# 一种针对彩色图像的泛盲掩密分析方案

陈丹陈原 王育民 (西安电子科技大学 ISN 国家重点实验室 西安 710071)

**摘 要:**基于 Farid(2002)年泛盲掩密分析算法对彩色图像检测的局限性,提出了新的泛盲掩密分析算法。该算法 利用彩色图像的亮度和色度进行分类特征矢量的选择。通过矢量方向滤波技术(VDF)对图像进行方向化处理,并以 滤波前后图像之间的矢量方向相关度为色度的分类特征。该方法大大减少了分类特征的数目,而且克服了 Farid 算 法忽略 RGB3 个颜色通道间相关性的缺陷。分别利用该算法和 Farid 的算法对 JSteg, F5 和 EZStego 软件掩密后的 图像进行检测,比较实验结果表明,对于不同长度的嵌入消息,该算法具有更好的检测性能。 关键词:彩色图像检测,泛盲掩密分析,矢量方向滤波器,矢量方向相关度,RBF 网络

中图分类号: TP391 文献标识码: A 文章编号: 1009-5896(2005)10-1542-07

## A Universal Blind Steganalysis for Color Images

Chen Dan Chen Yuan Wang Yu-min (State Key Lab. on ISN, Xidian University, Xi'an 710071, China)

**Abstract** On the basis of the limitation of Farid's universal blind steganalysis for color images, a new universal blind steganalysis scheme is proposed. In view of the luminance and the chrominance of the image, the class feature vector is extracted. A new approach is presented to extract the feature for the chrominance component that the image is directional processed using the vector directional filter, and then the vector direction correlation between the image and its directional filtered image is calculated. It reduces the number of the features enormously, and overcomes the fault of Farid's scheme that neglects the inherent correlation between the RGB channels, which is propitious to simplify the design of classifier and to improve the performance of the classifier. Applying the scheme and Farid's scheme to detecting the stego images produced by JSteg, F5, and EZStego respectively, the comparison of these simulation results shows that the scheme is quite more efficient for different size of messages embedded.

**Key words** Color images detecting, Universal blind steganalysis, Vector directional filters, Vector direction correlation, RBF network

## 1 引言

掩密分析(Steganalysis)是信息隐藏(Information hiding)领 域内的重要研究内容。它是掩密术(又称隐写术, Steganography)的对抗技术,主要是利用掩密术固有的弱点, 即信息的嵌入会引起载体数据分布特性或统计特性的改变, 分析各种可能的载体信息,旨在检测、提取或破坏非法隐藏 于这些载体数据中的秘密消息。掩密分析技术的研究成果不 但可以直接应用于军事领域和国防安全,而且也将推动掩密 技术的进一步发展。

迄今为止,多数的掩密分析方法都是针对某种特定掩密

算法的<sup>[1]</sup>。这些方法虽然能够达到较准确的分析结果,但是 灵活性和可扩展性都很差。由于隐藏学科自身的特点,掩密 算法可能并不公开,而且掩密算法将层出不穷,所以无法得 知要攻击算法的具体细节。Farid<sup>[2,3]</sup>首先将掩密分析看成两类 模式的识别问题,即通过合理选择特征和设计分类器来区分 没有隐藏消息的载体图像和隐藏了消息的掩密图像,并由此 提出泛盲掩密分析(Universal blind steganalysis)的概念。这一 概念的提出,使得我们可以针对某一类而不是某一个掩密算 法来检测隐藏消息的存在,从而大大提高了掩密分析算法的 适用范围和检测效率。

然而,对于彩色图像, Farid 提出分别计算图像 RGB3

<sup>2004-04-13</sup> 收到, 2005-03-14 改回

<sup>&</sup>quot;十五" 计划通信预研(41001040303)资助课题

个通道的统计量来构成区分载体图像和掩密图像的特征矢 量。这种做法不但因为选取了太多的特征而复杂度较高,而 且由于所提取的分类特征忽略了彩色图像RGB 3 个颜色通 道间的相关性而检测率较低。本文将分别从分类特征维数、 单个特征的有效性和特征之间的相关性3个方面分析Farid算 法对于彩色掩密图像检测的局限性,并提出新的分析算法, 即分别从图像的亮度和色度角度选择特征,这样不但可以综 合RGB分量上的改变量,而且更适合于分析变换域上的掩密 算法。对于色度分量,提出计算矢量方向相关度的方法,大 大减少了特征数目。分类器采用了RBF网络<sup>[4]</sup>,因为RBF网 络在解决非线性分类问题中具有收敛速度快、不易陷入局部 极小点、鲁棒性好等优点。利用本文算法和Farid的算法对 JSteg,F5和EZStego<sup>[5-7]</sup>掩密后的图像进行检测,实验结果表 明,对于不同长度的嵌入消息,本文算法的掩密检测率较 Farid的算法都有了大幅度的提高。

本文的组织结构如下:第2节介绍泛盲掩密分析的系统 模型。第3节分析彩色图像掩密算法的特点和 Farid 的泛盲 分析算法的缺陷。第4节提出彩色图像掩密分析方案,首先 从图像中分别提取亮度和色度分类特征,然后基于 RBF 神经 网络进行分类器设计。第5节是仿真实验,测试算法对 JSteg, F5和 EZStego 掩密后的图像的检测性能,并将其与 Farid 的 算法进行比较。第6节是全文的总结。

### 2 泛盲掩密分析系统模型

与一般的模式识别系统相同<sup>(4)</sup>, 泛盲掩密分析过程也分 为学习和判决两个部分,如图1所示。学习过程包括从大量 的训练样本图像(包括载体图像和掩密图像两类样本)中提取 并选择有利于分类的特征矢量,以此构造掩密分析分类器并 对分类器进行训练,直到满足一定的精度要求。判决过程是 利用学习过程中建立的分类器对被测图像进行分类。这样, 即使在不知道掩密算法的情况下,也可以通过比较和分析学 习样本,观察出由于消息嵌入所带来的数据特性变化,从而 提取出可用于分类的特征并设计分类器,最终获得测试图像 掩密与否的信息。

该模型中,分类特征的选择是非常重要的,它强烈地影 响到分类器的设计与性能,因为如果对不同类别所选的特征 的差别很大,则比较容易设计出高性能的分类器;反之,则 有可能使得分类器难于构造或者分类效率不高。因此,泛盲 掩密分析系统成功的关键就是选出利于分类的最有效特征。



#### 3 Farid 泛盲分析算法的缺陷

彩色图像是非常好的载体数据,可以为秘密信息的传输 提供丰富的载体资源。首先,因为数据量大,冗余成分也比 较多,只要通过精巧的设计,就能隐藏大容量的秘密消息。 其次,现成的彩色图像数量要比灰度图像多得多,经数码相 机或扫描仪形成的图像大都是彩色图像,用彩色图像作为载 体既资源丰富又方便快捷。所以,对于掩密分析而言,彩色 图像应是重要的检测对象。然而,Farid<sup>[2,3]</sup>的算法对于彩色掩 密图像的检测率却并不高。究其原因,主要是特征的选择不 够合理。

前面已经指出,分类特征的有效性是影响泛盲掩密分析 检测的关键因素。这里,我们分别从分类特征维数、单个特 征的有效性和特征之间的相关性 3 个方面对 Farid 算法所选 特征的有效性进行分析。

(1) 分类特征维数 一般而言,在样本数不是很多的情况下,高维特征空间无论是从计算的复杂程度还是对分类器性能来看都是不适宜的。通常,在设计分类器之前都要研究如何对其降维,这已经成为一个重要的课题<sup>[4]</sup>。Farid的算法首先是针对灰度图像提出的,其选取 12(*n*-1)个小波系数统计量和 12(*n*-1)个小波系数误差统计量,共 24(*n*-1)个统计量作为区分载体图像和掩密图像的特征矢量,其中*n*为小波分解级数。对于彩色图像,Farid提出分别计算RGB 3 个通道的统计量,即共有 24(*n*-1)×3 个统计量构成分类特征矢量。为了获取更多的分类信息,*n*不能太小,如*n*取 3,则有 144 个特征,勿庸置疑,这么多的特征将会导致分类器的高度复杂。

(2)单个特征的有效性 我们对Farid算法的每个分类特 征做方差分析<sup>[8]</sup>以检测其区分载体图像和掩密图像的有效 性。取 500 幅由数码相机拍摄的自然图像(剪裁成 640×480), 分别计算它们的小波系数统计量和小波系数误差统计量构 成特征矢量,然后用JSteg掩密软件在这 500 幅图像中分别嵌 入消息形成掩密图像,按照相同的方法计算特征矢量。表 1 给出了部分特征在方差分析中的F值和p值(按RGB 3 个颜色 通道分别排列)。从实验数据可以看出,并非所有的特征对于 分类都是有利的,其中约有 50%的特征(表中的黑体数据)对 分类是无效的。

(3) R, G, B3 通道的特征之间的相关性 如果分类特征

之间存在很强的相关性,那么所观测的数据将在很大程度上 有信息的重叠,这样不但会导致冗余数据过多而效率降低, 更不利于分类器的设计和性能。从上面的实验中,可以发现 一个显著的特点:表1的同一行(相同分解级和相同分解方向 上的统计量)上R,G,B所对应的F值或p值都非常接近, 也就是说,这些特征对分类具有相同的效应。这说明R,G,B 3 通道的特征数据之间存在着很强的相关性。为了进一步证 实这一点,分别求取这些特征之间的相关系数。表2给出了 R和G通道,R和B通道,G和B通道相应特征之间的相关 系数值(因篇幅所限仅列出了部分数据),其中相关系数的绝 对值大于等于0.5的个数占总体系数的99.31%,而大于等于 0.9 的个数达到总体个数的88.54%。由此可见,从Farid 算 法得到的样本数据中将含有大量冗余,进而不利于分类判 决。

事实上,现有的彩色图像掩密算法,为了减少感知性失 真,都会利用RGB 3 个通道的相关性来尽量减少颜色值的改 变。它们通常不会在 3 个通道独立地隐藏秘密信息,而是仅 在某一个通道上隐藏,或者是将图像的颜色值作为一个整体 或矢量<sup>[5]</sup>,通过调整颜色值的序号来嵌入信息。更多的算法 则是在变换域上隐藏信息,比如JSteg<sup>[6]</sup>,F5<sup>[7]</sup>和OutGuess<sup>[9]</sup>算 法,它们总是先将图像从RGB空间变换到YUV空间,然后分 别在Y,U,V通道上嵌入信息。显然,Farid分别计算RGB 3 个通道的统计量的做法是不合理的,因为对于空间域的掩密 算法,掩密图像在每个通道的改变量都很小,而检测率和改 变量是成比例的,改变量越大,检测率就越高;反之,则检 测率越低。而对于变换域的掩密算法,由于信息是在YUV空 间嵌入的,在RGB空间进行分析将会损失掉一些利于检测的 信息。为此,应重新考虑分类特征的选择。

#### 4 彩色图像掩密分析方案

#### 4.1 彩色图像特征选择算法

上一节分析了 Farid 的算法对于彩色掩密图像检测的局限性,本节我们将有针对性的提出新的分析算法。

考虑到彩色图像掩密算法对每个颜色通道的改变量都 很小的特点,将图像从 RGB 空间变换到 YUV 空间(即亮度 和色度空间)上来分析,变换公式为:

$$\begin{pmatrix} \mathbf{Y} \\ \mathbf{U} \\ \mathbf{V} \end{pmatrix} = \begin{pmatrix} 0.299 & 0.587 & 0.114 \\ -0.147 & -0.289 & 0.436 \\ 0.615 & -0.515 & -0.100 \end{pmatrix} \begin{pmatrix} \mathbf{R} \\ \mathbf{G} \\ \mathbf{B} \end{pmatrix}$$
(1)

其中 Y 分量是图像的光亮度信号,U 和 V 分量是两种色度信号。变换后,不但去除了分量之间的强相关性,还具有如下

优点:对于空间域的掩密算法,它综合了 RGB3 个通道上的 改变量,从而能够更多地获取掩密与否的信息量;对于变换 域上的掩密算法,则不会损失任何对检测有利的信息,因为 这些算法本身就是在 YUV 空间嵌入信息的。

特征选择的具体过程如图 2 所示。对于亮度信号,即Y 分量,它反映了彩色图像的灰度,我们沿用Farid的算法,计 算 24(n-1)个统计量,n取为 3。而对色度信号,因为有U和V 两个分量,如果分别计算统计特征,特征数目就会很大,而 且色度信号反映了彩色图像的色彩,分别处理将无法反映整 体色彩的改变量。因此,我们提出评价整体色彩改变量的方 法,即利用矢量方向滤波(VDF)技术<sup>[10,11]</sup>,对图像进行方向 化处理形成滤波图像,然后计算原图像和其方向滤波图像间 的矢量方向相关度。下面将详细介绍彩色图像的方向化处理 和矢量方向相关度的概念。



(1)像的方向化处理 图像的每个像素值都由一个3维 矢量构成,矢量的方向表示了颜色的色度<sup>[11]</sup>。为了获得图像 的色度信息,可以仅对矢量的方向进行滤波处理,这就是 VDF技术,VDF技术一般用于图像恢复、边缘检测和图像分 割等领域。由于VDF可以产生图像矢量方向的最优估计,也 就是色度值的最优估计,所以我们将其用来获取图像色度分 量的分类特征。这里我们简单介绍VDF方法的基本原理,详 细论述可以参考文献[11]。

设一个多通道信号为  $f(x): z^{l} \rightarrow z^{m}$ 。W是大小为 n 的滤 波处理窗口。W 中的像素记为  $x_{i}, i = 1, 2, \dots, n$ 。  $f(x_{i})$  是由 m个信号通道定义的矢量空间中的 m 维矢量,简记为  $f_{i}$ 。对 于彩色图像,有 l = 2, m = 3。

令输入集为 { $f_i$ , i = 1, 2, ..., n},则对应于  $f_i$ 的  $\alpha_i$  角定义为

$$\alpha_i = \sum_{j=1}^n A(\boldsymbol{f}_i, \boldsymbol{f}_j), \quad i = 1, 2, \cdots, n$$
(2)

其中  $A(f_i, f_j)$  是矢量  $f_i$ 和  $f_j$  之间的角度。一般情况下,  $0 \le A(f_i, f_j) \le \pi$ , 而对于彩色图像,  $0 \le A(f_i, f_j) \le \pi/2$ 。  $\alpha_i$ 的大小顺序:

$$\alpha_{(1)} \leq \alpha_{(2)} \leq \cdots \leq \alpha_{(r)} \leq \cdots \leq \alpha_{(n)}$$
(3)  
意味着同样顺序的  $f_j$ :

$$f^{(1)} \le f^{(2)} \le \dots \le f^{(r)} \le \dots \le f^{(n)}$$
 (4)

式(4)中第1项构成了 BVDF(Basic Vector Directional Filter)的 输出,前 r 项构成了 GVDF(Generalized Vector Directional Filter)的输出。

表1 部分特征在方差分析中的 F 值和 p 值

G和 B 通道相应	特征之间的相关系	数值				
(	3	В				
F	р	F	р			
57.3129	0.0000	120.1572	0			
585.5954	0	417.0951	0			
0.0950	0.7579	0.1003	0.7515			
7.5710	0.0060	7.3450	0.0068			
0.3694	0.5435	0.5693	0.4507			
7.5171	0.0062	8.1673	0.0044			
2.2407	0.1347	1.4460	0.2295			

表 2 R和G、R和B、 R

index	R		G		В	
	F	р	F	р	F	р
1	105.3913	0	57.3129	0.0000	120.1572	0
2	462.5126	0	585.5954	0	417.0951	0
3	0.0999	0.7520	0.0950	0.7579	0.1003	0.7515
4	7.3078	0.0070	7.5710	0.0060	7.3450	0.0068
5	0.4530	0.5011	0.3694	0.5435	0.5693	0.4507
6	6.1315	0.0134	7.5171	0.0062	8.1673	0.0044
7	1.0628	0.3028	2.2407	0.1347	1.4460	0.2295
8	3.1684	0.0754	2.9964	0.0838	2.9514	0.0861
9	165.7369	0	92.3306	0	188.9332	0
10	279.5582	0	484.7299	0	238.7617	0
11	0.0381	0.8454	0.0414	0.8387	0.0370	0.8475
12	8.3285	0.0040	8.9298	0.0029	8.3713	0.0039
13	0.1065	0.7443	0.0276	0.8682	0.1930	0.6605
14	0.7959	0.3725	2.0652	0.1510	0.1683	0.6817
15	0.0915	0.7624	0.2188	0.6401	0.1472	0.7013
16	0.6635	0.4155	0.4987	0.4803	0.0121	0.9123
17	865.7938	0	857.4566	0	852.1989	0
18	434.0759	0	423.1323	0	433.1511	0
19	2.6429	0.1043	2.7407	0.0981	2.7527	0.0974
20	17.6972	0.0000	17.8249	0.0000	17.8985	0.0000
21	9.8450	0.0018	9.8433	0.0018	9.8685	0.0017
22	18.4648	0.0000	14.4948	0.0001	9.2695	0.0024
23	1.2443	0.2649	1.3289	0.2493	1.6039	0.2056
24	0.6922	0.4056	1.4635	0.2267	2.1648	0.1415

index	RG		RB		GB	
	Cover	Stego	Cover	Stego	Cover	Stego
1	0.9923	0.9759	0.9832	0.9848	0.9875	0.9799
2	0.9173	0.9818	0.8044	0.9595	0.8979	0.9751
3	0.9940	0.9903	0.9932	0.9898	0.9952	0.9932
4	0.9947	0.9952	0.9982	0.9968	0.9957	0.9960
5	0.9401	0.9528	0.8651	0.8824	0.8792	0.8821
6	0.9555	0.9668	0.9020	0.9398	0.9287	0.9383
7	0.8697	0.9335	0.8817	0.8844	0.9157	0.8838
8	0.9691	0.9536	0.9361	0.9347	0.9403	0.9528
9	0.9873	0.9793	0.9868	0.9851	0.9835	0.9747
10	-0.0644	0.7836	0.8257	0.9387	-0.1194	0.7097
11	0.9991	0.9989	0.9963	0.9965	0.9957	0.9961
12	0.9968	0.9965	0.9941	0.9953	0.9879	0.9907
13	0.9663	0.9636	0.9377	0.9424	0.9466	0.9385
14	0.8333	0.8757	0.8869	0.8850	0.7587	0.8160
15	0.9438	0.9581	0.9322	0.9399	0.9491	0.9423
16	0.9437	0.9591	0.9417	0.9304	0.8961	0.9349
17	0.9924	0.9994	0.9909	0.9981	0.9904	0.9982
18	0.9581	0.9987	0.9369	0.9975	0.9478	0.9978
19	0.9999	0.9998	0.9980	0.9971	0.9977	0.9965
20	1.0000	0.9999	0.9993	0.9985	0.9992	0.9982
21	0.9774	0.9884	0.9666	0.9745	0.9746	0.9763
22	0.8993	0.8868	0.8248	0.8354	0.8426	0.8719
23	0.9741	0.9849	0.9573	0.9686	0.9774	0.9692
24	0.9402	0.9700	0.8686	0.9439	0.9053	0.9556

$$f^{(1)} = \text{BVDF}[f_1, f_2, \cdots, f_n]$$
 (5)

 ${f^{(1)}, f^{(2)}, \dots, f^{(r)}} = \text{GVDF}[f_1, f_2, \dots, f_n], 1 \le r \le n$  (6) 由此可见, BVDF 输出了与其他矢量之间角度和最小的矢量, GVDF 则是对它的推广,输出了角度和最小的矢量集合,以 做进一步的幅度处理。这里,因为仅需获取色度信息,所以 我们只对图像做 BVDF 处理,滤波窗口大小取为 5。

经过BVDF方向化处理之后,产生了新的方向滤波图像。 如果掩密图像改变了载体图像的色度,那么载体图像和掩密 图像与其相应的方向滤波图像之间的色度差异肯定是不同 的,基于此,我们定义矢量方向相关度的概念来度量色度的 改变量。

(2) 矢量方向相关度 设样本图像和其方向滤波图像的 像素值矢量分别为 f(i,j)和  $f^{*}(i,j)$ ,其中  $(i,j),1 \le i$ ,  $j \le N$  是像素点的位置, $N \times N$  是像素数目。由于彩色图像有  $0 \le A(f_i, f_j) \le \pi/2$ ,所以矢量方向相关度定义为

$$R(f,f^*) = 1 - \frac{1}{N^2} \sum_{i,j=1}^{N} \frac{2}{\pi} \cos^{-1} A(f,f^*)$$
(7)

综上所述,对于彩色图像,我们对亮度和色度分别提取 分类特征,即计算图像 Y 分量的 24×(3-1)=48 个小波统计 量和 VDF 滤波前后图像之间的矢量方向相关度 R(f,f\*), 这样共有 49 个特征作为区分载体图像和掩密图像的一组分 类特征矢量,特征数目大幅减少,而且综合考虑了 RGB 3 通道的相关性及彩色图像掩密算法的特点,这对简化分类器 设计和提高分类器性能是十分有利的。

#### 4.2 分类器设计

分类器采用RBF网络<sup>[4]</sup>,即径向基函数网络,是一种较 常用的前馈型神经网络。它可以实现对任意非线性函数的逼 近,而且用于模式分类具有收敛速度快、不易陷入局部极小 点、鲁棒性好和易于实现等优点,基本结构如图 3 所示:



这种网络的特点是:只有一个隐层,隐层单元采用径向 基函数作为输出特性,输入层到隐层之间的权值均固定为1; 输出节点为线性求和单元,隐层到输出节点之间的权值可 调,因此输出为隐层的加权求和。

由于掩密分析是一种两类模式的识别问题,所以相应的 RBF网络结构比较简单,如图4所示。图中x为输入矢量,*n*  表示隐层输出矢量, $w_i^1$ 表示第i个隐层神经元的权值矢量,即隐层神经元权值矩阵 $W^1$ 的第i行, $b^1$ 是隐层的阈值矢量, $W^2$ 是输出权值矢量, $b^2$ 是输出阈值。网络的输入维数为m,隐层神经元个数为s个,输出个数为1,隐层神经元采用高斯函数作为传递函数,输出层的传递函数为线性函数。

我们根据训练样本采用迭代方法确定隐层神经元的个数,即每迭代一次就增加一个神经元,直到平方和误差下降 到目标误差以下或者隐层神经元个数达到最大值时迭代停止。输出单元的权值直接由最小二乘法计算。



### 5 仿真实验

我们将分别测试上述算法对JSteg、F5 和EZStego<sup>[5-7]</sup>掩密 后的图像的检测能力。采用由数码相机拍摄的自然图像(640 ×480)进行仿真实验。取 500 幅这样的图像作为未嵌消息的 载体图像,存为JPEG(或GIF)格式,然后用JSteg或F5(或 EZStego)软件在每幅图像中嵌入消息,形成掩密图像。这样, 共 1000 幅图像构成了训练样本图像集。类似地,另取 1000 幅图像构成测试样本图像集。

首先,从训练样本图像集中计算分类特征矢量集,即分别计算每幅图像的亮度统计量和色度特征,然后用该特征矢量集训练RBF网络,目标误差精度设为10<sup>-3</sup>,最大迭代次数 设为1000。当 RBF 网络训练完成之后,再通过测试样本图 像集来检验分类效果。为比较性能,采用 Farid 的算法对同 样的样本集进行检测,分别提取图像 RGB 3 通道的统计量构 成特征矢量(小波分解级数 n 也取为 3),并以此训练 RBF 网 络,误差精度和最大的迭代次数设为相同。

图 5 - 图 7 分别给出了该算法和 Farid 的算法对用 JSteg, F5 和 EZStego 软件嵌入不同长度消息后的图像的检测结果。 图 5 和图 6 获得了相近的实验数据,是因为 JSteg 和 F5 都是 JPEG 压缩域的掩密软件,它们都分别在图像的 Y,U 和 V 分量上嵌入了信息。而 EZStego 由于是在空间域上嵌入信息, 所以检测性能稍差。设检测阈值为 0.6,对于 JSteg 软件,嵌 入消息达 40%以上时,就能被本文算法有效检测,而 Farid 的算法只能检测嵌入消息达 80%以上的图像;对于 F5 软件 嵌入消息达 35%以上,可被本文算法检测,而 Farid 的算法 只能检测嵌入消息达 85%以上的图像;对于 EZStego 软件, 嵌入消息达 70%以上,可被本文算法检测到,而 Farid 算法 几乎不能检测掩密的图像。可以看出,本文的算法具有更好 的检测性能。



图 7 用 EZStego 掩密后图像的检测率

为了更清楚地说明本文所选的特征矢量优于 Farid 的算法,我们再从统计学的角度进行分析和比较。实际上,各类 样本之所以能够分开是因为它们位于特征空间中的不同区 域,显然这些区域之间距离越大类别可分性就越大,样本也 就越容易被正确地分类。因此,我们采用文献[4]第八章(第 186页)中定义的类别可分性判据来衡量所选特征的有效性。 设 x<sup>(0)</sup>, x<sup>(1)</sup>分别是从载体图像和掩密图像中计算出的特征矢 量,则可分性判据定义为

$$\boldsymbol{J}_{d}(\boldsymbol{x}) = \operatorname{tr}(\boldsymbol{S}_{\boldsymbol{w}}^{-1}\boldsymbol{S}_{\boldsymbol{b}})$$
$$\boldsymbol{S}_{\boldsymbol{w}} = \sum_{i=0}^{1} P_{i} \frac{1}{n_{i}} \sum_{k=1}^{n_{i}} (\boldsymbol{x}_{k}^{(i)} - \boldsymbol{m}_{i}) (\boldsymbol{x}_{k}^{(i)} - \boldsymbol{m}_{i})^{\mathrm{T}}$$
$$\boldsymbol{S}_{b} = \sum_{i=0}^{1} P_{i} (\boldsymbol{m}_{i} - \boldsymbol{m}) (\boldsymbol{m}_{i} - \boldsymbol{m})^{\mathrm{T}}$$
(8)

其中,  $S_w$ 为类内离散度矩阵,  $S_b$ 为类间离散度矩阵,  $m_i$ (i = 0, 1)是各类样本的均值向量, m是样本集的总均值向量,  $P_i$ (i = 0, 1)是先验概率。分别以上面实验中得到的特征矢量样本集 计算可分性判据 $J_d$ 值。表 3 给出了本文算法和Farid算法的 $J_d$ 值比较, 对于不同长度的嵌入消息,本文算法的 $J_d$ 值普遍比 Farid算法的高。实验数据再次表明,本文所选的特征比Farid 所选特征更利于区分载体图像和掩密图像。

#### 6 结束语

彩色图像是非常重要的掩密分析对象。本文在分析了 Farid 算法对彩色图像的局限性的基础上,提出新的掩密分析 算法,即分别从图像的亮度和色度角度选择特征,这样不但 可以综合 RGB 分量上的改变量,而且更适合于分析变换域 上的掩密算法。首次提出以 VDF 滤波前后图像之间的矢量 方向相关度为色度分类特征的思想,大大减少了特征数目, 这对于简化分类器设计和提高分类器性能是十分有利的。分 类器采用了 RBF 网络,其在解决非线性分类问题中具有收敛 速度快、不易陷入局部极小点、鲁棒性好和易于实现的优点。 实验结果表明,本文算法的掩密检测性能较 Farid 的算法有 了大幅度的提高。

从仿真结果进一步可以看出,一旦确定了所使用的分类器,泛盲掩密分析的性能将主要取决于分类特征的选择。所以,进行掩密分析时,必须针对掩密算法的具体特点选择相应的图像特征,才能获得尽可能高的检测效率。

嵌入消息	长度(%)	10	20	30	40	50	60	70	80	90	100
Jsteg 2	Farid 算法	0.1763	- 0.0688	0.2236	0.2790	0.5687	0.6934	1.9493	2.4998	-2.7743	3.5528
	本文 算法	0.5910	0.6433	0.8967	- 0.7222	0.9814	1.6457	2.9898	3.0107	3.4571	3.8489
F5 F3 Far 算道 算道	Farid 算法	0.1035	0.1496	0.2091	- 0.2474	0.2673	0.6484	0.9276	0.7698	3.2806	6.5817
	本文 算法	0.2625	0.2872	0.6501	0.7341	1.3158	1.4601	- 2.9873	3.8522	6.7916	8.9263
EzStego	Farid 算法	0.0234	0.0198	0.0135	0.0742	0.1148	- 0.2126	0.3689	0.7491	0.8823	1.5479
	本文 算法	0.0627	- 0.0437	0.0981	0.1066	0.1743	0.2099	0.4124.	0.6999	0.9157	2.4631

= 2	米미ㅋ	ᅚᄭᄮᆄᆇ		しょた六
オマ・ショ		/ + + +	11店/	Er 40
	×		11110/1	PL371

## 参考文献

- Fridrich J, Goljan M. Practical steganalysis state of the art. Proc. SPIE Photonics West, Electronic Imaging 2002. Security and Watermarking of Multimedia Contents, San Jose, California, Jan. 2002: 1 – 13.
- [2] Farid H. Detecting hidden messages using higher-order statistical models. Proc. Int'l Conf. Image Processing, , Rochester, NY, IEEE Press 2002: 905 – 908.
- [3] Lyu S, Farid H. Detecting hidden messages using higher-order statistics and support vector machines. Proc. 5<sup>th</sup> Int'l Workshop on Information Hiding, Noordwijkerhout, the Netherlands, Springer -Verlag, 2002: 340 – 354.
- [4] 边肇祺,张学工等编著.模式识别(第二版).北京:清华大学 出版社,2000,第8章,第11章.
- [5] Romana Machado, EZStego. <u>http://www.stego.com</u>.
- [6] Derek Upham, JPEG-Jsteg-V4. <u>http://www.funet.fi/pub/c</u> rypt/steganography/jpeg-jsteg-v4.diff.gz.

- [7] Andreas Westfeld, F5. <u>http://wwwrn.inf.tu-dresden.de/~Westfeld/</u> <u>f5.html.</u>
- [8] 魏宗舒等著. 概率论与数理统计教程. 北京:高等教育出版社, 1983: 241-253.
- [9] Provos N. OutGuess universal steganography. August 1998.
- [10] <u>http://www.outguess.org/</u>.
- [11] Trahanias P E and Venetsanopoulos A N. Vector directional filters ——A new class of multichannel image processing filters. *IEEE Trans. on Image Processing*, 1993, 2(10): 528 – 534.
- [12] Trahanias P E, Karakos D, Venetsanopoulos A N. Directional processing of color images: theory and experimental results. *IEEE Trans. on Image Processing*, 1996, 5(6): 868 – 880.
- 陈 丹: 女,1976年生,博士生,研究方向为信息隐藏、网络安 全.
- 陈 原: 女, 1978年生, 博士生, 研究方向为密码学、信息安全.
- 王育民: 男,1936年生,教授,博士生导师,研究方向为编码理 论、密码学、信息安全.