

#### 4 结束语

提出了一种基于深度信息的静态手势识别的方法。提出一种优化方案通用评估数据库中几种特征提取方法对手势的识别率和 CPU 时间的手势识别算法,其中包含了手势大小、光照和旋转变化等因素的影响。从识别精度和速度两方面,综合比较了如 Hu 矩、Zernike 矩、伪 Zernike 矩、傅里叶描述符和 Gabor 特征等几种常见的图像特征描述符。手势识别采用多层感知器,具有结构灵活、识别速度快的特点。为了提高识别精度和减少计算量,特征向量和神经网络均通过基于 NSGA-II 的多目标进化算法进行调整。在对手语识别的进一步探究中,对提出方法的有效性进行了验证。基于深度图像的方法也更容易推广应用到其他计算机视觉技术领域,具有一定的通用性。

参考文献:

- [1] 侯亭亭,肖秦琨,杨永侠.基于动态贝叶斯网络的手势识别[J].国外电子测量技术,2015,34(1):36-39.
- [2] 赵一丹,肖秦琨,高嵩.基于模糊神经网络和图模型推理的动作识别[J].计算机技术与发展,2018,28(8):80-85.
- [3] GUO Yun,LIU Chunping,GONG Shengrong.Improved algorithm for Zernike moments[C]//International conference on control automation and information sciences.Changshu,China: IEEE,2015:307-312.
- [4] 吴良圆,魏书宁,周棒棒,等.基于 PCA 降维的分层超限学习机手势识别方法[J].电子测量技术,2017,40(3):82-88.
- [5] 蒲兴成,王涛,张毅.基于改进 Hu 矩算法的 Kinect 手势识别[J].计算机工程,2016,42(7):165-172.
- [6] CHAUDHARY A,RAHEJA J,DAS K,et al.Intelligent approaches to interact with machines using hand gesture recognition in natural way: a survey [J].International Journal of Computer Science & Engineering Survey,2011,2(1):122-133.
- [7] HASAN H,KAREEM S.Static hand gesture recognition using neural networks [J].Artificial Intelligence Review,2014,41(2):147-181.
- [8] AOWAL M A,ZAMAN A S,RAHMAN S M M,et al.Static hand gesture recognition using discriminative 2D Zernike moments[C]//IEEE region 10 conference.Bangkok,Thailand: IEEE,2015:1-5.
- [9] LEUNG F H F,LAM H K,LING S H,et al.Tuning of the structure and parameters of a neural network using an improved genetic algorithm [J].IEEE Transactions on Neural Networks,2003,14(1):79-88.
- [10] 何阳清,葛元,王林泉.应用几何矩和边缘检测的手势识别算法[J].计算机工程,2005,31(15):165-166.
- [11] ZERNIKE V F.Beugungstheorie des schneidenver-fahrens und seiner verbesserten form der phasenkontrastmethode[J].Physica,1934,1(7):689-704.
- [12] CHONG C W,RAVEENDRAN P,MUKUNDAN R.An efficient algorithm for fast computation of Pseudo-Zernike moments [J].International Journal of Pattern Recognition & Artificial Intelligence,2003,17(6):1011-1023.
- [13] ZHANG Dengsheng,LU Guojun.A comparative study of curvature scale space and Fourier descriptors for shape-based image retrieval [J].Journal of Visual Communication & Image Representation,2003,14(1):39-57.
- [14] 付余伟.基于非高斯 Gabor 滤波器的掌纹识别算法研究[D].昆明理工大学,2017.
- [15] HUANG D Y,HU W C,CHANG S H.Vision-based hand gesture recognition using PCA+gabor filters and SVM [C]//5th international conference on intelligent information hiding and multimedia signal processing.Kyoto,Japan: IEEE,2009:1-4.
- [16] 路艳雪.NSGA-II 多目标优化算法的改进及应用研究[D].太原:太原理工大学,2017.
- [17] SEADA H,DEB K.U-NSGA-III: a unified evolutionary optimization procedure for single,multiple,and many objectives: proof-of-principle results [C]//International conference on evolutionary multi-criterion optimization.[s.l.]: [s.n.],2015:34-49.
- [18] KOHAVI R,JOHN G H.Wrappers for feature subset selection [J].Artificial Intelligence,1997,97(1-2):273-324.