文章编号 2095-1531(2015)04-0582-07

# 基于人脸姿态估计的虚拟眼镜试戴技术

卢 洋,王世刚*,赵文婷,武 伟
(吉林大学 通信工程学院,吉林 长春 130012)

**摘要**:为解决用户线上眼镜的最优选购,提出了一种基于人脸姿态估计的虚拟眼镜试戴技术。首先采用肤色模型与形状模型结合的算法对场景中的人脸区域进行检测,然后根据人眼在人脸中的几何位置关系实现人眼的精确定位;进一步利用人眼对称性先验知识来估计脸在三维空间中的姿态信息,即人脸与正面的偏移角度;最后,依据人眼位置和人脸姿态将眼镜图像融合到眼睛区域,即完成眼镜的虚拟试戴。该方法为3D环境下客户与商品之间虚拟视觉化的实现提供了一种可靠的技术支撑和应用思路。

**关 键 词**:姿态估计;图像融合;人脸检测;人眼定位
**中图分类号**:TP391.4 **文献标识码**:A doi:10.3788/CO.20150804.0582

# Technology of virtual eyeglasses try-on system based on face pose estimation

LU Yang, WANG Shi-gang*, ZHAO Wen-ting, WU Wei
(*College of Communication Engineering*, *Jilin University*, *Changchun* 130012, *China*)
* *Corresponding author*, *E-mail*: *wangshigang@vip.sina.com*

**Abstract**: In this paper, we present a new virtual eyeglasses try-on system based on face pose estimation, and provides technical support for the realization of the user online optimal commodity purchase. Firstly, we use color model combining shape model for face detection in the scene. Then, we use face geometry relationship to achieve precise positioning of the human eye. We furtherly estimate the face pose information based on a priori knowledge of the symmetrical of eyes in the face region, which is the deviation angle of face with the front side. Finally, we integrate the images of glasses to face and achieve the virtual eyeglasses try-on system according to calculating the angle and location information of eyes. This method provides a reliable technical support and application for the future implementation of virtual visualization between customers and goods in 3D environment.

**Key words**: pose estimation; image fusion; face detection; eye location

**收稿日期**:2015-03-06;**修订日期**:2015-04-20
**基金项目**:教育部博士点基金资助项目(No. 20120061110091)

## 1　引　言

随着互联网和电子商务技术的快速发展,网上购物受到更多人的青睐,越来越多商家把在线销售作为实体店销售的一种补充方式。传统的网络购物方式已满足不了消费者的需求,最主要的问题是消费群体无法看到本人试穿的效果,导致判断错误,造成不必要的经济开销。人们传统购物的"真实性"理念与网络购物的"虚幻性"特点不能得到有机的统一。而在线网络试穿、虚拟试戴等技术的出现弥补了网络购物的缺陷,使网络购物更加真实化、人性化,成为近年来机器视觉研究领域的热点[1]。如何真实、快速地实现虚拟眼镜试戴是本文研究的主要目的。

以往的人脸检测技术仅仅停留在判断目标视频序列或图像中是否存在人脸,以及进一步确定存在人脸的位置及大小等。为了实现虚拟眼镜试戴,仅检测到人脸是远远不够的,还需要获取眼睛的轮廓信息。解决此类问题需要对人脸的特征点进行定位,它是研究图像中人脸的前提。而人眼定位是眼镜虚拟试戴的关键技术,大多研究方法集中于对正面人脸的检测以及人眼定位。本文针对不同姿态人脸的估计提出了一种几何方法,能够准确地估计出人脸的姿态,从而实现眼镜的虚拟试戴。

## 2　人脸检测及定位

人脸检测作为人脸信息识别中一项极为关键的技术,在生物信息识别、安防监测中都得到了广泛的应用,由于其重要性和必需性,近年来在机器视觉、行为分析等课题领域得到了普遍的重视[2-3]。人脸检测的核心内容就是找寻人脸区域区别于其他区域的特征,人脸最基本的特征有颜色特征(肤色、发色、唇色、灰度),几何特征(轮廓、五官、对称性),统计特征(均值、方差、相关系数、直方图、镶嵌图特征)等。目前已有的人脸检测算法有快速检测算法和基于模板的方法[4]。前者的主要算法基于先验知识的方法,后者的主要算法基于后验知识学习和训练的方法。

由于论文的目标是实现不同姿态下人脸图像的检测,因此算法要求对人脸的旋转具有鲁棒性,故本论文采用基于先验知识的方法进行人脸检测。人脸区域中,肤色是占主导地位的像素值,由于它很好的适应性,可以适应人脸各种角度的旋转;并且不受人脸面部表情以及细节特征的改变而改变,具有很好的稳定性,所以本论文选择肤色模型来进行人脸检测。考虑到 RGB 色彩空间内人脸肤色受光照影响严重,并且肤色与非肤色的重叠性较大[6-7],而 YIQ 颜色空间利用人眼的色分能力,使人眼对不同的色彩具有很明显的区分度;并且 YIQ 色彩空间在处理彩色信息时冗余度较少。综上,本论文选择基于 YIQ 色彩空间建立肤色模型进行人脸检测。

在 YIQ 彩色模型中,Y 是指颜色的亮度,而 I 和 Q 是指色调,用于描述图像色彩的饱和度。RGB 与 YIQ 之间的对应关系如式(1):

$$\begin{bmatrix} Y \\ I \\ Q \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ 0.596 & -0.274 & -0.322 \\ 0.211 & -0.523 & 0.312 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}. \tag{1}$$

由此定义肤色模型:对于彩色图像中的像素点 $p$,将其从 RGB 模型变换到 YIQ 模型,如果满足不等式(2),则点 $p$ 为肤色像素,反之为非肤色像素[5]。

$$\frac{(I_p - m)^2}{r^2} - \frac{Q_p^2}{q^2} \leqslant 1 \ , \tag{2}$$

式中,$I_p$ 和 $Q_p$ 分别代表 $p$ 点对应的 $I$ 分量和 $Q$ 分量,$r$ 和 $q$ 分别代表 $I$ 分量和 $Q$ 分量的范围,$m$ 代表 $I$ 分量的均值。

以 $I$ 代表像素分量,选择的 YIQ 空间的肤色范围为(5,150),对图 1(a)运行结果如图 1(b)所示。然后根据面积约束剔除分割出的小目标区域;最后将潜在所有的连通区域提取出来,计算其横向对称度、纵向对称度、椭圆度等特征,融合 3 个特征计算每个区域的得分,取得分最大的区域为人脸区域,从而得到人脸的精确定位,效果如图 1(c)所示。


(a) 原始图像 (b) 肤色粗定位结果
(a) Origin image (b) Skin coarse location

(c) 基于面积、形状约束的人脸检测结果
(c) Based on the area and shape constraints

图1 人脸检测及定位

Fig. 1 Face detection and location

## 3 人脸姿态估计

在人机交互、行为分析等领域人脸检测已经不仅仅满足于检测出人脸区域,姿态估计得到了越来越多的重视。人脸姿态估计是对目标视频序列或图像中获取人脸在三维坐标系空间中的角度信息。要判断一幅人脸图像的角度信息,就需要确定图像中人脸是否在某一个方向上存在着偏转。而将高维空间上一幅复杂人脸图像的数据与姿态角度概念建立联系则需要进行一系列的处理。因此提出了一种几何方法对人脸姿态进行准确估计。

### 3.1 基于形状和空间位置关系约束的人眼定位

人眼定位是人脸检测中的关键步骤,对姿态估计效果影响很大。多数人脸检测中的第一步就是实现人眼定位,精确的人眼定位为人脸特征值的提取以及人脸信息分析奠定了坚实的基础。本文在进行人眼定位之前,对待检测人眼区域进行平滑滤波操作[6,8]。现有的基于人脸几何特征进行人眼定位的算法仅考虑人脸的几何约束特征,忽略了图像尺寸大小对于定位的影响。故本文针对这点不足采用一种图像归一化的算法,经实验可知此算法对不同尺寸的图像以及不同大小的人眼来说具有普遍适用性。基于形状和空间位置关系约束的人眼定位方法,对于不同尺寸的图像,其约束关系包括眼睛黑色区域面积,眼睛黑色区域的像素个数,眼睛黑色区域下方一定区域不存在其他黑色区域等。

通常,人眼的初定位是根据眼睛在二值化人脸图像中的几何位置确定的,判定通常基于以下几条人脸几何约束条件:

(1)人双眼中心距的限定范围:人的双眼中心距离不是一个随意的值,而是在某一确定的距离范围内波动:就一幅尺寸为160 pixel×120 pixel的图像来说,正常的人双眼距离应在像素个数区间(20,50)内;

(2)人眼下方一定范围内不允许出现其他黑色区域:这意味着人双眼下方不允许存在其他人体器官,因此对应于人脸二值化图像中就是不允许出现其他黑色区域;

(3)眼睛黑色区域像素值个数范围确定:正常的人眼大小大致活动在一个固定的范围,故在对应的二值化人脸图像中眼睛黑色区域的大小也应在某个确定范围内波动,过大的黑色区域不可能是人眼;

(4)人眼黑色区域外接矩形的形状要求宽大于或者等于高:从生理学可知,由于人眼的特定形状是椭圆或近似圆形,故其外接矩形的形状必定是宽大于或者等于高,所以高远大于宽的外接矩

(a) 人脸几何结构图 (b) 脸部区域YIQ肤色检测
(a) Figure structure of face (b) Skin colour test of face in YIQ

(c) 眼睛定位结果
(c) Test result of eye position

图2 人眼定位

Fig. 2 Eye location

形对应的黑色区域不可能是眼块。

根据图 1(c)脸的定位结果,在人脸区域使用 YIQ 肤色模型进一步肤色检测,可知:人脸中的眉毛、眼睛全都不属于肤色范围,所以在二值化图像中以黑洞的形式存在。然后结合人脸的几何分布结构,找到额头所在的位置,向下偏移一定的距离寻找对称椭圆区域即可定位到眼睛。如图 2 所示,为人眼定位各部分结果图。

经实验测试分析,经过以上的人眼约束条件,人眼检测的准确率高达 90% 。

### 3.2　人脸姿态估计的数学描述

假设不同姿态的人脸图像已经被定位,并将图像缩放为宽为 $W$ 高为 $H$。这些像素点的灰度值组成了一个 $W \times H = p$ 维的图像向量 $x$,$x$ 所在的空间被称为外观空间 $R^p$。因为人脸在旋转过程中有 3 个自由度:左右摆动、上下摆动、图像内旋转,因此定义自由度为向量 $y$,而 $y$ 所在的空间被称为姿态空间 $R^q$,$q \leqslant 3$。人脸外观图像的均值向量为 $x_m$,相应的姿态均值向量为 $y_m$。经过均值归一化的人脸图像样本集合为 $X = \{x_1, \cdots, x_N\}$,相应的姿态样本集合为 $Y = \{y_1, \cdots, y_N\}$。其中 $x_i^{\partial} = x_i - x_m$,$y_i^{\partial} = y_i - y_m$,$N$ 为样本总数[10-11]。

因此,对于一个处于任意姿态的人脸样本 $t$ 来说,姿态估计的任务就是要计算出自由度向量 $y$ 的各个参数值,实现虚拟试戴。为了实现准确且逼真的试戴效果,本论文仅考虑图像内的旋转这个自由度。

### 3.3　基于人眼对称性先验的人脸姿态估计

姿态估计是人脸跟踪、人脸三维模型初始化等应用中至关重要的环节。已存在的姿态估计算法可以分为以下三类:(1)基于几何形状分析:这类方法在建立的人脸模型的基础上将模型与新输入的人脸图像建立对应关系。由于人脸几何结构不考虑不同姿态人脸的可感知差异,如鼻子的偏移和人脸的对称性等,就可以利用精确的人脸特征点定位方法来进行人脸姿态估计;(2)基于流形学习的方法:因为人脸头部运动具有 3 个自由度,将同一目标人物的不同姿态在高维图像空间中构成一个低维流形,进而找到人脸流形在低维空间的嵌入,从而估计出人脸姿态;(3)基于分类器的方法:将姿态估计同化为一个模式分类识别的问题,利用分类器将不同姿态进行分类、学习来对人脸进行姿态估计[9,12]。

本文在基于形状的几何分析方法基础上,提出了一种新的基于人脸对称性先验的姿态估计方法。人脸的形状近似于椭圆形,当人脸只绕着垂直方向左右旋转时,可以近似看成一个刚体,椭圆的一致性保持良好。因此,采用人眼对称性先验的方法进行人脸姿态估计,建立如图 3 所示的椭圆模型。

在图 3(a)中,点 $B$,$C$ 分别表示左眼,右眼所在矩形中的中心点,经过 $BC$ 的直线交椭圆于 $A$,$D$ 两点,做线段 $BC$ 的垂直平分线,交椭圆于 $E$,$F$ 两点,直线 $EF$ 与直线 $AD$ 相交于点 $O$。

令:

$$d_1 = \overline{BO}, \tag{3}$$

$$d_2 = \overline{CO}. \tag{4}$$

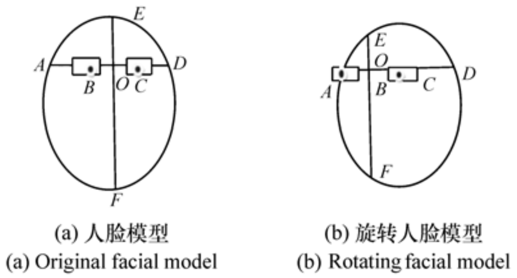


图 3　人脸椭圆

Fig. 3　Elliptical model of face

当人脸绕着 $EF$ 轴向左转动时,$d_1$ 逐渐增大,$d_2$ 逐渐减小;同理,当人脸绕着 $EF$ 轴向右转动时,$d_1$ 逐渐减少,$d_2$ 逐渐增大。令 $\theta$ 表示人脸绕着 EF 直线的转动角度,如果用大量样本建立 $d_1$,$d_2$ 与姿态参数 $\theta$ 之间的对应关系,即可根据未知图像中人脸的位置参数估计出人脸的姿态。由于考虑到眼镜试戴的实际效果,所以转动一定的角度即可,过大的角度会导致试戴效果失真。假设正面的人脸对应的角度为零度,向左偏移表示负方向,右偏表示为正方向。图 4 为人脸在不同角度下的投影。

通过分析不同的人脸图像,发现对于所有正面的人脸图像,两只眼睛的中心点与脸的中位线

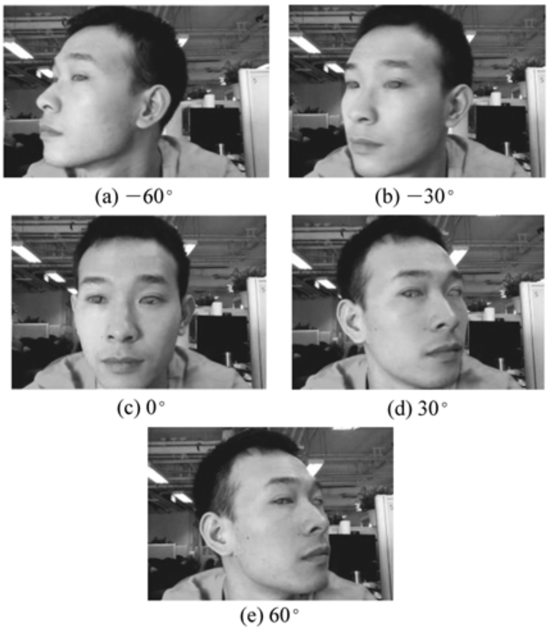


图4　不同姿态的人眼定位效果图

Fig. 4　Eye location rendering of different attitude angle

之间的距离相等，如图5所示。

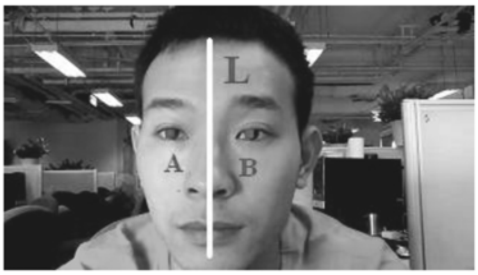


图5　实际人脸结构示意图

Fig. 5　Diagram of actual face structure

在图5中，直线 $L$ 为脸的中位线，$A$ 为左眼中心点，$B$ 为右眼中心点。设点 $A$ 到直线 $L$ 的距离为 $dist_1$，$B$ 到直线 $L$ 的距离为 $dist_2$，定义两眼到中位线距离差 $dist = dist_1 - dist_2$。对于正面人脸图像 $\theta=0$ 一般有 $dist \approx 0$，理想情况下满足 $dist=0$；当脸往左边偏转时 $\theta<0$，则 $dist>0$；往右偏转时 $\theta>0$，则有 $dist<0$。因此，人脸偏移角度 $\theta$ 可表示为 $\theta=f(dist)$。函数 $f(x)$ 一般为非线性函数，为了模拟出该函数，采集到不同角度的人脸样本，其角度和距离差的关系曲线如图6所示。

从上图6可以看出，人脸偏移角度 $\theta$ 与距离差 dist 呈现出很强的线性关系，通过最小二乘拟

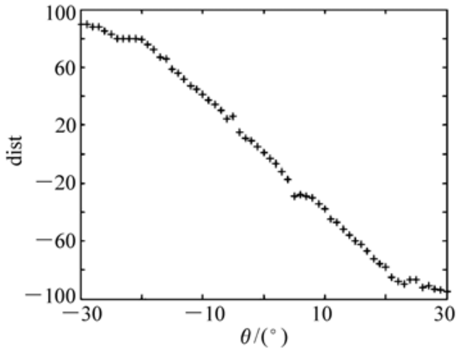


图6　距离差与角度的关系图

Fig. 6　Diagram of distance difference vs. angle

合得到如下关系函数：$\theta=(1.279-dist)/3.577$。依据这种线性关系，准确定位出人眼的位置则快速高效的得到人脸姿态的估计，从而有助于完成最终的眼镜试戴任务。

## 4　抠图处理与实验分析

数字图像抠图问题为数字图像处理中比较典型的一类问题，分为蓝色背景下的蓝屏抠图以及一般背景下的图像抠图。其中，一般背景抠图通常需要 Trimap 模板的输入，用来初步指定一些已知的前景、背景和未知区域。然而，由于抠图问题的欠约束性，以及假设方式的局限性，通常的抠图方法会与真实值存在较大偏差，这也严重影响了图像的最终合成效果。同时，目前抠图领域也缺乏标准且有效的像素级评价系统。本文将问题进行了简化，避免复杂的背景，将眼镜的背景设定为纯色，与镜框和镜片不冲突的颜色。因此，根据定位到的眼睛所在的位置以及估计出的人脸姿态直接将眼镜图像融合到眼睛所在的区域即可完成试戴任务。

实验用到的眼镜图像是通过 OpenGL 实现 $-90°\sim90°$ 偏转范围内的任意姿态的眼镜图像，满足实验要求。

本实验环境为：CPU 为 Intel Core i5，M450，2.4 GHz，内存为 2.00GB 的 Win7 中 MATLAB2013A，每幅图片的平均处理时间为 1.1 s；而将相同的实验图采用 Adaboost 模型和 ANN 模型进行处理，处理时间分别为 1.5 和 1.0 s。

对比于 Adaboost 算法来说,本文的算法在检测时间上有很大的优势,可以实现更快速有效的检测;对比与 ANN 模型算法来说,虽说检测时间略慢于该算法,但是 ANN 模型算法是采用分类器学习的方法进行人脸检测,该算法虽然检测速度快,但是训练速度过慢,所以对于网络大数据情况下实现快速有效的检测是很难的。综上,本文的算法在有效性和快速性上都有更大的优势。

对于眼镜虚拟试戴问题,源图像为眼镜图像库中任意选择的图像,目标图像为用户提供的正面人脸图像,其试戴效果如图 7 所示。


图 7　男士实验结果

Fig. 7　Experiment results for man

如图 7 和图 8 所示,本论文分别选取了两位不同的目标对象进行实验,针对每个目标对象扭转角度的不同,共对 5 个不同旋转角度(分别为:0°、－25°、23°、－23°、27°)的人脸进行检测以及选取相对应的眼镜,进行融合得到最终的试戴效果图。通过不同姿态试戴效果的对比,可以看出本文提出的算法试戴效果很好地保留了眼镜的颜色,且与目标图像融合自然,试戴效果逼真。

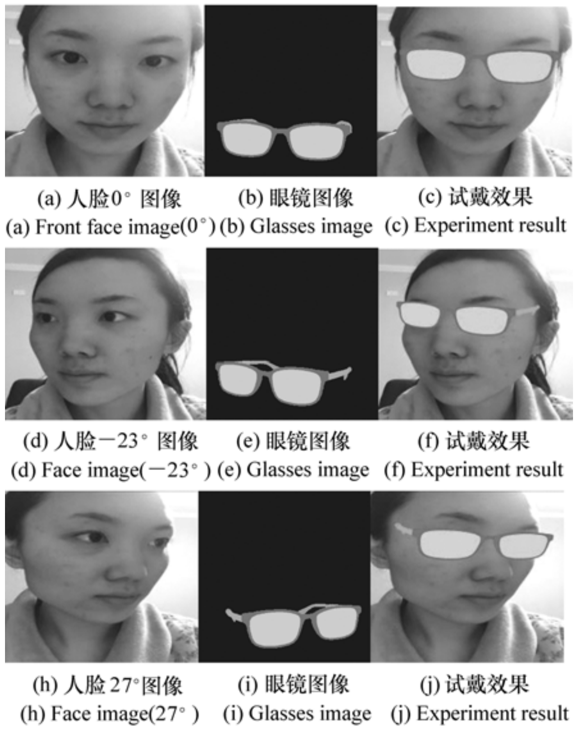


图 8　女士实验结果

Fig. 8　Experiment results for lady

## 5　结　论

本文对人脸检测、定位以及人脸姿态估计进行了研究,提出了基于姿态估计的眼镜虚拟试戴技术。实验结果表明,该算法速度快,在成像距离一定的情况下,检测效率高,试戴的效果自然逼真,具有很好的实用性;但本文依靠的肤色模型目前只能处理黄种人肤色。今后的重点研究工作将会提升人脸的检测率,提高人脸特征点的精度和鲁棒性。图像抠图和图像合成最早出现在影视作品中,如何生动逼真自然的融合图像成为研究的目标。目前,虚拟试戴、试穿的研究甚少,但随着科技的不断发展,其必将具有良好的应用前景。

**参考文献:**

[1]　ZHANG Z,TONG T. Research on computer aided glasses selection and virtual try on[J]. *Computer Engineering and De-*

*sign*,2007,30:45-46.

[2] HJELMAS E. Face detection:a survey[J]. *Computer Vision and Image Understanding*,2001,83:236-274.

[3] HUANG L,JIANG H,HUO G. Face recognition based on multi-scale feature extraction and multiple regression analysis [J]. *Optics & Optoelectronic Technology*,2012,6:90-93.

[4] SHEN L,LANG B,ZHU M. Face recognition method based on artificial neural network[J]. *Chinese J. Liquid Crystal and Displays*,2011,6:836-840.

[5] QUAN X,ZHANG L. Research on the face detection based on skin color model and eye location[J]. *Science Technology and Engineering*,2010,31:7822-7824 +7829.

[6] JI Q,HU R. 3D Face pose estimation and tracking from a monocular camera[J]. *Image and Vision Computing*,2002,20(7):499-511.

[7] NIE X,TAN Z,GUO J. Face illumination compensation based on wavelet transform[J]. *Optics and Precision Engineering*. 2008,1:150-155.

[8] HUANG B,TANG J. Novel method of eye detection in facial images[J]. *Chinese J. Liquid Crystal and Displays*,2009,2:278-282.

[9] LIU M,GUO D,MA J,*et al.*. Face pose estimation based on elliptical model and neural network[J]. *J. Jilin University* (Science Edition),2008,7-11.

[10] LI J,YANG J. Eyeglasses try-on based on improved poisson equation[C]. 2011 International Conference Multimedia Technology. Hangzhou,China,2011:3058-3061.

[11] BLACK M J,YACOOB Y. Tracking and recognizing rigid and non2rigid facial motions using local parametric models of image motion[C]. Fifth International Conference on Computer Vision. Cambridge,MA. USA,IEEE. 1995: 3742381.

[12] SCHWERDT K,CROWLEY J L. Robust face tracking using color[C]. Fourth IEEE International Conference on Automatic Face and Gesture Recognition. Grenoble,Frane,2000:902-905.

**作者简介:**

卢 洋(1991—),女,河南焦作人,硕士研究生,2010年于吉林大学获得学士学位,主要从事图像视频处理方面的研究。E-mail:407851137@qq.com

赵文婷(1989—),女,吉林省吉林市人,博士研究生,2012年于吉林大学获得学士学位,主要从事视频图像处理方面的研究。E-mail:985817166@qq.com

王世刚(1961—),男,吉林长春人,教授,博士生导师,1983年于东北大学获得学士学位,1997年于吉林工业大学获得硕士学位,2001年于吉林大学获得博士学位,主要从事图像与视频信号智能处理方面的研究。E-mail:wangshigang@vip.sina.com

武 伟(1982—),女,吉林长春人,博士研究生,2005年、2009年于长春工业大学分别获得学士、硕士学位,主要从事立体视频方面的研究。E-mail:114188246@qq.com