

# 基于马尔可夫链模型的图像拼接篡改盲取证算法

卜江 郑彬 陈海洋

(国防科学技术大学人文与社会科学学院 长沙 410074)

**摘要** 利用马尔可夫链对图像 DCT 变换系数的相关性进行建模,提取模型的状态转移概率矩阵作为特征向量,然后采用支持向量机分类器对自然图像和篡改图像进行分类。实验结果表明,该算法的平均准确率达到 86%,性能优于 Ng 提出的基于双相干特征的拼接图像检测算法。

**关键词** 盲取证,离散余弦变换,马尔可夫链,支持向量机

**中图分类号** TP391 **文献标识码** A **DOI** 10.11896/j.issn.1002-137X.2015.12.051

## Forensic Method Based on Markov Chain Model for Splicing Images

BU Jiang ZHENG Bin CHEN Hai-yang

(College of Humanities and Social Sciences, National University of Defense Technology, Changsha 410074, China)

**Abstract** Markov chain is used to model the correlation of discrete cosine transform coefficients, and the state transition probability matrix is extracted as feature vector. Then the support vector machine classifier is used to discriminate natural images and doctored images. Experimental results show that the average accuracy rate can achieve 86%. The detection performance of our method is better than that of the method using bi-coherence features proposed by Ng.

**Keywords** Blind forensics, Discrete cosine transform, Markov chain, Support vector machine

## 1 引言

随着现代数字技术的发展以及低成本和高分辨率数码相机的普及,数字图像在日常生活和工作中得到了广泛的使用。数字化后的图像数据显然具有传统模拟时代无法比拟的许多优点,比如易于创作、易于存储、易于发布等,同时也提高了图像的欣赏质量。然而数字图像的便利性和不安全性是并存的,数字图像在带给人们方便的同时,也引入了一些潜在的风险,它们很容易被篡改。虽然大多数情况下人们修改文件都有合法的目的,但是有一些别有用心的人会通过篡改原始作品的内容,达到损害他人名誉、混淆视听等目的,甚至造成严重的后果。例如对 X 光医学图像的一个不经意的修改可能会造成误诊;作为法庭证据的照片如果被恶意修改就可能扭曲法律事件的真实面貌,使好人蒙冤而坏人却逍遥法外。因此迫切地需要发展能够对数字图像的真伪进行检测的鉴别技术。

传统的数字图像鉴别技术主要有两种:数字签名和数字水印。但是传统的鉴别技术都很容易造成图像鉴别盲区,因为它们都需要图像提供方进行摘要的提取和数据的嵌入,即鉴别需要图像提供方进行配合,从而导致两种方法的应用受到限制。针对传统数字签名和数字水印鉴别技术适用面窄引起的技术需求,近几年刚发展起来一种数字图像盲取证技术,它是一种新的图像鉴别思路,旨在不依赖任何预签名提取或预嵌入信息的前提下,对图像的真伪和来源进行鉴别。目前

利用盲取证技术来检测恶意篡改的图像取得了一定的成果,提出的算法包括复制-粘贴检测<sup>[1,2]</sup>、模糊检测<sup>[3]</sup>、相机参数不一致检测<sup>[4]</sup>、彩色插值检测<sup>[5]</sup>、拼接图像检测<sup>[6]</sup>等。在拼接图像检测研究中,文献[13]提出了一种基于人脸色温的篡改检测算法,但该算法只适用于人脸拼接篡改的图像;文献[14]利用颜色滤波矩阵来估计原图的拼接位置,但选取的特征较为单一,导致篡改图像检测的准确率较低。

本文将篡改图像检测看作是一个两类模式识别问题,即按图像是否被拼接篡改分为自然图像和篡改图像两类。篡改伪造图像的逼真程度再高,也会由于篡改操作而影响图像的内在统计特性,甚至破坏这种内在统计特性的一致性。因此本文从分析自然图像的统计特性出发,利用马尔可夫链模型对图像 DCT 交流系数之间的相关性进行建模,提取模型的状态转移概率矩阵作为特征向量,然后将该特征向量送入支持向量机(Support Vector Machine, SVM)分类器,实现对自然图像和篡改图像的分类和检测。实验结果表明,本文算法的检测准确率较高,性能优于文献[6]所提的算法。

本文第 2 节介绍图像篡改检测的特征提取;第 3 节简要介绍支持向量机分类器;第 4 节是实验结果与比较;最后总结全文。

## 2 特征提取

受马尔可夫模型用于图像隐写分析的启发<sup>[7]</sup>,本文利用马尔可夫链模型,提取 DCT 块内系数之间的相关性,提取相

到稿日期:2014-12-01 返修日期:2015-03-29 本文受总装备部预先研究项目,国防科技大学校级项目资助。

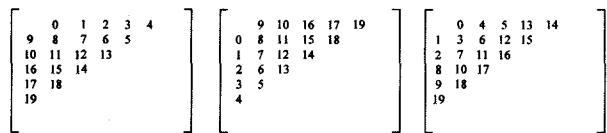
卜江(1983-),男,博士,讲师,主要研究方向为多媒体数据处理、数字图像盲取证等, E-mail: veron9@163.com; 郑彬(1985-),女,硕士,助教,主要研究方向为多媒体数据处理、数字图像盲取证等; 陈海洋(1987-),男,硕士,助教,主要研究方向为多媒体数据处理、数字图像盲取证等。

应的转移概率矩阵作为特征。

马尔可夫特征的提取步骤详细描述如下：

(1) 图像分块 DCT 变换。将彩色图像的亮度通道(或灰度图像)按  $8 \times 8$  做分块 DCT 变换, 得到图像的 DCT 系数矩阵, 步骤(2)~步骤(5)的操作都是针对块操作的。

(2) 以 3 种顺序扫描每个 DCT 块, 形成 3 个一维向量。对每个  $8 \times 8$  DCT 块的中低频 AC 系数按照图 1 所示的顺序展开成 3 个一维向量  $F_h, F_v$  和  $F_z$  (其中下标  $h, v, z$  分别代表水平扫描、垂直扫描和 ZigZag 扫描)。图 1 中以  $8 \times 8$  的 DCT 块为例, 仅显示 DCT 块的中低频 AC 系数, 其中的数字表示对其进行扫描的顺序。



(a)局部水平扫描 (b)局部垂直扫描 (c)ZigZag 扫描

图 1 扫描 DCT 块的 3 种顺序

(3) 计算块差分向量。对 3 个一维向量  $F_h, F_v$  和  $F_z$  取整, 并求绝对值, 然后用式(1)~式(3)分别求其一步差分矩阵, 得到 3 个一维差分向量  $D_h, D_v$  和  $D_z$  :

$$D_h(i) = |[F_h(i)]| - |[F_h(i+1)]| \quad (1)$$

$$D_v(i) = |[F_v(i)]| - |[F_v(i+1)]| \quad (2)$$

$$D_z(i) = |[F_z(i)]| - |[F_z(i+1)]| \quad (3)$$

其中,  $i \in [0, \dots, C-1]$ ,  $C$  为一个 DCT 块中所选取的中低频 AC 系数的总数, 取  $C=20$ ,  $[\cdot]$  为取整操作,  $|\cdot|$  为取绝对值操作。

(4) 阈值处理, 即将  $D_h, D_v$  和  $D_z$  中的系数值钳位于  $[-T, T]$  范围内, 将小于  $-T$  的系数全部改为  $-T$ , 大于  $T$  的系数全部改为  $T$ 。阈值处理的目的是降低计算马尔可夫特征的维数, 对于在  $[-T, T]$  范围的系数值, 计算得到的马尔可夫状态转移概率矩阵的维数为  $(2T+1) \times (2T+1)$ 。

(5) 计算块马尔可夫转移概率矩阵。利用马尔可夫链模型, 根据式(4)~式(6)来分别计算每个 DCT 块差分向量  $D_h, D_v$  和  $D_z$  的转移概率矩阵:

$$M_h \{D_h(i+1)=n | D_h(i)=m\} = \frac{\sum_{i=0}^{C-1} \delta(D_h(i)=m, D_h(i+1)=n)}{\sum_{i=0}^{C-1} \delta(D_h(i)=m)} \quad (4)$$

$$M_v \{D_v(i+1)=n | D_v(i)=m\} = \frac{\sum_{i=0}^{C-1} \delta(D_v(i)=m, D_v(i+1)=n)}{\sum_{i=0}^{C-1} \delta(D_v(i)=m)} \quad (5)$$

$$M_z \{D_z(i+1)=n | D_z(i)=m\} = \frac{\sum_{i=0}^{C-1} \delta(D_z(i)=m, D_z(i+1)=n)}{\sum_{i=0}^{C-1} \delta(D_z(i)=m)} \quad (6)$$

其中,  $m, n \in \{-T, -T+1, \dots, 0, \dots, T\}$

$$\delta(A=m, B=n) = \begin{cases} 1, & \text{if } A=m \text{ and } B=n \\ 0, & \text{otherwise} \end{cases} \quad (7)$$

(6) 计算水平转移概率矩阵、垂直转移概率矩阵和 Zig-

Zag 转移概率矩阵。对每种扫描顺序的所有 DCT 块的转移概率矩阵求平均, 分别得到水平转移概率矩阵  $\bar{M}_h$ 、垂直转移概率矩阵  $\bar{M}_v$  和 ZigZag 转移概率矩阵  $\bar{M}_z$ 。

(7) 计算全局转移概率矩阵。对水平转移概率矩阵、垂直转移概率矩阵和 ZigZag 转移概率矩阵求平均, 得到全局马尔可夫转移概率矩阵  $M = (\bar{M}_h + \bar{M}_v + \bar{M}_z) / 3$ , 最终形成  $(2T+1) \times (2T+1)$  维特征。

需要指出的是: (1) 图 1(c) 是 ZigZag 顺序, 它描述的是 DCT 域频率由小到大的顺序, 对于原始图像, 按照此顺序展开的系数相邻之间具有较高的连续性和相关性。而图像拼接操作会在一定程度上破坏这种连续性和相关性。另外我们还采用了图 1(a) 和图 1(b) 两种扫描顺序, 目的是尽可能多地提取 DCT 频率系数的空间相关性。(2) 块 DCT 变换系数的 AC 系数基本集中在 0 附近, 为了进一步缩小系数的范围, 进行块差分向量的计算, 所有块差分向量的系数都很小, 只要选择适当的阈值  $T$ , 就能既保证丢失的信息极少, 又极大地降低计算的复杂性。

### 3 支持向量机分类器

在提取对拼接篡改敏感的特征向量后, 一般学习型的通用图像盲取证算法需要设计一个有效的分类器, 以实现真实图像和拼接图像的有效区分。本文算法的分类器基于支持向量机(SVM)。支持向量机是建立在统计学习理论和结构风险最小原理基础上的, 其基本思想是对于非线性可分样本, 将其输入向量经非线性变换映射到一个高维空间  $Z$  中, 在变换后的空间中寻找最优的分界面(超平面), 使其推广能力最好。SVM 在解决小样本、非线性及高维数的模式识别问题方面表现出许多特有的优势, 已经在模式识别、函数逼近和概率密度估计等方面取得了良好效果<sup>[8]</sup>。

SVM 常用的核函数主要有: 线性、多项式、径向基和 S 形核函数。本文实验中采用径向基核函数进行分类。

### 4 实验结果与比较

#### 4.1 图像篡改检测数据库

实验采用 Columbia 大学提供的篡改图像检测评估数据库(详见文献[9]), 它是由 Columbia 大学 DVMM 实验室建立的、用于篡改图像盲取证算法的标准测试库。该数据库中包括了 1845 幅内容多样的图像, 大小为  $128 \times 128$  像素, 图像格式为 BMP, 这些图像主要来自于 CalPhotos 图像库<sup>[10]</sup>, 另外的一小部分的图像由 DVMM 实验室自己拍摄。这些图像根据内容可分为以下 5 个子类: 均匀纹理、均匀平滑、纹理-纹理、平滑-平滑和纹理-平滑。

#### 4.2 实验参数设置

在提取马尔可夫特征时, 对图像进行  $8 \times 8$  的分块 DCT 变换, 阈值  $T=3$ , 提取全局转移概率矩阵  $7 \times 7=49$  维, 即用于分类的特征向量是 49 维。

分类器采用目前广泛使用的 LibSVM<sup>[11]</sup>。LibSVM 是台湾大学林智仁(Lin Chih-Jen)副教授等设计开发的一个简单、易于使用和快速有效的 SVM 开源软件包, 可有效用于分类

问题、回归问题以及分布估计等问题。本文实验采用径向基函数作为核函数,实验中需要确定的两个重要参数是惩罚因子  $C$  和核函数的参数  $\gamma$ 。 $C$  的默认值为 1,  $\gamma$  的默认值为  $\frac{1}{k}$ , 其中  $k$  为样本的特征个数。通过 5 次交叉验证的方法来选取参数  $C$  和  $\gamma$  的最优值。每次实验中,随机选择数据库中 50% 的自然图像和 50% 的篡改图像用于训练 SVM 分类器,用余下的一半图像作为测试集,分别进行 20 次实验,取平均值作为本文算法在该数据库上的检测结果。

### 4.3 分类结果及比较

采用本文算法对 Columbia 大学篡改图像检测数据库进行测试,实验结果如表 1 所列。利用全局马尔可夫转移概率矩阵作为特征向量,对真实图像和篡改图像进行分类,达到了 85.41% 的真正率 TPR (TP Rate)、87.69% 的真负率 TNR (TN Rate)、86.54% 的准确率 (Accuracy) 和 87.67% 的精确率 (Precision)。

表 1 本文算法的实验结果

批示	本文算法
TP Rate (%)	85.41
TN Rate (%)	87.69
Accuracy (%)	86.54
Precision (%)	87.67

本文算法与 Ng 提出的拼接图像检测算法<sup>[6]</sup>的比较结果如表 2 所列,两种算法都是在 Columbia 大学篡改图像检测数据库上进行测试。从表 2 中可以看到,本文算法的分类真正率、准确率和精确率均比 Ng 所提算法的真正率、准确率和精确率有较大程度的提高。

表 2 本文算法与 Ng 算法的比较

	特征	TPR (%)	Accuracy (%)	Precision (%)
Ng 算法	双相干特征	59.21	62.59	63.54
	双相干、预测误差、边缘特征	80.98	71.48	68.14
本文算法	转移概率矩阵特征	85.41	86.54	87.67

**结束语** 本文提出了一种基于马尔可夫链模型的篡改伪造图像盲取证算法,该算法在分析自然图像统计特性的基础上,利用马尔可夫链模型对图像 DCT 交流系数之间的相关性进行建模,提取马尔可夫链的状态转移概率矩阵作为特征向量,采用支持向量机实现对自然图像和篡改图像的有效分类。实验结果表明,针对 Columbia 大学提出的图像篡改检测数据库,本文算法性能明显优于所提算法的性能。

### 参考文献

[1] Fridrich J, Soukal D, Lukas J. Detection of copy-move forgery in digital images [C]// Proceedings of Digital Forensic Research Workshop. Cleveland, OH, 2003

[2] 吴琼, 李国辉, 孙韶杰, 等. 基于小波和奇异值分解的图像复制伪造区域检测 [J]. 小型微型计算机系统, 2008, 29(4): 730-733  
Wu Q, Li G H, Sun S J, et al. Detection of copy forgery regions in the image based on wavelet and singular value decomposition

[J]. Journal of Chinese Computer Systems, 2008, 29(4): 730-733

[3] Sutcu Y, Coskun B, Sencar H T, et al. Tamper detection based on regularity of wavelet transform coefficients [C]// Proceedings of IEEE ICIP. 2007

[4] Hsu Y F, Chang S F. Image splicing detection using camera response function consistency and automatic segmentation [C]// Proceedings of ICME. 2007

[5] Popescu A C, Farid H. Exposing digital forgeries in color filter array Interpolated Images [J]. IEEE Transactions on Signal Processing, 2005, 53(10): 3948-3959

[6] Ng T T, Chang S F, Sun Q B. Blind detection of photomontage using higher order statistics [C]// Proceedings of IEEE International Symposium on Circuits and Systems. Vancouver, Canada, 2004

[7] 崔霞, 童学锋, 宣国荣, 等. 基于双向马尔可夫模型的 JPEG 图象隐写分析 [J]. 计算机工程与应用, 2007, 43(23): 32-34  
Cui X, Tong X F, Xuan G R, et al. JPEG steganalysis based on bidirectional markov model [J]. Computer Engineering and Applications, 2007, 43(23): 32-34

[8] 边肇祺, 张学工, 等. 模式识别 (第 2 版) [M]. 北京: 清华大学出版社, 2000  
Bian Z Q, Zhang X G, et al. Pattern Recognition (2nd edition) [M]. Beijing: Tsinghua University Press, 2000

[9] Ng T T, Chang S F. A data set of authentic and spliced image blocks; ADVENT Technical Report; 203-2004-3 [R]. Columbia University, 2004

[10] Calphoto. A database of photos of plants, animals, habitats and other natural history subjects [OL]. [http://calphotos.berkeley.edu/cgi/img\\_query?query\\_src=photos\\_index & where-lifeform=Animal-Mammal](http://calphotos.berkeley.edu/cgi/img_query?query_src=photos_index&where-lifeform=Animal-Mammal)

[11] Chang C C, Lin C J. LIBSVM: A Library for Support Vector Machines [OL]. <http://www.csie.ntu.edu.tw/~cjlin/libsvm>

[12] Carvalho T J, Riess C, Angelopoulou E, et al. Exposing digital image forgeries by illumination color classification [J]. IEEE Transactions on Information Forensics and Security, 2013, 8(7): 1182-1194

[13] 杨健梅, 黄添强, 江伟坚. 基于人脸色温的拼接图像篡改检测 [J]. 山东大学学报 (工学版), 2013, 43(5): 24-30  
Yang J M, Huang T Q, Jiang W J. Splicing image tamper detection based on human face color temperature [J]. Journal of Shandong University (Engineering Science), 2013, 43(5): 24-30

[14] 王波, 孔祥维, 尤新刚. 利用颜色滤波矩阵估计原图的拼接定位盲取证方法 [J]. 电子与信息学报, 2012, 34(1): 200-206  
Wang B, Kong X W, You X G. A blind forensics method for image splicing based on original image estimation using color filter array interpolation [J]. Journal of Electronics & Information Technology, 2012, 34(1): 200-206