

在频谱直方图中应用 DT-CWT 的人脸检测技术

孙粤辉,杜明辉

SUN Yue-hui, DU Ming-hui

华南理工大学 电子与信息学院,广州 510640

School of Electronic and Information Engineering, South China University of Technology, Guangzhou 510640, China

E-mail: yuehui.sun@gmail.com

SUN Yue-hui, DU Ming-hui. DT-CWT applied in spectral histogram space for face detection. Computer Engineering and Applications, 2007, 43(20): 86-89.

Abstract: A novel face detection method based on spectral histogram constructed by Dual Tree-Complex Wavelet Transform (DT-CWT) and Support Vector Machine (SVM) is presented in this paper. Face images can be properly represented according to good directional selectivity on different scales of DT-CWT. Then spectral histograms are constructed by DT-CWT and other filters convolving with original images. Spectral histograms have intrinsic generalization of image representations. An efficient and robust classification function is obtained by training SVM using histograms as input vectors. Experimental results show that DT-CWT outperformed Gabor wavelet transform especially on computational redundancy and speed. And DT-CWT applies in face detection gives the satisfying detection accuracy.

Key words: Dual Tree-Complex Wavelet Transform (DT-CWT); spectral histogram; Support Vector Machine (SVM); face detection

摘要:提出了一种利用双树-复小波变换(DT-CWT)构成频谱直方图并通过支持向量机(SVM)进行人脸检测的方法。在人脸图像的表示中,基于DT-CWT在不同尺度上具有的很好方向选择性,对原始图像滤波,并与其它滤波器卷积滤波后得到原始图像的不同频率特征一起构成频谱直方图,该直方图在图像的表示上具有很好的本质扩展性。通过支持向量机(SVM)对频谱直放图向量进行分类训练,得到了有效区分人脸与非人脸的分类函数。实验显示,DT-CWT具有与Gabor变换类似的性质,而且计算冗余度更小,计算速度更快。应用DT-CWT频谱直方图的人脸检测算法取得了令人满意的结果。

关键词:双树-复小波变换(DT-CWT);频谱直方图;支持向量机(SVM);人脸检测

文章编号:1002-8331(2007)20-0086-04 **文献标识码:**A **中图分类号:**TP391.4

1 引言

人脸检测作为人脸自动识别系统中重要的一环,在机器视觉研究中具有重要意义,在身份鉴别、安全监控、视频检索和人机交互等方面都有广泛的应用。由于光照条件的改变、人脸姿势位置的不同、面部表情变化、遮挡以及成像距离的远近这些因素的影响,使人脸检测技术成为一个复杂的极具有挑战性的任务。所以寻找一个能将人脸模式与非人脸模式有效区分开的函数是人脸检测的研究核心。近年来,随着图像处理、人工智能、机器视觉及模式识别等学科的研究领域不断拓宽、深入,人脸检测技术得到了很大的发展,新理论和新方法陆续出现^[1]。因为人脸图像的复杂性,显式地表征人脸特征具有一定困难,所以基于学习的统计模型的方法(也称基于数据驱动的方法)越来越受到重视。它的主要特点是从训练数据中学习建立起人脸在图像中所有可能位置、尺度以及方位的模型,利用训练好的模型对测试图像进行检测分类。然而,在实际

应用中,训练集的大小始终有限,算法的普遍适用性就成为了基于统计学习方法的研究关键,它主要是由算法中使用的特征表示方法与分类器所决定的。

本文提出了一种基于统计学习的人脸检测算法,利用频谱直方图^[2,3]作为图像的特征表示,使用支持向量机(SVM)作为分类器。最近,Xiuwen Liu等人提出了使用频谱直方图进行对象分类和识别的方法,该方法对于纹理分析行之有效,对图像的代表拥有良好的本质扩展性^[4]。本文中使用的脸检测算法,在使用频谱直方图对图像的表示上进行了改进,利用双树-复小波变换(DT-CWT)取代Gabor小波变换。频谱直方图能够通过统计采样算法有效地将类似的图像分组,能够很好体现图像的本质特征^[4]。基于结构风险最小化的统计学习理论的SVM通过对训练集中提取得到的图像特征进行分类训练,得到在各种情况下能够有效区分图像中人脸与非人脸模式的决策函数。

文章接下来的部分组织如下:在第2章对Gabor变换滤波

基金项目:广东省自然科学基金(the Natural Science Foundation of Guangdong Province of China under Grant No.05006593)。

作者简介:孙粤辉(1979-),男,博士研究生,研究方向:机器视觉/统计模式识别;杜明辉,教授,博士生导师,主要研究方向:生物医学信号处理、图像处理等。

器与 DT-CWT 作主要的回顾,并通过实验对比,给出 DT-CWT 替代 Gabor 变换成为图像表示的原因;第 3 章描述了使用 DT-CWT 所得到的图像的频谱直方图表示;包括了训练与检测阶段的人脸检测算法在第 4 章中进行介绍;第 5 章给出了检测结果与实验数据,并对结果与数据进行分析说明,并对算法中存在的问题进行了讨论,指出下一步的工作方向;文章最后对算法进行总结。

2 Gabor 小波变换与 DT-CWT

人脸检测的算法中^[9],Gabor 小波变换在图像的特征表示中通常作为关键的滤波器。Gabor 小波能够从图像中提取方向性的频率特征。具体方式是将图像与一系列在方向与尺度上参数化的 Gabor 小波核进行卷积,得到一组结果作为图像的特征表,卷积滤波后得到的图像数与使用的 Gabor 小波的核数量一致。但是,重复的二维卷积计算使得运算速度非常慢。

DT-CWT 是最近提出的一种小波变换,在不同的尺度上拥有固定的 6 个不同方向,可以有效区分正负频率。而且分解后的图像维数根据尺度以 4 的倍数下降,因此冗余度比 Gabor 变换更少,计算速度更快。

2.1 Gabor 小波变换

Gabor 小波是在空间域中由高斯函数调制的复指数函数。二维 Gabor 函数的核能有效地描述哺乳动物视皮层简单细胞感受域剖面,实现对视皮层感受域数学表达^[6]。因此 Gabor 小波函数能表示出空间上良好的局部特征和方向性。设 $\sigma_x=\sigma_y=\sigma$, Gabor 小波滤波器可以定义如下^[5]:

$$\psi_{\mu,v}(z) = \frac{\|k_{\mu,v}\|^2}{2\pi\sigma^2} e^{-\frac{\|k_{\mu,v}\|^2 \|z\|^2}{2\sigma^2}} \left[e^{izk_{\mu,v}} - e^{-\frac{\sigma^2}{2\|k_{\mu,v}\|^2}} \right] \quad (1)$$

式(1)中, μ 与 v 分别定义为 Gabor 核的方向与尺度, $z=(x,y)$ 为给定的像素, $\|\cdot\|$ 是范数算子,向量 $k_{\mu,v}$ 由一个高斯包络函数约束,定义如下:

$$k_{\mu,v} = k_v e^{i\phi_\mu} = \begin{pmatrix} k_v \cos\phi_\mu \\ k_v \sin\phi_\mu \end{pmatrix}, k_v = 2^{\frac{v+1}{2}} \pi \quad (2)$$

$$\phi_\mu = \mu \frac{\pi}{n}, \mu = \{1, \dots, n\}$$

式(2)中, k_v 表示频率 v , ϕ_μ 表示方向 μ , n 表示的是所使用的方向总数。那么,图像 I 的 Gabor 变换 G_{out} 为:

$$G_{out} = I * \psi_{\mu,v} \quad (3)$$

2.2 双树-复小波变换(DT-CWT)

Kingsbury 博士提出的复小波^[7,8]在形状上与 Gabor 小波类似。如图 1 所示,DT-CWT 由两组实数滤波器树构成,树 a 与树 b ,产生复数系数的实部与虚部。为了补偿信号的延迟,设计了奇偶长度的双正交线性相位滤波器。

Kingsbury 总结了 DT-CWT 具有如下的特性^[7]:

- (1)近似的移不变性;
- (2)在二维上有着与 Gabor 滤波器类似的很好的方向选择性(在更高维上同样适用);
- (3)可以通过很短的线性相位滤波器完美重构信号;
- (4)冗余度有限,与尺度的数目无关,对于一维信号而言,与分解后信号长度比值为 2(对于 m 维而言,比值为 $2m$);

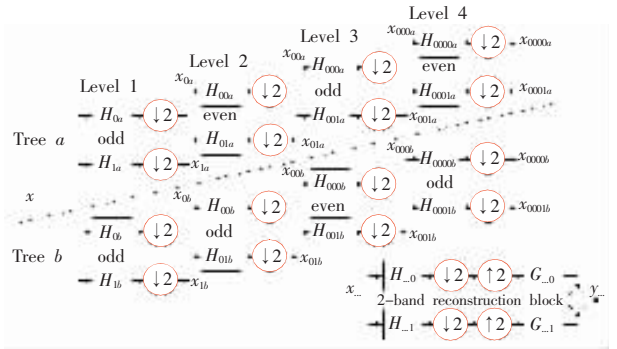


图 1 DT-CWT

此图由 N.Kingsbury 博士提供^[7]

(5)计算效率高,对一维信号而言,计算量为简单 DWT 的两倍(m 维而言,计算量为 $2m$ 倍)。

DT-CWT 滤波器能够使每一级的信号都完美地被重建。该变换可以区分正负频率并且产生 6 个方向为 $\pm 15^\circ, \pm 45^\circ, \pm 75^\circ$ 的子带,如图 2 所示,可以看出与 Gabor 是具有同样方向性。



图 2 DT-CWT 的滤波脉冲频率响应

前两行分别为响应的实部和虚部,第三行为响应的幅度

但是这些方向是固定的,而 Gabor 变换中子带信号可以在任意方向上进行计算。DT-CWT 对图像 $f(x)$ 的展开如下^[7,8]:

$$f(x) = \sum_k W_\phi(j_o, k) \phi_{j_o, k}(x) + \sum_i \sum_{j>j_i} \sum_k W_\psi(j, k) \psi_{j, k}(x) \quad (4)$$

式(4)中 $i=\pm 15^\circ, \pm 45^\circ, \pm 75^\circ$, 尺度函数 $\phi_{j_o, k}$ 与小波函数 $\psi_{j, k}$ 是复数, $W_\phi(j_o, k)$ 表示尺度化的系数而 $W_\psi(j, k)$ 则是变换得到的小波系数。因此,对应于 $i=\pm 15^\circ, \pm 45^\circ, \pm 75^\circ$ 表示的方向,得到了 6 个子带上的信号。图像的 DT-CWT 是一种接近严谨的结构,上下两层之间的信号采样之比为 4。而实验中 Gabor 滤波器使用 4 个频率尺度和 6 个与 DT-CWT 同样的方向参数,将得到 24 个子带信号,所以子带上的信号与原始信号的采样数是一致的,另外,重复进行的二维卷积运算在计算上十分耗时,系统的速度将被拖慢。

人脸图像分别经过 Gabor 滤波器与 DT-CWT 的幅度响应如图 3 所示。图 3(a)中人脸图像大小为 64×64 ,经过 4 个不同频率尺度,每个尺度 6 个方向的 Gabor 函数变换滤波后得到了 24 个 64×64 的子带矩阵,这些矩阵被成为一个 Gabor 集。使用同样的尺度与方向参数,图像经过 DT-CWT 处理后,尽管同样产生 24 个矩阵,但是第一个尺度频率下的 6 个矩阵的维数只有原始图像的 1/2,接下来的尺度频率中的矩阵又是上一个尺度中矩阵维数的 1/2,以此类推。因此,对应于一个大小为 64×64 的原始图像的 DT-CWT 集具有 6 个不同方向的子带,每个子带中的 4 个小波系数矩阵大小分别为 $32 \times 32, 16 \times 16, 8 \times 8$ 及 4×4 ,维数下降了很多。如图 3 所示,DT-CWT 在面部特征的提取方面有更好的性能,较 Gabor 而言,冗余度更小,处理速度更快。

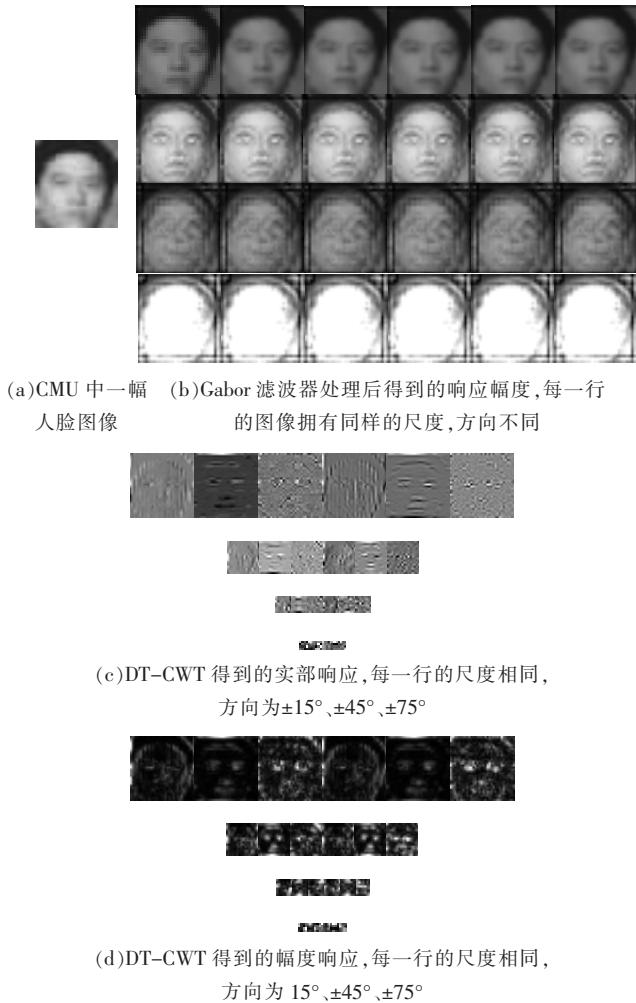


图3 经过 Gabor 滤波器与 DT-CWT 幅度响应的人脸图像

3 图像的频谱直方图表示

由文献[2]可知,给定输入图像 I 和一组滤波器 $\{F^{(\alpha)}, \alpha=1, \dots, K\}$,原始图像与滤波器卷积得到的子带图像 $I^{(\alpha)}$ 是通过线性卷积计算得到的:

$$I^{(\alpha)}(v) = F^{(\alpha)} * I = \sum_u F(u)I(v-u) \quad (5)$$

每个子带图像 $I^{(\alpha)}$ 的频谱直方图为:

$$H_{I^{(\alpha)}}(z) = \frac{1}{|I| \cdot \sum_v \delta(z - I^{(\alpha)}(v))} \quad (6)$$

将原始图像 I 与所有给定的滤波器的卷积后计算得到的 $H_{I^{(\alpha)}}$ 并置放在一起就得到了原始图像的频谱直方图:

$$H_I = (H_{I^{(1)}}, \dots, H_{I^{(K)}}) \quad (7)$$

由此,每幅图像都可以由滤波后得到的直方图表示。

人脸图像的频谱直方图是相近的,而对于不同类别的图像,它们的直方图也不相同。此外,通过利用随机采样算法证明^[9],将图像映射到频谱直方图空间的方法确实可以体现图像的特征本质。这种图像的表达是非参数化的,而且对于区分不同的模式十分有效。不仅如此,直方图对图像局部特征的细微变动甚至类似旋转与镜像的变化并不敏感,这样可以大幅度减少人脸与非人脸训练集中图像所需要的数量。

为了得到图像有效的频谱直方图表示,需要确定一组滤波器。参考文献[2,3]以及前文介绍的 DT-CWT,本文使用了 3 种不同类型的滤波器:

(1)4 个梯度滤波器, $D_x=[0, -1, 0], D_y=[0, -1, 0], D_{xx}=[-1, 2, -1]$ 与 $D_{yy}=[-1, 2, -1]$ 。

(2)4 个 LoG 滤波器, $\sigma=1/4, 1/2, 2, 32/18$ 。

(3)24 个 DT-CWT 滤波器,分为 4 个频率尺度,每个尺度上的方向为 $\pm 15^\circ, \pm 45^\circ, \pm 75^\circ$ 。

人脸与非人脸图像样本的频谱直方图如图 4 所示。

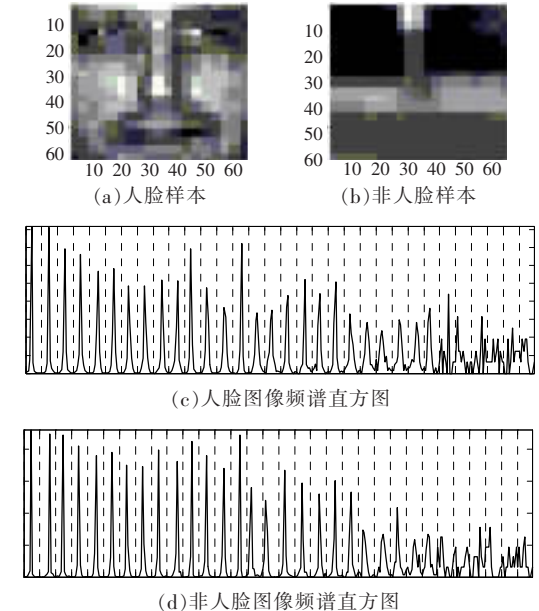


图4 人脸与非人脸图像样本及它们的频谱直方图表示

4 人脸检测算法

确定了图像的统计特征之后,为了将其用于接下来的分类、分割以及识别模块,必须定义一种距离度量函数。由于直方图属于离散密度函数,所以本文中采用的一种更适合比较两个频谱直方图的距离度量是 x^2 函数,该函数被广泛用于直方图之间的比较^[10],定义如下:

$$d_x^2(\mathbf{x}, \mathbf{y}) = \sum_i \frac{(x_i - y_i)^2}{x_i + y_i} \quad (8)$$

式(8)中, \mathbf{x} 与 \mathbf{y} 都是通过式(7)从不同的图像中计算得到的频谱直方图。训练集中的人脸与非人脸图像利用第 3 章介绍的滤波器计算得出频谱直方图后,将得到的直方图向量作为输入来训练以式(8)为核函数的 SVM。训练好的 SVM 用于对测试图像的图像窗口进行分类^[11]。

由于分类函数是针对 19×19 图像窗口设计的,所以需要所有的人脸都大致在窗口的内部。文中的人脸检测算法通过以 1.1 的倍数对测试图像连续向下采样的方式构建了 7 层的高斯金字塔。与此类似的分层方法曾被 Rowley 等人采用^[12]。每一层的金字塔在检测阶段都是单独处理的。对于每层的图像,一个 19×19 的图像窗口以 2 个像素的步长移动,遍历整幅图像。通过选定的滤波器计算出来的频谱直方图向量存储起来。将所有存储的直方图构成向量矩阵送入训练好的 SVM 进行分类。当一个图像窗口被分类为 人脸,窗口的坐标被记录下来,因此经

过 SVM 分类可以得到一组窗口的坐标。在后处理阶段,应用一些判断准则将这些小的区域连接起来,最后在输出图像上的人脸区域附近画一个矩形框表示检测得到的人脸位置。

5 初步的实验结果与讨论

实验中算法的训练阶段,2 549 幅正面的人脸图像采集自 ORL、CMU、Yale B 人脸数据库,并在平面内作了 $[-15^\circ, +15^\circ]$ 旋转处理,以增加算法对平面内旋转的人脸检测的鲁棒性。非人脸图像大约 4 546 幅,通过 bootstrap 自举算法从非人脸的自然图像中采集。将训练集中的所有图像以及测试集中的图像通过直方图均衡算法处理,将光照不均对检测算法的影响减至最低。实验里将本文提出的人脸检测算法应用于从训练图像中分离出来的 CMU+MIT data set。

在基于 CMU+MIT 测试集的实验中,实验结果如表 1 所示。图 5 给出了本文的人脸检测算子在测试图像中的检测结果。该算法在测试中展现出了良好的性能。

表 1 本文算法与其他经典算法结果比较

	检测率/%	错检数
本文算法	93.1	87
Rowley, Baluja, Kanade ^[1]	90.7	116
Schneiderman ^[1]	92.0	100

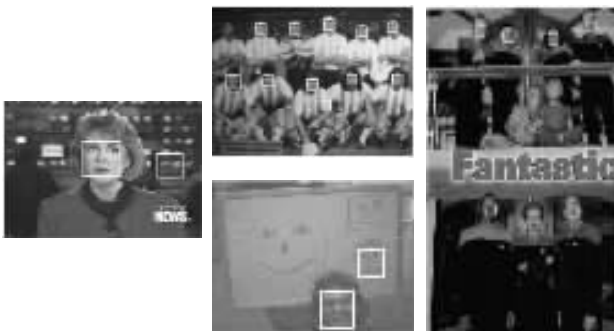


图 5 人脸检测算法对 CMU test set 中的测试图像的检测结果

与目前存在的一些人脸检测算法比较(例如文献[13, 14]),本文提出的算法在运算速度上相对还是慢很多。对一幅测试图像进行检测,通常需要几分钟的时间。如何应用类似 Adaboost 的算法训练学习得出更有效的分类器是算法改进的主要方向。另外提高算法速度就是将图像窗口以更多像素的步长移动而不是以目前的 2 个像素。但是,这样算法的性能将随着检测区域数量的下降而受到影响。改进的方式是使窗口自适应的进行移动而不是以固定的步长。

本文提出的将 DT-CWT 应用于人脸检测只是对 DT-CWT 在图像的识别与分类研究中的初步尝试。本文主要是用 DT-CWT 取代 Gabor 滤波器来计算频谱直方图,DT-CWT 在基于统计方法的人脸检测中是否能更多地应用还需要进一步研究,而且如何将 DT-CWT 的特点更好的体现在模式分类算法中也是接下来的研究方向。对频谱直方图空间中的特征有效地提取也可以更好地改进算法的性能以及系统运算的速度。

6 总结

本文提出了一种新颖的基于频谱直方图表示的人脸检测算法。首先,使用包括了 LoG 以及 DT-CWT 的一组滤波器捕捉人脸图像不同尺度与方向上的空间和频域特性;然后,原始图像与这些滤波器的卷积响应构成多维的直方图空间;最后,将直方向量矩阵输入到训练好的 SVM 中进行分类。初步的实验结果显示了频谱直方图对于人脸检测而言是一个很好的图像特征表示。而且,这种人脸检测方法具有一定程度上的旋转不变性,算法鲁棒性很好。(收稿日期:2006 年 11 月)

参考文献:

- [1] Yang M H, Kriegman D J, Ahuja N. Detecting faces in images: a survey[J]. IEEE Trans PAMI, 2002, 24(1): 34-58.
- [2] Liu Xiu-wen, Wang De-liang. A spectral histogram model for texton modeling and texture discrimination[J]. Vision Research, 2002, 42(23): 2617-2634.
- [3] Waring C A, Liu Xiu-wen. Face detection using spectral histograms and SVMs[J]. IEEE Trans Systems, Man, and Cybernetics—Part B Cybernetics, 2005, 35(3): 467-476.
- [4] Vapnik V. The nature of statistical learning theory[M]. New York: Springer-Verlag, 1995.
- [5] Liu Cheng-jun, Wechsler H. Gabor feature based classification using the enhanced fisher linear discriminant model for face recognition[J]. IEEE Trans Image Processing, 2002, 11(4): 467-476.
- [6] 桑农, 唐奇伶. 视皮层感知机制在轮廓提取中的应用[C]//第十二届全国图象图形学学术会议论文集. 北京: 清华大学出版社, 2005: 226-230.
- [7] Kingsbury N G. The dual-tree complex wavelet transform: a new efficient tool for image restoration and enhancement[C]//Proc European Signal Processing Conf, 1998: 319-322.
- [8] Kingsbury N G. Shift invariant properties of the Dual-Tree Complex Wavelet Transform[C]//Proc IEEE Conf on Acoustics, Speech and Signal Processing, Phoenix, AZ, 1999, 3: 1221-1224.
- [9] Liu Xiu-wen, Srivastava A, Wang De-liang. Intrinsic generalization analysis of low dimensional representations[J]. Neural Networks, 2003, 16(5/6): 537-545.
- [10] Schiele B, Crowley J L. Object recognition using multidimensional receptive field histograms[C]//Fourth European Conference on Computer Vision, ECCV '96, 1996, 1: 610-619.
- [11] Osuna E, Freund R, Girosi F. Training support vector machines: an application to face detection[C]//Conf on Computer Vision and Pattern Recognition, San Juan, Puerto Rico, 1997: 130-136.
- [12] Rowley H A, Baluja S, Kanade T. Neural network-based face detection[J]. IEEE Trans PAMI, 1998, 20(1): 23-38.
- [13] Lienhart R, Kuranov A, Pisarevsky V. Empirical analysis of detection cascades of boosted classifiers for rapid object detection[C]//Proc DAGM Pattern Recognition Symposium, Madgeburg, Germany, 2003: 297-304.
- [14] Viola P, Jones M J. Robust real-time face detection[J]. International Journal of Computer Vision, 2004, 57(2): 137-154.