

# 手势分割问题的多目标优化模型及其进化求解方法

刘可, 巩敦卫<sup>†</sup>

(中国矿业大学 信息与电气工程学院, 江苏 徐州 221116)

**摘要:** 提出一种手势分割问题的多目标优化模型, 并给出该模型的进化求解方法. 建立模型时, 以像素点的位置作为决策变量, 以像素点的颜色与人手肤色的差值作为目标函数. 此外, 根据手部像素点的位置具有相关性, 建立多目标分布估计算法来求解上述模型, 以得到最佳的像素点集, 从而形成人的手势. 实验结果表明了所提出模型和方法的有效性.

**关键词:** 多目标优化; 分布估计算法; 手势分割; 进化计算

**中图分类号:** TP273

**文献标志码:** A

## Multi-objective optimization model and its evolution-based solution for gesture segmentation problems

LIU Ke, GONG Dun-wei<sup>†</sup>

(School of Information and Electrical Engineering, China University of Mining and Technology, Xuzhou 221116, China)

**Abstract:** A multi-objective optimization model for the problem of gesture segmentation is proposed, and a method of solving the above model based on evolutionary algorithms is presented. When building the model, the positions of a series of pixels are taken as the decision variable, and the differences between the color of the pixels and that of a hand are taken as the objective functions. In addition, a multi-objective estimation of the distribution algorithm is presented based on the correlation among the positions of the hand pixels to solve the model so as to get the best pixel set, thus forming the human gesture. The experimental results show the effectiveness of the proposed model and algorithm.

**Keywords:** multi-objective optimization; estimation of distribution algorithm; gesture segmentation; evolutionary computation

### 0 引言

基于肤色的图像分割, 是解决手势分割问题的主要方法之一. YCbCr 是肤色分割中常用的颜色空间<sup>[1]</sup>, 该空间中的 Y 分量表示亮度, Cb 和 Cr 分量表示色度. 文献[2]指出, 存在最佳匹配肤色的 Cb 或 Cr 数值. 基于此, 如果从所有像素点中选择最接近肤色的, 则有可能将手势与背景分割开, 这是一个优化问题. 许多学者采用进化算法求解优化问题<sup>[3-4]</sup>; Deb 等<sup>[5]</sup>提出了 NSGA-II; Zhang 等<sup>[6]</sup>提出了 RM-MEDA.

基于肤色的图像分割, 首先需要选择颜色空间; 然后, 通过训练样本集建立肤色模型, 进而区分出肤色与非肤色像素点. 在文献[7]中, 肤色的 Cb、Cr 分量需同时满足下列条件:  $77 \leq Cb \leq 127$ ,  $133 \leq Cr \leq 173$ .

在文献[8]中, 肤色的 Cb、Cr 分量分布在与 Y 分量有关的区域内; 如果像素点的 HSV 向量分布在一个三维空间内, 则可以识别为肤色像素点. 文献[9]采用单峰高斯模型, 表示 CbCr 平面上的肤色分布. 文献[10]采用高斯混合模型, 表示 CbCr 平面和 RGB 空间中的肤色分布. 文献[11]对离散化的颜色空间中的每个单元赋予一个肤色概率值, 建立肤色的直方图模型. 但是, 肤色在颜色空间中的分布具有随机性, 对于出现概率较小的手部颜色, 已有的肤色分割方法存在局限性.

为了能够识别出现概率较小的手部颜色, 本文建立手势分割问题的多目标优化模型, 并给出求解该模型的方法. 建立模型时, 以像素点的位置为决策变量,

收稿日期: 2015-09-16; 修回日期: 2016-01-04.

基金项目: 国家自然科学基金项目(61403155).

作者简介: 刘可(1979—), 男, 博士生, 从事进化计算的研究; 巩敦卫(1970—), 男, 教授, 博士生导师, 从事智能优化与控制、基于搜索的软件工程、交互式遗传算法、区间多目标进化计算等研究.

<sup>†</sup>通信作者. E-mail: dwgong@vip.163.com

以像素点的颜色与人手肤色的差值为目标函数. 此外, 根据手部像素点的位置具有相关性, 建立多目标分布估计算法来求解上述模型, 以得到最佳的像素点集, 以此形成人的手势, 从而解决手势分割问题. 对于出现概率较小或较大的手部颜色, 本文方法均能够进行识别. 因此, 与已有的肤色分割方法相比, 本文方法具有一定的优势.

## 1 手势分割的多目标优化模型

对于一幅手势图像, 根据该图像中像素点的Cb和Cr分量, 选取一定数量的像素点, 可以把手部区域从手势图像中分割出来. 为了选取手势图像中位于手部区域的若干像素点, 建立如下多目标优化模型:

$$\min \mathbf{F}(x) = (f_1(x), f_2(x)); \text{ s.t. } x \in 2^X. \quad (1)$$

其中: 决策空间  $X$  为图像中所有像素点位置的集合,  $2^X$  为  $X$  的幂集; 决策变量  $x$  为选取的不重合像素点位置的集合,  $x = (x^1, x^2, \dots, x^n)$ ,  $x^i = (x_i, y_i)$ ,  $i = 1, 2, \dots, n$ ,  $n$  为选取的像素点个数.

现在给出目标函数  $f_1(x)$  和  $f_2(x)$  的数学表示.  $f_1(x)$  使得像素点的Cb分量接近于最佳匹配肤色的Cb分量, 其表达式为

$$f_1(x) = \sum_{i=1}^n |\text{Cb}(x^i) - \overline{\text{Cb}}|/n. \quad (2)$$

其中:  $\text{Cb}(x^i)$  为  $x^i$  的Cb分量,  $\overline{\text{Cb}} = 109.38$  为最佳匹配肤色的Cb分量<sup>[8]</sup>.

目标函数  $f_2(x)$  使得像素点的Cr分量接近于最佳匹配肤色的Cr分量, 其表达式为

$$f_2(x) = \sum_{i=1}^n |\text{Cr}(x^i) - \overline{\text{Cr}}|/n. \quad (3)$$

其中:  $\text{Cr}(x^i)$  为  $x^i$  的Cr分量,  $\overline{\text{Cr}} = 152.02$  为最佳匹配肤色的Cr分量<sup>[8]</sup>.

通过求解优化模型(1), 能够得到一系列像素点的位置; 根据这些像素点的Cb和Cr分量, 形成它们的取值区间; 根据某像素点的Cb和Cr分量是否在上述区间内, 确定该像素点是否属于手部区域, 从而将手部区域从手势图像中分割出来.

## 2 用于手势分割的分布估计算法

本节采用分布估计算法求解模型(1), 以得到一系列像素点的位置. 为此, 首先, 基于手部像素点位置的相关性, 建立候选解的概率分布模型; 然后, 通过在搜索域中对该模型高斯采样, 得到一个规模为  $N$  的临时种群; 最后, 根据序值和拥挤距离<sup>[5]</sup>, 从规模为  $N$  的父代种群和临时种群形成的合并种群中选择优势个体, 得到规模为  $N$  的子代种群. 重复上述过程, 直到满足算法终止条件, 得到一或多个非被占优解. 可以看出, 建立候选解的概率分布模型是模型求解的关键.

鉴于此, 下面给出概率分布模型的建立方法.

考虑模型(1)的候选解  $x$ , 它包含  $n$  个像素点, 每一像素点可以用其位置唯一表示. 如果  $x$  是模型(1)的优化解, 则基于该解, 能够形成一个手部区域. 由于形成手部区域的像素点与相应的手指对应, 可以按照手指的位置将这些点分成5类, 且每一类与一个手指位置对应.

现在考虑某一类像素点, 该类像素点与某一手指对应. 将该手指沿手掌方向延伸, 直至手掌根部, 能够近似形成一条直线, 于是这类像素点将分布在该直线两侧. 这样, 候选解  $x$  包含的像素点将分布在5条直线两侧. 记这5条直线分别为  $l_1, l_2, \dots, l_5$ , 分布在这些直线两侧的像素点形成的集合分别为  $A_1, A_2, \dots, A_5$ .

为了确定  $x$  包含的像素点所属的集合, 采用一维局部主分量分析法<sup>[6]</sup>将  $x$  中的像素点聚类成5类  $A_j, j = 1, 2, \dots, 5$ .

对于模型(1)的多个解, 利用非支配排序方法<sup>[5]</sup>可以选出一或多个优化解, 该优化解包含的所有像素点位于手部区域内, 因此, 模型(1)的优化解集包含的所有不重合位置点也位于手部区域内. 优化解集包含的所有不重合位置点分布在手部区域以内  $l_j$  的两侧. 根据文献[6],  $A_j$  的像素点在直线  $l_j$  上投影落在某一直线段  $\varphi_j$  上,  $\varphi_j = c_j + \theta u_j$ ,  $\underline{d}_j \leq \theta \leq \overline{d}_j$ ,  $c_j = \sum_{x^i \in A_j} x^i / |A_j|$ ,  $u_j$  为  $V_j$  的最大特征值对应的特征向量, 而

$$V_j = \sum_{x^i \in A_j} (x^i - c_j)(x^i - c_j)^T / |A_j| - 1,$$

$$\underline{d}_j = \min_{x^i \in A_j} d^i, \overline{d}_j = \max_{x^i \in A_j} d^i, d^i = (x^i - c_j)^T u_j.$$

于是, 直线段  $\varphi_1, \varphi_2, \dots, \varphi_5$  即为生成  $x$  各分量的概率分布模型.

根据上面给出的概率分布模型, 在决策空间中进行搜索多次高斯采样<sup>[6]</sup>得到多个  $x^i$ ,  $n$  个互不重合的  $x^i$  形成一个候选解,  $N$  个候选解形成一个临时种群.

## 3 实验研究

### 3.1 实验过程

本文采用文献[12]的图像库进行实验. 图像文件名称由 hand1\_0\_bot\_seg\_1\_cropped, hand1\_0\_bot\_seg\_2\_cropped, ..., hand5\_z\_dif\_seg\_5\_cropped 改为 1, 2, ..., 2515. 将本文所提出方法与两种手势分割方法<sup>[7-8]</sup>进行比较. 在文献[7]中, 肤色在 YCbCr 空间的Cb、Cr分量需同时满足下列条件:  $77 \leq \text{Cb} \leq 127$ ,  $133 \leq \text{Cr} \leq 173$ . 在文献[8]中, 肤色在 YCbCr 空间的Cb、Cr分量分布在限定的区域内. 对于每一种方法分割的手部区域, 定义正确检测到的肤色像素点数量与真实的肤

色像素点数量的比为正检率;检测到肤色的非肤色像素点数量与真实的肤色像素点数量的比为误检率.这样,对于图像库的每一幅图像,都能得到每种方法的正检率和误检率.通过比较正检率和误检率,能够反映本文方法与两种已有方法的图像分割效果.

在手势图像中,如果组成背景图像元素中存在某些特殊元素,则这些特殊元素的颜色与人手的颜色特征非常接近,这一背景可称为复杂背景.在手势分割中,把复杂背景与肤色分割开具有一定难度.如果把一些复杂背景区域识别为手部区域,则会增加手势分割结果的误检率.为了研究本文方法的误检率,对于图像库的每一个手部区域,都在手部区域以外填充了背景图像.背景图像的Cb和Cr分量数值分别是70和200,它们接近肤色的Cb和Cr分量数值,并且文献[7]和文献[8]的方法可以把背景图像与肤色分割开.这样,所有手势图像的尺寸均是 $800 \times 800$ ,在图像中手部区域居中,周围是填充的背景图像.由于背景图像的Cb和Cr分量数值接近肤色的Cb和Cr分量数值,如果背景图像的Y(亮度)分量数值是一般情况下的数值,则背景图像的颜色均与人手的颜色特征非常接近.由此可见,上述背景图像属于复杂背景.

由于图像库中的所有手部区域都有绿色边缘,绿色边缘的Cr分量一般小于等于124,本文借此把绿色边缘从手部区域中分割出,进而得到准确的2515个手部区域.对图像库中的2515个手部区域均进行背景图像填充.对于手势图像133.png、1024.png、1838.png、2063.png、2506.png的手部区域,填充了背景图像后,得到图1(a)~图1(e).它们的手部区域分别来自5个人,其中背景图像的Y分量数值是140.

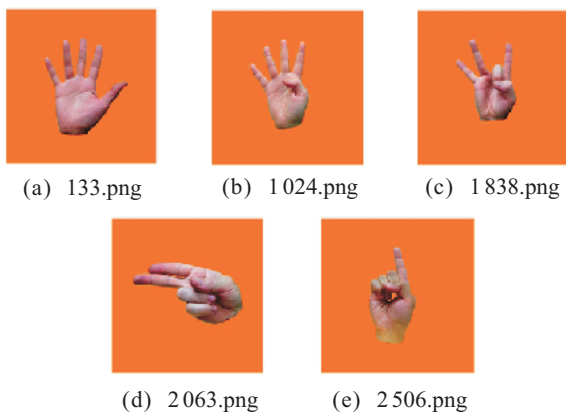


图1 5幅手势图像

对于本文提出的方法,首先,对手势图像进行预处理,得到像素点的Cb和Cr分量数值;其次,建立模型(1);再次,采用分布估计算法求解模型(1),得到一个或多个非被占优解;最后,根据最优解集确定人手的Cb和Cr取值区间,进而分割出手部区域.

选取像素点的坐标数量作为决策变量的维数.考

虑到维数过大将会导致模型求解的复杂度增加,延长求解时间,因此,对上述图像应进行必要的压缩,通过每间隔8行和8列进行像素点采样,以减小手势图像的尺寸.压缩后,每幅图像包含的像素点个数为 $100 \times 100$ .对于压缩后的图像,通过YCbCr颜色空间对图像依次处理,得到图像中所有像素点的Cb和Cr分量数值,从而完成手势图像的预处理.

在本文算法中,通过对概率分布模型进行高斯采样得到候选解,因此,概率分布模型对应的优化解集容易陷入局部最优.决策变量的维数增大,将使出现这种情况的可能性增大.根据模型(1),决策变量应代表一定数量的像素点.这样,实验中采用100维向量作为决策变量,通过决策变量 $x$ 能够得到代表50个像素点位置和它们的Cb和Cr数值,进而得到对应的 $f_1(x)$ 和 $f_2(x)$ .

利用本文算法求解上述优化模型可得到一个或多个非被占优解,或者说得到一个最优解集.这个最优解集对应一个不重合像素点位置的集合.多次求解模型(1)可以得到多个最优解集.本文定义,在最优解集对应的像素点集中,手部像素点的数量与像素点总数量之比是最优解集的正确率.在本文算法中需要设置的参数有:种群规模和进化代数.在进行进化计算的实验时,种群规模分别取30、40、50、60、100,进化代数分别取60、80、100、120、150.因此,共25种算法参数的设置.在每一种参数设置下,对于任一幅手势图像,都进行20次求解模型(1)的实验.实验结果表明,种群规模从30增大到50时,正确率的平均值从67%迅速提高到94%;种群规模从50增大到100时,正确率没有明显提高,但是实验所需时间的平均值从6s明显增大到22s.进化代数从60增大到100时,正确率的平均值从76%迅速提高到93%;进化代数从100增大到150时,正确率没有明显提高,但是实验所需时间的平均值从7s明显增大到16s.对于本文所提出的手势分割方法,最优解集的正确率越大,手势分割的结果越准确.一般情况下,为了实现手势的快速识别,在保证手势分割准确的前提下,手势分割的图像处理过程越快越好.这就要求在实验中,在保证进化计算准确的前提下,计算所需的时间尽量少.考虑到进化计算中正确率与所需时间的关系,在下面的实验中,种群规模取50,进化代数取100.

一般来说,进化算法均具有随机性,多次运行的计算结果不一定相同.本文所提出算法属于进化算法,因此,利用所提出算法求解模型(1),多次计算得到多个最优解集,这些最优解集一般互不相同.如果这些最优解集互不相同,则它们代表的不重合像素点位置的集合也可能互不相同.或者说,多次运行本文算法

求解模型 (1), 得到的多个不重合像素点位置的集合不一定相同. 由于最优解集和其像素点集均具有随机性, 本文通过 20 次进化计算得到 20 个最优解集, 并得到 20 个最优解集对应的像素点位置, 以及所有像素点位置重复选取的次数. 像素点位置重复选取的次数越大, 表明其误差越小. 根据文献 [12] 提供的图像库, 估计压缩后的图像中, 手部区域像素点的个数应大于 500. 这样, 选择的像素点个数应大于 500. 为了减小误差, 应选择重复次数较大的像素点位置作为手部像素点. 大于某一重复次数的像素点数量, 随着重复次数的增加而减少. 本文在重复次数大于 1 的像素点集中, 选择大于某一重复次数的像素点集, 对应的像素点数量大于 500 并且最接近 500. 如果重复次数大于 1 的像素点集, 对应的像素点数量小于 500, 则选择该像素点集.

最后, 根据选择的像素点位置得到手势图像的 Cb、Cr 取值区间, 并以此分割手部区域. 实验采用的电脑配置是: Intel Core i3-3240, 3.40 GHz CPU, 1.85 GB 内存.

### 3.2 结果与分析

本文对填充了背景的手势图像分别进行 20 次计算, 得到 20 个最优的像素点集. 对于图 1(a) 的 20 个像素点集, 能够得到所有像素点位置重复选取的次数, 结果如图 2(a) 所示. 对于图 1(b)~图 1(e), 执行相同操作, 结果如图 2(b)~图 2(e) 所示. 对于图 2, 选择的像素点位置如图 3 所示.

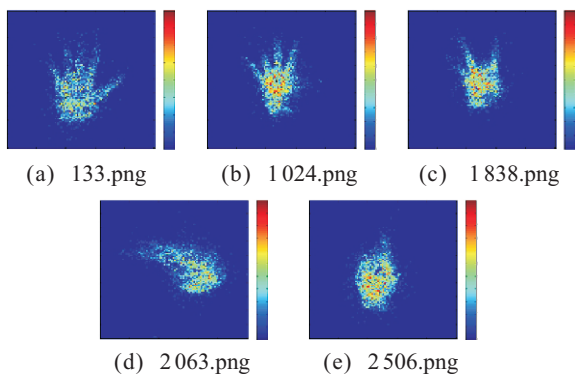


图 2 对应于图 1 的像素点位置重复选取的结果

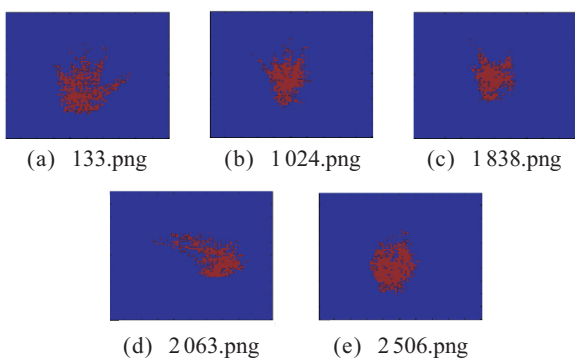


图 3 根据图 2 选择的像素点

根据图 3(a)~图 3(e) 所示的像素点位置, 对于图 1(a)~图 1(e), Cb 的取值区间依次是 110-133、104-129、112-136、103-137 以及 90-133; Cr 的取值区间依次是 133-167、128-176、132-168、127-171 以及 129-174; 以此分割手部区域, 结果的正检率依次是 99.82%、99.70%、98.29%、98.61%、99.42%, 误检率均是 0.

对于已填充背景的 2515 个手势图像, 采用本文方法进行计算, 得到手部区域 Cb 取值下界和上界, 以及 Cr 取值下界和上界. 以此分割手部区域, 计算结果的正检率如图 4(a) 所示. 图 4(a) 中, 横坐标是手势图像文件的名称, 纵坐标是其正检率. 2515 个手势图像的计算结果中, 有 112 个存在误检率且数值较大; 它们的 Cb 取值下界是 70, Cr 取值上界是 200, 正检率接近 100%, 如图 4(a) 所示. 这表明, 112 个计算结果没有把手与背景分割开. 对于把手与背景分割开的 2403 个计算结果, 正检率的平均值是 96.16%, 误检率均是 0. 采用本文方法完成一幅图像的手势分割实验, 所需的时间为 4~5 s.

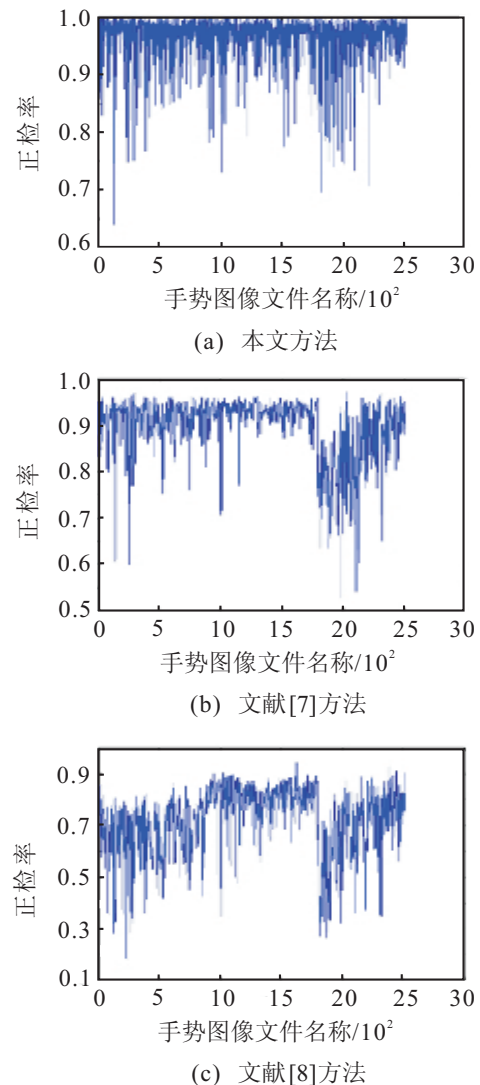


图 4 3 种方法手势分割结果的正检率

本文分别根据文献[7]和文献[8]的方法,在已填充背景的2515个手势图像中分割手部区域,计算结果的正检率如图4(b)和图4(c)所示.图4(b)和图4(c)中,横坐标是手势图像文件的名称,纵坐标是其正检率.对于2515个手势图像,文献[7]和文献[8]的方法误检率均为0,两种方法的所有计算结果均把人手与背景分割开.对于文献[7]的方法,2515个正检率的平均值是89.93%;对于文献[8]的方法,2515个正检率的平均值是71.52%;本文方法正检率的平均值高于文献[7]和文献[8]的方法.

对于图4(a)和图4(b)所示的实验结果,两种方法处理同一幅手势图像可以得到一对正检率,共2515对正检率,其中任意一对正检率均存在一定的差异.本文使用假设检验对这些差异的显著性进行验证.对于图4(a)和图4(b)所示的任意一对正检率,本文采用非参数检验中的Wilcoxon带符号秩检验方法进行两两比较,显著性水平取0.05.在2287对正检率中,图4(a)的正检率优于图4(b)的正检率;在另外的228对正检率中,图4(b)的正检率优于图4(a)的正检率.假设检验结果表明,图4(a)的正检率显著优于图4(b)的正检率.对于图4(a)和图4(c)所示的实验结果,两种方法处理同一幅手势图像也可以得到一对正检率,共2515对正检率,其中任意一对正检率均存在一定的差异.对于图4(a)和图4(c)所示的任意一对正检率,本文采用同样的假设检验方法对其差异的显著性进行验证,显著性水平取0.05.在2500对正检率中,图4(a)的正检率优于图4(c)的正检率;在另外的15对正检率中,图4(c)的正检率优于图4(a)的正检率.假设检验结果表明,图4(a)的正检率显著优于图4(c)的正检率.由此可见,本文所提出的方法优于文献[7]和文献[8]的方法.

#### 4 结 论

本文提出了手势分割问题的多目标优化模型,并给出了求解该模型的进化优化方法.建立模型时,以像素点的位置作为决策变量,以像素点的颜色与人手肤色的差值作为目标函数.此外,根据手部像素点位置的相关性,建立了求解该模型的多目标分布估计算法,以得到最佳的像素点集,并以此形成人手,从而解决手势分割问题.将所提出方法应用于实际的人手分割问题,并与已有方法比较,实验结果表明,该方法能够很好地解决手势分割问题.

#### 参考文献(References)

- [1] Attila L, Tamas S. User adaptive hand gesture recognition system with interactive training[J]. *Image and Vision Computing*, 2005, 10(2): 1102-1114.
- [2] Zhi C, Jung K, Jian L, et al. Real-time hand gesture recognition using finger segmentation[J]. *The Scientific World Journal*, 2014, 14(3): 267872.
- [3] Deb K, Jain H. An evolutionary many-objective optimization algorithm using reference-point-based nondominated sorting approach, Part I: Solving problems with box constraints[J]. *IEEE Trans on Evolutionary Computation*, 2014, 18(4): 577-601.
- [4] Sanghamitra B, Arpan M. An algorithm for many-objective optimization with reduced objective computations: A study in differential evolution[J]. *IEEE Trans on Evolutionary Computation*, 2015, 19(3): 400-413.
- [5] Deb K, Pratap A, Agarwal S, et al. A fast and elitist multi-objective genetic algorithm: NSGA-II[J]. *IEEE Trans on Evolutionary Computation*, 2002, 6(2): 182-197.
- [6] Zhang Q F, Zhou A M, Jin Y. RM-MEDA: A regularity model based multiobjective estimation of distribution algorithm[J]. *IEEE Trans on Evolutionary Computation*, 2007, 12(1): 41-63.
- [7] Chai D, Ngan K N. Face segmentation using skin color map in videophone applications[J]. *IEEE Trans on Circuits and Systems for Video Technology*, 1999, 9(4): 551-564.
- [8] Garcia C, Tziritas G. Face detection using quantized skin color regions merging and wavelet packet analysis[J]. *IEEE Trans on Multimedia*, 1999, 1(3): 264-277.
- [9] Michael J J, James M R. Statistical color models with application to skin detection[J]. *Int J of Computer Vision*, 2002, 46(1): 81-96.
- [10] Phung S L, Bouzerdoum A, Chai D. Skin segmentation using color pixel classification: Analysis and comparison[J]. *IEEE Trans on Pattern Analysis and Machine Intelligence*, 2005, 27(1): 148-154.
- [11] Chai D, Bouzerdoum A. A Bayesian approach to skin color classification in ycbcr color space[C]. *Proc of IEEE Region Ten Conf*. Singapore: IEEE, 2000: 421-424.
- [12] Barczak A L C, Reyes N H, Abastillas M, et al. A new 2D static hand gesture colour image dataset for ASL gestures[J]. *Research Letters in the Information and Mathematical Sciences*, 2011, 15(1): 12-20.

(责任编辑:李君玲)