基于傅里叶频谱分析和稀疏Logistic回归的反照片欺骗算法

李 翼　谭晓阳

南京航空航天大学计算机科学与工程系，南京210016
E-mail: lfly2008@nuaa.edu.cn, x.tan@nuaa.edu.cn

摘要： 活体检测是生物特征识别系统中检测和拒绝仿冒身份特征的一项重要功能，而照片欺骗是人脸识别中一种最常见的入侵方式，对基于人脸识别的安全系统构成很大的威胁。在人脸照片和真实人脸图像的差异主要反映在中低频谱的假定下，本文提出一种新的基于二维离散傅里叶分析和稀疏logistic回归的反照片欺骗算法。该方法首先用DoG（Difference of Gaussian）滤波技术来提取图像数据的主要信息，然后基于稀疏logistic回归技术来对输入图像进行快速和有效的判别。为了处理训练数据中的类不平衡问题，采用early stopping技术来提高logistic回归模型的鲁棒性。本文也对所提方法中的关键参数（如稀疏度）对性能的影响进行了研究。在大型测试集上的实验结果表明了所提方法的有效性。

关键词： 活体检测, 反照片欺骗,人脸识别, DoG 滤波器,二维离散傅里叶变换,稀疏Logistic回归

An Anti-Photo Spoof Method in Face Recognition Based on the Analysis of Fourier Spectra with Sparse Logistic Regression

Yi Li　　Xiaoyang Tan

Dept. of Computer Science and Technology,Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China
E-mail: lfly2008@nuaa.edu.cn, x.tan@nuaa.edu.cn

Abstract: Among others, spoofing with photos is one of the most common manner to intrude a face recognition system. In this paper, we presents a novel method to deal with this problem, based on the observation that the difference between a photo and a real face usually leads to different distribution behavior in the frequency domain. In particular, we propose to first use a DoG (Difference of Gaussian) filter on the given image to preserve rich information for the subsequent stages while suppressing less-discriminative energy as much as possible. A two-dimensional discrete Fourier transformation is then applied on the filtered image, which produces an input to a sparse logistic regression model to give the final determinant. Furthermore, we adopt an early stoping strategy to prevent the logistic model from overfiting when our training set has class imbalance problem. We also investigate the influence of degree of sparsity on the performance of the system. Extensive experiments on a large scale testing set verify the feasibility and effectiveness of the proposed method.

Key Words: Live detection, anti photo spoof, face recognition, DoG filter, two-dimensional discrete Fourier transform, Sparse Logistic Regression

1　引言

在各种基于生物特征的应用中，人脸识别 [1]由于其符合人类自身区分不同人的方式，且用户友好（即不需要刻意或令使用者反感的配合等），受到了广泛的欢迎。生物特征识别系统的使用安全性是人们普遍关注的问题，人们对生物特征识别系统的信心和接受很大程度上取决于系统的鲁棒性、低错误率和抗欺骗能力 [2]。

生物特征识别系统中，最常见的欺骗方式是发生在用户界面上 [2]。冒名顶替者使用某种具有相同表现形式的假冒特征来入侵系统。在人脸识别应用中，合法用户的人脸图片、视频以及三维模型等常被用于身份伪造，而其中照片是最常见的欺骗方式。这一问题通常被称为活体检测问题（liveness detection或者liveness testing），其目的是判断获取到的生物特征是否来自一个已授权的、有生命的、在现场的、真实的人 [3]。

研究者们通过分析真实人脸与照片人脸的图片信息差异，提出了大量反照片欺骗的活体检测方法，如根据头部的三维深度信息去判断是否是一个真人 [4]; 对采集到的二维图像进行傅里叶变换分析 [5]; 分析真实人脸的非刚性运动变化，如采用线性光学流的方法捕捉三

维人脸细微的运动信息 [6]，发声时嘴唇部位的动作分析 [7]，眨眼动作分析 [8]，眼睛区域的变化 [9]，头部运动变化 [10]等; 采用多模式判别的方法，如分析人脸的红外图像 [11]，联合声音识别的方法 [12]等。

我们认为，一种理想的活体检测方法至少应该满足如下一个或多个条件：1、不需要用户的主动配合；2、不需要额外的设备；3、计算快速、高效，计算的时空代价在可接受范围内；4、活体检测所采用的特征易于捕获；5、有较强的鲁棒性，能够克服光照变化和抗部分遮挡干扰；6、可作为独立的模块，在不需要改变其他模块的前提下，加入已有的人脸识别系统。

本文提出了一种可用于基于人脸识别的身份认证应用的反照片欺骗活体检测方法。该方法通过首先对图像进行DoG（Difference of Gaussian）滤波的预处理，从预处理过的二维图像中提取出傅里叶变换特征，再采用稀疏logistic回归模型来判断身份认证中采集到的图像是真实人脸还是照片人脸。实验结果表明，本文提出的方法在不添加额外的辅助设备、不需要用户的主动配合、实现简单、计算量小且功能独立的情况下，能够很好的解决基于人脸识别的身份认证中照片欺骗的问题。

本文以下第2节首先对人脸图像的成像特点与傅里叶变换的关系进行介绍，第3节详细描述所提方法，第4节给出实验结果，最后是结论。

2 人脸的成像特点与傅里叶变换

在典型的基于人脸识别的身份认证中，人脸图像的成像结果往往受到三方面的影响 [18]：1、人脸的内部特征（如人脸皮肤的反射属性、3D结构属性，人脸表情等）；2、外部的成像条件（如所处的环境内的光照情况，人脸部存在的饰物、毛发的遮挡等）；3、图像采集设备的属性（如照相机的焦距、光圈、分辨率等因素）。如果将采集人脸图像所处环境的光源想象成是理想的点光源，则按照Lambertian模型，人脸的生成图像表示成:

$$I(x,y) = f_c(\rho_{(x,y,z)} h^T_{(x,y,z)} s_{(x,y,z)}) \qquad (1)$$

其中，$\rho_{(x,y,z)}$是三维空间中人脸表面一点的坐标，$h^T_{(x,y,z)}$是人脸表面一点所在平面的法向量，$s_{(x,y,z)}$是点光源在人脸表面一点的方向和强度，fc是摄像头的成像函数，$I(x,y)$是三维空间中一点(x,y,z)对应的映射到成像平面的图像点。注意到如果人脸识别中采集到的人脸图像是由照片成像的，则成像前的物体是一个平面的物体，照片中的每一点所处平面的法向量可近似为一个固定的常数。即使是照片存在一定程度的向内、向外弯曲和水平方向、垂直方向内的旋转，其各点所在平面的法向量$h^T_{(x,y,z)}$也不符合真实人脸面部的凹凸情况。因此从理论上说，我们可以利用照片成像的这个特点来区分照片和真人图像。

但在实践中，由于人脸表面各点所在平面的法向量不易直接计算，但该量实际反映的是对象表面纹理分布，不同的对象表面纹理对光照反射程度的不同，从而造成不同的成像结果。我们可以利用这一特点来区分照片和真实人脸：真人人脸与照片人脸由于纹理不同而往往在光照反射程度会有很大的不同，进而造成成像差异。在实践中我们可以方便地对采集得到的二维图像进行傅里叶变换来捕捉这种差异。图1给出了真实人脸、照片人脸以及在频域内的对比示例。

基于上述分析，最近Jiangwei Li等人提出一种对单张人脸图片或者人脸图片序列进行傅里叶变换分析来判断是否是真实人脸的方法 [5]。该方法基于这样的假设：照片会小于真实人脸的大小并且照片是平的，因此由照片得到的人脸图像比真实人脸得到的人脸图像含有更少的高频部分信息。针对这个特点，他们对图像在一定的频率范围内的高频信息所占的比重采用阈值的方法进行判断，即大于高频阈值的人脸图像为真人人脸图像，否则为照片人脸图像。

这种方法有不少的优点，如不需要用户的主动配合、不需要添加辅助设备、计算简单等。但是在现实应用场景中，即使照片也可能存在大量高频信息（脸部的痣、眼镜、胡须等遮挡以及照片反光等光照情况而带来的高频噪声等），极大降低了高频信息的区分能力。此外，基于阈值的分类器过于简单而难以有效利用训练样本中包含的有用信息。针对上述两个问题，我们基于一个新的（往往在现实世界中更合理的）假定，从而提出一种区分照片和真人的方法。

3 基于DoG滤波和稀疏Logistic回归的反照片欺骗算法

3.1 基本思想

我们所提出的方法建立在如下观察之上：1）照片（或视频）图像在摄像头中的图像实际上是一种二次成像；2）现代CCD感光器的主要噪声来源为与光强平方根成比例的短噪声(short noise)。从直观上说，照片二次成像相当于对图像造成更大的模糊效果–增强了低频信息分布，而损失大量的中频细节信息。换言之，真人图像由于是一次成像因而比照片包含更多的中频细节信息。所以，有理由认为中频带信息可能具有更

(a) 真实人脸　(b) 真实人脸的傅里叶频谱　(c) 照片人脸　(d) 照片人脸的傅里叶频谱

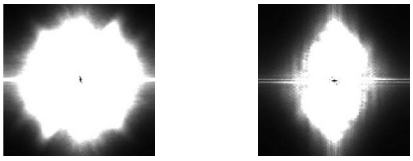
图 1 真实人脸、照片人脸以及在频率域内的对比

大的区分能力。为此，我们首先利用DoG（Difference of Gaussian）滤波器来抽取这部分信息。

本文将反照片欺骗问题视为一个两类分类问题。为了充分利用大量训练样本中的统计信息，我们在经过DoG滤波和频率域变换后的图像基础上构建一个两类分类器来进行判定。本文中选择logistic回归作为分类器。该分类器性能与线性SVM相当 [17]，但是无需保存支持向量模板而在检测效率上优于后者。

3.2 DoG滤波

DoG滤波器是一种获取结果带通的方法。好的空间域内的细节信息对于分类器做出判断非常重要。DoG中涉及到的两个主要参数是内外高斯滤波器的方差，分别记为σ_0和σ_1。一般内高斯取值较窄（σ_0小于等于1个像素），而外高斯σ_1取2-4个像素 [13]，取决于实际空间域内的频率信息分布。在本文中这两个参数的取值为$\sigma_0 = 0.5$和$\sigma_1 = 1$。其处理效果见图2。



(a)真实人脸的傅里叶频谱　(b)照片人脸的傅里叶频谱

图 2 经过DoG滤波后，真实人脸与照片人脸的傅里叶频谱对比

3.3 稀疏Logistic回归

在进行了DoG滤波之后，我们将输入图像进行二维离散傅里叶变换。为了更好的刻画真实人脸与照片人脸在频率域的统计差异，我们采用logistic回归机作为分类器。Logistic回归属于一类判别型模型，通过直接最大化类标号的后验概率来学习一个线性分类函数，与其他线性分类器如SVM不同的是，该方法无需保存任何训练样本，在检测时的效率大大提高。

本文中，我们基于训练数据构建了一个稀疏logistic模型 [14]，该模型在传统logistic回归的似然函数函数上增加了一个基于l_1范数的正则化项，以防止模型陷入过拟合。

4 实验

4.1 实验配置

对输入的图片，我们首先采用 [15]中的方法进行人脸检测，截出人脸区域。再采用 [16]中眼睛定位的方法在人脸区域内，进行眼睛定位。根据定位得到的眼睛坐标，将人脸图片规整化到64像素×64像素。经过规整化的灰度人脸区域进行DoG滤波之后，对这些图像采用二维离散傅里叶变换，再用logistic回归分类器判断是否是真实人脸。活体检测的整个流程如图3。

图 3 活体检测流程

4.2 数据收集

正、反例样本包括真人人脸图像序列和照片人脸图像序列。采样的基本原则是尽量减少真人人脸所含的动态信息，增加照片人脸所含有的动态信息。

正例样本使用网络摄像头收集了9个真人人脸的图像序列,涵盖了有无眼镜遮挡以及性别、年龄等变化因素。每个人都被要求正视网络摄像头，自如眨眼，表情自然，无明显的动作和表情变化。只为每个人收集一组正例样本。

反例样本是用这9个人的照片采集得到的图像序列。采集反例样本所使用的照片包括四寸照片（6.8cm×10.2cm）和五寸照片（8.9cm×12.7cm）两种大小以及打印机打印和传统冲洗两种质地的照片。在所使用的照片中，人脸部区域占整个照片区域的大小超过2/3。反例样本尽可能地涵盖了人脸识别中照片欺骗的各种方式：在网络摄像头前，照片保持正对、沿水平轴旋转、沿垂直轴旋转、上下方向的向内向外弯曲、左右方向的向内向外弯曲，如图4中五种情况。在每种情况下，照片还存在上下左右前后的空间位置移动。

正、反例样本均使用相同的网络摄像头在同一地点和同一时间段内采集，光照条件均为室内非控光源。正、反例样本的图片序列采集速度为20fps，每组图片采集500张。图片的分辨率为640像素×480像素。

图 4 反例样本采集方式(从左至右分别为：静止、上下旋转、左右旋转、上下弯曲、左右弯曲等五种情况)

经过人脸检测、眼睛定位以及手工筛选（删除规整化后人脸明显的大角度倾斜的情况）后，得到的正例样本为9组真人人脸图片序列，共计3548张真实人脸图片，反例样本为45组照片人脸图片序列，共计21910张照片人脸图片。部分正、反例样本示例如图5。

在实验中，我们选取了一部分正例样本和反例样本作为训练集。从9组真人图片中随机选取4组，共计1700张图片作为正例训练样本，从45组照片人脸图片中随机选取15组照片人脸图片，共计7243张图片作为反例训练样本。将剩余的、未参加过训练的样本全部作为测试集。

(a) 规整化后的真实人脸图片序列

(b) 规整化后的照片人脸图片序列

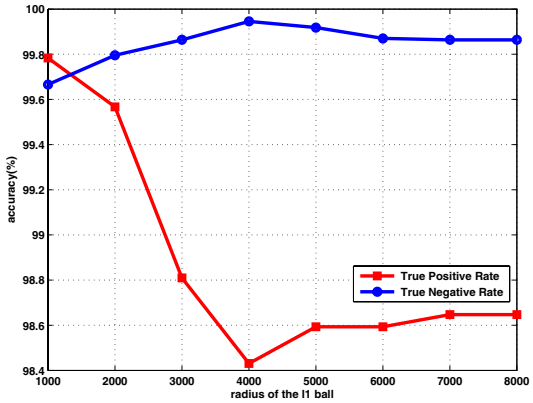
图 5 采集到的真实人脸、照片人脸图片序列示例

4.3 实验结果

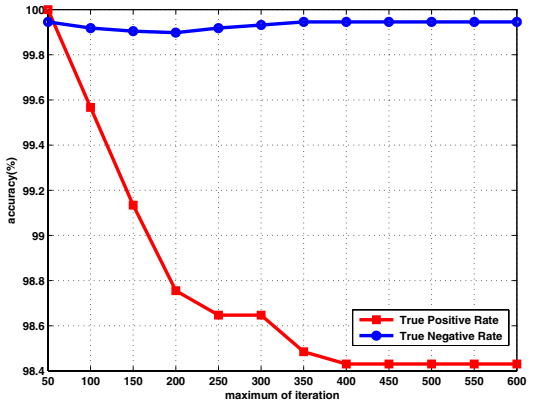
4.3.1 参数分析

我们首先研究了稀疏logistic回归中的稀疏度，即不同l_1范数正则化的球半径（radius of l_1 ball）对检测性能的影响，实验结果如图6(a)所示。可见，l_1范数正则化强度对于正反例测试样本的检测性能有很大的影响，考虑到应用于人脸识别中的反照片活体检测的目的是希望在保证真实人脸的正确检测率的前提下，尽可能的提高照片人脸的正确检测率，我们在后续实验中选择l_1范数正则化强度4000，此时图片人脸的拒绝率提高到足够的高，接近100%（误接受率低于千分之一），而真实人脸图片的通过率稳定在≥98%的情况。

由于本文训练集存在严重的类不平衡问题（正例样本1700,反例样本7243），如此训练的logistic回归器将对反例有极大的偏置。为了解决这个问题，我们在训练中采用了early stopping技术，即在logistic回归训练收敛之前结束训练。图6(b)给出了不同的最大迭代次数对分类器性能的影响。可见，当迭代50次时结束训练可以得到较好的泛化能力。而继续训练则将导致true positive rate（即真人图片正确检测率）的下降，表明发生了对照片图像的偏置。



(a) 不同l_1范数正则化的球半径对分类器性能的影响



(b) l_1范数正则化的球半径设为4000时，不同的最大迭代次数对分类器性能的影响

图 6 稀疏logistic回归训练参数对于实验结果的影响

4.3.2 结果分析

在训练集上训练得到的logistic模型用1848张真实人脸图片和14667张照片人脸图片进行测试，在同样的

表 1 不同的活体检测方法分类正确率对比

分类器	True Positive	False Positive	True Negative	False Negative
文献 [5]方法	96.27%	0.73%	99.27%	3.73%
未DoG滤波的稀疏logistic回归方法	86.58%	0.18%	99.82%	13.42%
DoG滤波的稀疏logistic回归方法	**100.0**%	**0.05**%	**99.95**%	**0.0**%

数据集上和 [5]中方法进行了对比。实验结果如表1所示。由结果可见采用DoG滤波及稀疏logistic回归方法显著提高了正确检测率。

5 结论

本文提出了一种可用于基于人脸识别的身份认证应用的反照片欺骗活体检测方法。该方法通过首先对图像进行DoG（Difference of Gaussian）滤波的预处理，进一步提取傅里叶变换特征，再采用稀疏logistic回归模型来判断身份认证中采集到的图像是真实人脸还是照片人脸。实验结果表明，本文提出的方法在不添加额外的辅助设备、不需要用户的主动配合、实现简单、计算量小且功能独立的情况下，能够很好的解决基于人脸识别的身份认证中照片欺骗的问题。

致谢

本研究部分受国家自然科学基金60773060，江苏省自然科学基金BK200922660，以及教育部留学人员基金的支持。

References

1 X.Tan, S.Chen, Z.-H. Zhou, and F. Zhang. Face Recognition from a Single Image per Person: A Survey. *Pattern Recognition*, 39(9): 1725-1745, 2006.

2 Anil K. Jain, Karthik Nandakumar and Abhishek Nagar. Biometric Template Security. *Advances in Signal Processing*, January, 2008.

3 Schuckers, Stephanie A C. Spoofing and Anti-Spoofing Measures. *Info. Sec. TR*, 4(7): 56-62, 2002.

4 T. Choudhury, B. Clarkson et al., Multimodal Person Recognition using Unconstrained Audio and Video. *AVBPA)*, 176-181, 1999.

5 J. Li, Y. Wang, T. Tan, A.K.Jain. Live Face Detection Based on the Analysis of Fourier Spectra. *Proceedings of SPIE*, 296-303, 2004.

6 K. Kollreider, H. Fronthaler and J. Bigun. Evaluating liveness by face images and the structure tensor. *AIAT*, 75-80, 2005.

7 Robert W.Frischholz, Ulrich Dieckmann. BioID: A Multimodal Biometric Identification System. *IEEE Computer*, 2(33): 64-68, 2000.

8 Gang Pan, Lin Sun, Zhaohui Wu, Shihong Lao. Eyeblink-based Anti-Spoofing in Face Recognition from a Generic Webcamera.*ICCV, 2007*, 1-8, 2007.

9 Hyung-Keun Jee, Sung-Uk Jung, and Jang-Hee Yoo. Liveness Detection for Embedded Face Recognition System. *IJMMS*, 235-238, 2006.

10 Robert W. Frischholz, Alexander Werner. Avoiding Replay-Attacks in a Face Recognition System using Head-Pose Estimation. *AMFG*, 234-235, 2003.

11 Diego A. Socolinsky, Andrea Selinger. A Comparative Analysis of Face Recognition Performance with Visible and Thermal Infrared Imagery. *CVIU*, 72-114, 2003.

12 Girija Chetty, Michael Wagner. Liveness Verification in Audio-Video Speaker Authentication. *AICSST*, 358-363, 2004.

13 X.Tan and B.Triggs. Enhanced Local Texture Feature Sets for Face Recognition under Difficult Lighting Conditions. *AMFG07*, 168-182, 2007.

14 Jun Liu, Jianhui Chen, Jieping Ye. Large-Scale Sparse Logistic Regression. *ICML*, 2009.

15 P. Viola and M. Jones. Robust real-time face detection. *IJCV*, 57(2): 137-154, 2004.

16 X. Tan, F. Song, Z. Zhou, S. Chen1. Enhanced Pictorial Structures for Precise Eye Localization Under Uncontrolled Conditions. *CVPR*, 57(2): 137-154, 2009.

17 Trevor Hastie, Robert Tibshirani, Jerome Friedman. *The Elements of Statistical Learning*, Springer, 2001.

18 山世光. 人脸识别理论与应用研究. 信息技术快报, 第3卷第10期, 2005.