2018-2019春季
信息隐藏课程
第9讲 通用隐写分析

中国科学院信息工程研究所
INSITUTE OF INFORMATION ENGNEERING CAS

SKLOIS
信息安全国家重点实验室

赵险峰

中国科学院信息工程研究所
信息安全国家重点实验室

2018年11月

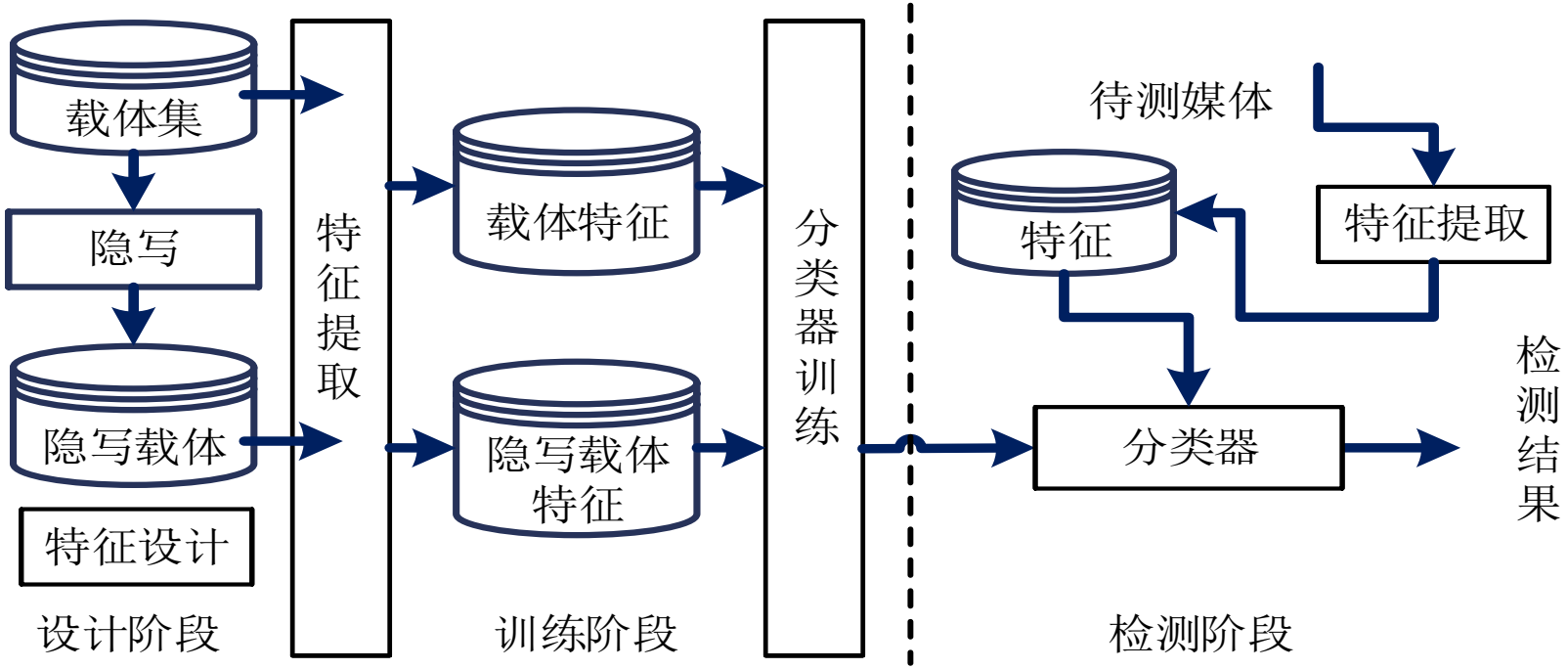
纲要

1. 基本概念与过程
2. 分类问题与SVM
3. 基于SPAM特征的分析
4. 基于Markov特征的分析
5. 基于融合校准特征的分析
6. 文献阅读推荐

1-1 基本概念（通用隐写分析）

- 通用隐写分析（Universal Steganalysis）中使用的分类特征适用于识别多个或者多类隐写，主要分为针对空域隐写与压缩编码域隐写的两类方法
- 通用隐写分析在特征提取中，充分考虑了隐写对载体邻域相关性造成的影响等普遍事实，计算得到高维分类特征，基于支持向量机、线性分类器等手段进行隐写样本的识别
- 由于以上分类手段均需要基于已标注样本进行监督学习（Supervised Learning），因此，通用隐写分析与专用分析一样，也可能需要知道被分析算法、样本的基本情况或者参数
- 假设分析者知道算法与载体情况（假设算法或软件被敌手获得）有利于设计出更安全的隐写，因此，基于监督学习的通用隐写分析有很好的安全验证作用
- 但若对将检测样本的算法与媒体特性不知，隐写分析者只能通过判断和猜测构造训练样本，进行盲检测

1-2 通用隐写分析基本过程



2-1 分类问题与判别函数

- 假设存在 M 个模式 $\omega_1, \omega_2, \cdots, \omega_M$，这些模式一般可由从观测样本中提取的 n 维特征 $\boldsymbol{x} = (x_1, x_2, \cdots, x_n)^{\mathrm{T}}$ 表征。模式识别的目的是，基于检测特征 $\boldsymbol{x}$ 判别其所属的模式 ω_i，这一般基于判别函数完成
- 判别函数是定义在 n 维空间中的一个单值函数。对应以上 M 个模式，可以分别记为 $g_1(\boldsymbol{x}), g_2(\boldsymbol{x}), \cdots, g_M(\boldsymbol{x})$，若 $\boldsymbol{x}$ 属于 ω_i 类，有：$g_i(\boldsymbol{x}) > g_j(\boldsymbol{x})$，$i, j = 1,2, \cdots, M$，$j \neq i$
- 几何上看，基于判别函数可形成分割各类模式特征向量的分界面，在 ω_i 与 ω_j 的分界面上有：$g_i(\boldsymbol{x}) = g_j(\boldsymbol{x})$

2-2 二分类与多分类

- 如果 $M = 2$，称以上模式识别问题为二分类问题，当$M > 2$，称为多分类问题，分类方法被称为分类器
- 如果以上判别函数容易实现，则可以实现多分类。但是，一般只有两个模式的二分类问题比较容易解决，因此，多分类问题往往用二分类方法解决：
  - 方法：构造$C_M^2 = \frac{M(M-1)}{2}$个二分类，每次试图分类出属于 ω_i 与 ω_j 类的样本，经过C_M^2种检测后，如果一个样本被分到 ω_i 的次数最多，就判别其属于该类模式

2-3 样本标注与分类器训练

- 一般先假设判别函数的基本参数形式，参数通过训练样本集进行估计。这种训练过程称为监督学习（Supervised Learning）
- 在学习开始前，需要对样本类型进行标注，标注的样本集可以表示为：$S = \{(\boldsymbol{x}_1, y_1), (\boldsymbol{x}_2, y_2), \cdots, (\boldsymbol{x}_N, y_N)\}$，其中，$\boldsymbol{x}_i$ 是样本的模式特征向量，y_i 是它的模式类别标识
- 以下介绍常用的二分类线性分类器与支持向量机判别函数的确定方法，其中的$y_i \in \{+1, -1\}$

2-4 线性分类器

- 如果判别函数是线性的，相应的分类器就称为线性分类器。对于模式 ω_i，线性判别函数可以表示为

$$g_i(\boldsymbol{x}) = w_{i1}x_1 + w_{i2}x_1 + \cdots + w_{in}x_1 + w_{i0}, \qquad i = 1,2,\cdots,M$$

- 记 $\boldsymbol{w}_i = (w_{i1}, w_{i2}, \cdots, w_{in})^{\mathrm{T}}$，则上式可表示为

$$g_i(\boldsymbol{x}) = \boldsymbol{w}_i^{\mathrm{T}}\boldsymbol{x} + w_{i0}, \qquad i = 1,2,\cdots,M$$

- 在只有两类模式的情况下，$g_1(\boldsymbol{x}) = \boldsymbol{w}_1^{\mathrm{T}}\boldsymbol{x} + w_{10}$，$g_2(\boldsymbol{x}) = \boldsymbol{w}_2^{\mathrm{T}}\boldsymbol{x} + w_{20}$，若$\boldsymbol{x}$属于$\omega_1$，有$g_1(\boldsymbol{x}) > g_2(\boldsymbol{x})$，即有$(\boldsymbol{w}_1^{\mathrm{T}} - \boldsymbol{w}_2^{\mathrm{T}})\boldsymbol{x} + (w_{10} - w_{20}) > 0$。令$\boldsymbol{w} = \boldsymbol{w}_1^{\mathrm{T}} - \boldsymbol{w}_2^{\mathrm{T}}$，$w_0 = w_{10} - w_{20}$，得到**分界超平面**

$$g(\boldsymbol{x}) = \boldsymbol{w}^{\mathrm{T}}\boldsymbol{x} + w_0 = 0$$

- 线性二分类器的决策方法是：如果$g(\boldsymbol{x}) > 0$，判断$\boldsymbol{x}$属于ω_1；否则属于ω_2；以下将说明，投影函数$g(\boldsymbol{x})$的值反映了特征到分界面的距离

2-5 线性分界的几何解释

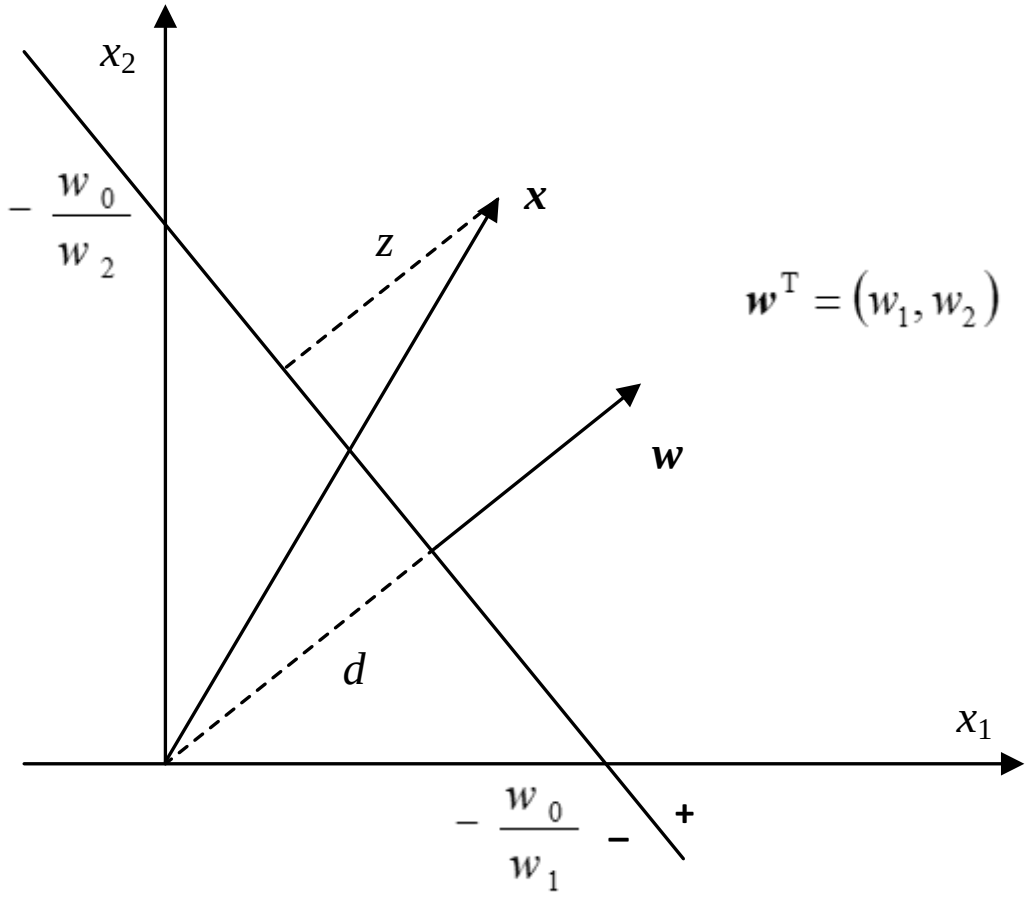
- 若 $\boldsymbol{x}$ 在分界面上，$g(\boldsymbol{x})=0$，对分界面上的任意两点$\boldsymbol{x}_1,\boldsymbol{x}_2$，有 $\boldsymbol{w}^{\mathrm{T}}\boldsymbol{x}_1+\mathrm{w}_0=\boldsymbol{w}^{\mathrm{T}}\boldsymbol{x}_2+\mathrm{w}_0=0$，即$\boldsymbol{w}^{\mathrm{T}}(\boldsymbol{x}_1-\boldsymbol{x}_2)=0$
- $\boldsymbol{x}_1-\boldsymbol{x}_2$可代表分界面方向上（而不是分界面上）任一向量，说明 **$\boldsymbol{w}$ 垂直于分界面，与向分解面的投影平行**
- 从图中可以得知，原点到分界面的距离为$d=\dfrac{|w_0|}{\sqrt{w_1^2+w_2^2}}=\dfrac{|w_0|}{\|\boldsymbol{w}\|}$, (1)
- 将任一向量 $\boldsymbol{x}$ 向分界面投影，记投影点上的向量为 $\boldsymbol{x}_p$（位于分界面），投影距离为 z，有$\boldsymbol{x}=\boldsymbol{x}_p+z\times\dfrac{\boldsymbol{w}}{\|\boldsymbol{w}\|}$，由于

$$g(\boldsymbol{x})=g\left(\boldsymbol{x}_p+z\times\frac{\boldsymbol{w}}{\|\boldsymbol{w}\|}\right)=\boldsymbol{w}^{\mathrm{T}}\left(\boldsymbol{x}_p+z\times\frac{\boldsymbol{w}}{\|\boldsymbol{w}\|}\right)+w_0=z\times\|\boldsymbol{w}\|$$

- 因此，$z=\dfrac{g(\boldsymbol{x})}{\|\boldsymbol{w}\|}$ (2)
  - 说明 $g(\boldsymbol{x})$的值可以度量特征到分界面的距离

2-6 线性分界的几何解释

注意：这里坐标上的分量x_1, x_2不同于前页的二维的 $\boldsymbol{x}_1, \boldsymbol{x}_2$



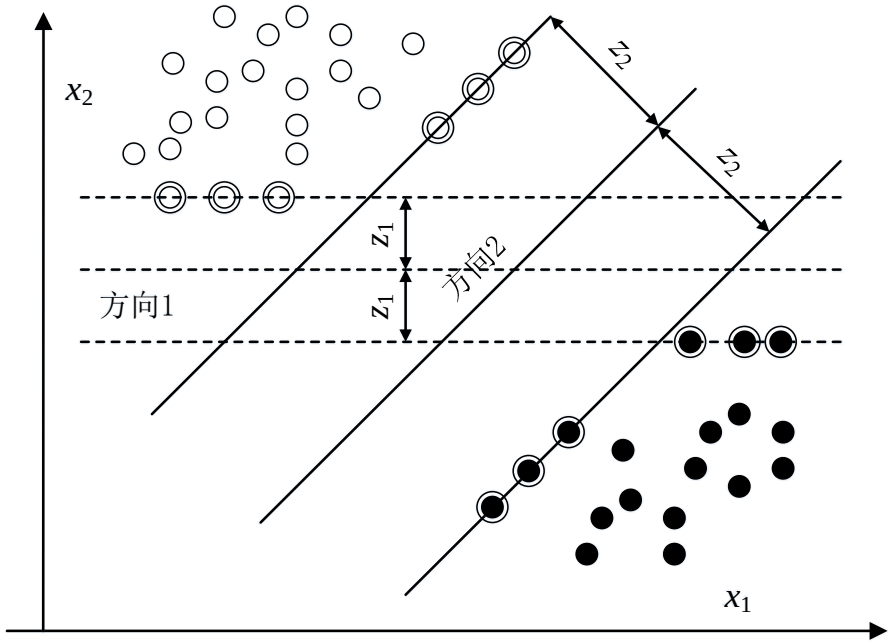
$$\boldsymbol{w}^{\mathrm{T}} = (w_1, w_2)$$

2-7 支持向量机

- 支持向量机（Support Vector Machine, SVM）是一类分界得到优化的分类器，分界准则是使得不同模式向量离分界面的距离都最大，并且分界面不偏向任何一类模式
- 间隙（Margin）的概念是指，在分界面的垂直方向上不同类向量的距离，例如下图中，方向1分界面的间隙是 $2z_1$，方向2分界面的间隙是$2z_2$，其中，距离分界面最近的点 $\boldsymbol{x}$ 被称为支持向量（Support Vector），这些点用圆圈进行了标注
- 线性情况下

  z_i 可以用 $g(\boldsymbol{x})/\|\boldsymbol{w}\|$ 表示，$\boldsymbol{x}$为支持向量

  (根据前页公式(2))



2-8 二分类线性SVM（优化问题）

- 需要使得$g(\boldsymbol{x})$满足：对$\boldsymbol{x} \in \omega_1$，$g(\boldsymbol{x}) \geq C$；对$\boldsymbol{x} \in \omega_2$，$g(\boldsymbol{x}) \leq -C$。这可以表示为

$$\boldsymbol{w}^{\mathrm{T}}\boldsymbol{x} + w_0 \geq 1, \qquad \boldsymbol{x} \in \omega_1$$
$$\boldsymbol{w}^{\mathrm{T}}\boldsymbol{x} + w_0 \leq -1, \qquad \boldsymbol{x} \in \omega_2$$

- 以上右侧是正负1原因：由于不等式中全部系数待定，并且乘上常数比例不等式仍成立，因此将右侧限定为正负1不失一般性。在此情况下，对距离分界面最近的向量有：$g(\boldsymbol{x}) = 1$，$\boldsymbol{x} \in \omega_1$；$g(\boldsymbol{x}) = -1$，$\boldsymbol{x} \in \omega_2$。因此间隙为(根据前面公式(2))

$$\frac{1}{\|\boldsymbol{w}\|} + \frac{1}{\|\boldsymbol{w}\|} = \frac{2}{\|\boldsymbol{w}\|}$$

- 最大化间隙需要最小化$\|\boldsymbol{w}\|$，在获得训练样本后，这等价于求解以下优化问题：$\text{minimize } J(\boldsymbol{w}, w_0) = \frac{1}{2}\|\boldsymbol{w}\|^2$

$$\text{subject to } y_i\left(\boldsymbol{w}^{\mathrm{T}}\boldsymbol{x}_i + w_0\right) \geq 1, \qquad i = 1,2,\cdots,N$$

2-9 不可分情况下的SVM（问题）

- 每个样本向量 $\boldsymbol{x}_i$ 存在三种情况：$y_i(\boldsymbol{w}^{\mathrm{T}}\boldsymbol{x}_i + w_0) \geq 1$并正确分类，$0 \leq y_i(\boldsymbol{w}^{\mathrm{T}}\boldsymbol{x}_i + w_0) < 1$并正确分类，$y_i(\boldsymbol{w}^{\mathrm{T}}\boldsymbol{x}_i + w_0) < 0$并错误分类，分别对应正确位于支持向量之后、处于分界面与支持向量之间以及超出分界面的三种情况，它们可以用以下不等式刻画：

$$y_i(\boldsymbol{w}^{\mathrm{T}}\boldsymbol{x}_i + w_0) \geq 1 - \xi_i$$

- 其中，ξ_i 为松弛变量，以上三类情况对应 $\xi_i = 0$，$0 < \xi_i \leq 1$，$\xi_i > 1$，在优化中，既希望间隙较大，也希望 ξ_i 的值总和较小，则优化问题变为：$\text{minimize } J(\boldsymbol{w}, w_0) = \frac{1}{2}\|\boldsymbol{w}\|^2 + C\sum_{i=1}^{N}\xi_i$

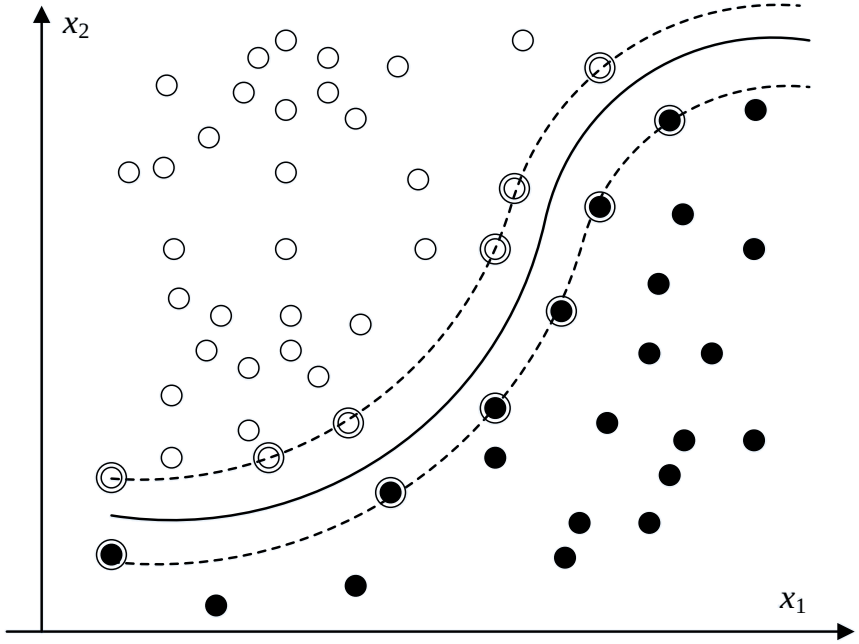
$$\text{subject to } y_i(\boldsymbol{w}^{\mathrm{T}}\boldsymbol{x}_i + w_0) \geq 1 - \xi_i, \qquad \xi_i \geq 0, i = 1,2,\cdots,N$$

- C 是控制目标函数两项之间关系的参数，控制对错分样本的惩罚程度，也称**惩罚系数**，这样求解得到的分类器称为软间隙（Soft Margin）SVM，在求解中一般需要在一个范围内进行搜索C

2-10 非线性SVM

- 对很多分类问题，采用线性分类器是不可分，但是采用非线性分类器是可分的
- 为了获得非线性的分界面，需要将模式向量变换到新的空间：$\boldsymbol{x} \longmapsto \phi(\boldsymbol{x}) \in H$，在这个空间中按照线性处理，并要求 H 是Hilbert空间，即存在内积操作并对应原空间一个函数K:

$$\langle \phi(\boldsymbol{x}), \phi(\boldsymbol{z}) \rangle = K(\boldsymbol{x}, \boldsymbol{z})$$



2-11 非线性SVM（求解）

- 在模式识别的优化问题中，向量之间往往仅仅存在内积操作，例如：

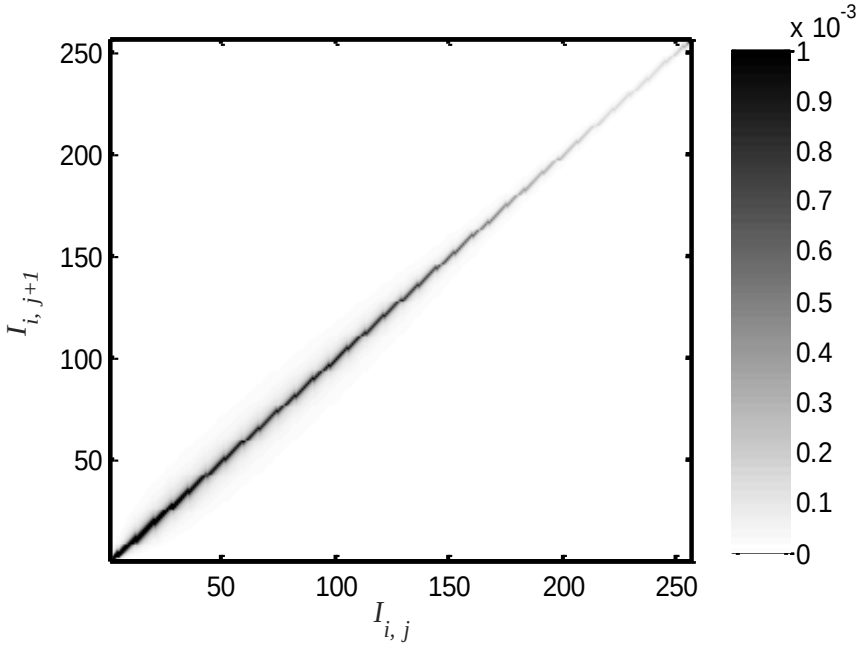
$$\max_{\boldsymbol{\lambda}} \left(\sum_{i=1}^{N} \lambda_i - \frac{1}{2} \sum_{i,j} \lambda_i \lambda_j y_i y_j K(\boldsymbol{x}_i, \boldsymbol{x}_j) \right)$$

$$\text{subject to } \sum_{i=1}^{N} \lambda_i y_i = 0,\ 0 \le \lambda_i \le C, i = 1,2,\cdots,N$$

- 函数K被称为核函数，线性情况下是普通内积，常用的核函数有：
  - 多项式核：$K(\boldsymbol{x}, \boldsymbol{z}) = (\boldsymbol{x}^{\mathrm{T}} \boldsymbol{z} + 1)^q, q > 0$
  - 高斯核：$K(\boldsymbol{x}, \boldsymbol{z}) = \exp(-\gamma \|\boldsymbol{x} - \boldsymbol{z}\|^2), \gamma > 0$
  - 双曲核：$K(\boldsymbol{x}, \boldsymbol{z}) = \tanh(\beta \boldsymbol{x}^{\mathrm{T}} \boldsymbol{z} + \sigma)$
- 核函数的参数（如γ, q等）也一般要在训练中进行搜索

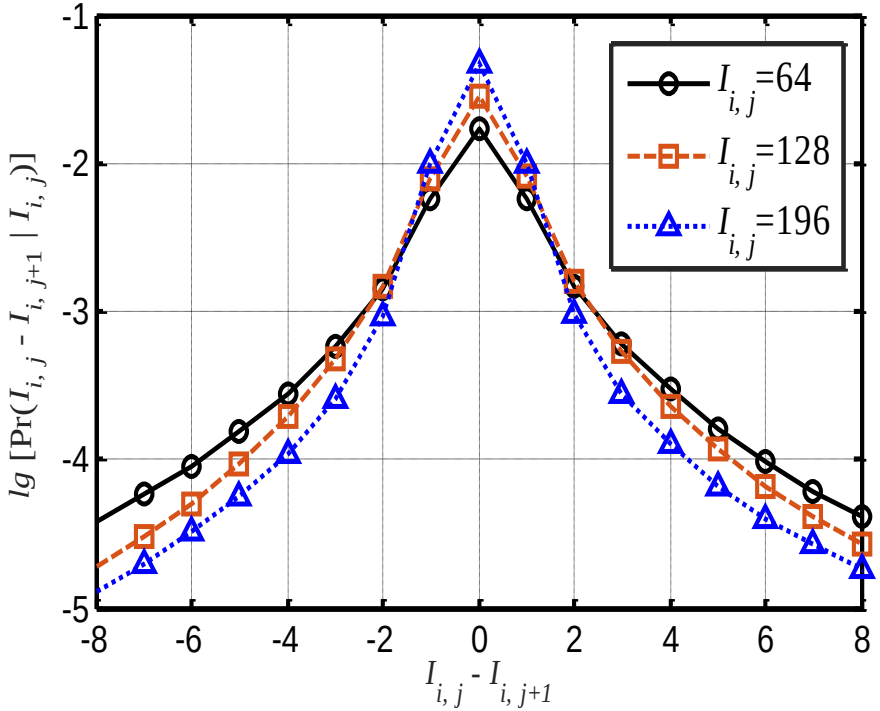
3-1 SPAM特征的分析（现象1）

- 针对自然拍摄的数字图像，一个基本事实是，图像的邻域像素具有非常大的相关性
- 记$m \times n$图像的像素为$\{I_{i,j} | I_{i,j} \in \{0,1,\cdots,255\}, i \in \{1,\cdots,m\}, j \in \{1,\cdots,n\}\}$，以上事实部分表现为，对$\Pr(I_{i,j} = x, I_{i,j+1} = y)$与$\Pr(I_{i,j} = x, I_{i+1,j} = y)$，当$x, y$接近的时候较大



3-2 SPAM特征的分析（现象2）

- 由于隐写主要改变了载体信号的噪声分布，这个现象可以更好地通过邻域像素的差值分布反映，对水平方向差的条件概率 $\Pr\left(I_{i,j} - I_{i,j+1} | I_{i,j}\right)$，可以发现在$\Pr\left(0 | I_{i,j}\right)$处分布最密
- 右图有：$I_{i,j} - I_{i,j+1} \in [-8,8]$，$[-8,8]$ 的限制是将考察范围限制在典型范围，不但有利于分类（压制内容），也有利于降低提取的特征维度
- 结论：自然图像的像素在很大程度上符合Markov模型



3-3 SPAM特征的分析（原理）

- SPAM分析方法原理：针对隐写对空间域邻域相关性的破坏，计算像素不同方向上的差值并进行统计分布特征的提取，通过非线性SVM进行隐写图像的识别，主要面向检测图像空间域隐写
- 记$\{\leftarrow, \rightarrow, \downarrow, \uparrow, \nwarrow, \searrow, \swarrow, \nearrow\}$为像素的8个方向，它们用于标记相邻像素在这些方向上的差值，如$D_{i,j}^{\rightarrow} = I_{i,j} - I_{i,j+1}$为为从左向右水平方向的相邻像素差值
- SPAM特征分为一阶与二阶两类特征，主要刻画8个方向上一阶与二阶邻像素差值转移概率的变化

3-4 SPAM特征的分析（1阶特征提取）

- 统计8个方向的相邻像素差值一阶转移概率，例如在从左向右水平方向统计$M_{u,v}^{\rightarrow}=\Pr\left(D_{i,j+1}^{\rightarrow}=u|D_{i,j}^{\rightarrow}=v\right)$，统计范围由$u,v\in[-T,T]$限定的截断区间限定，称$T$为截断长度
- 如果$\Pr\left(D_{i,j}^{\rightarrow}=v\right)=0$，则$M_{u,v}^{\rightarrow}=0$。类似地得到$M_{u,v}^{\leftarrow}$，$M_{u,v}^{\downarrow}$，$M_{u,v}^{\uparrow}$，$M_{u,v}^{\searrow}$，$M_{u,v}^{\nwarrow}$，$M_{u,v}^{\swarrow}$，$M_{u,v}^{\nearrow}$。
- 为了降低总体特征维度并增加特征的稳定性，对相同一对$u,v\in[-T,T]$进行以下合并：

$$F_{u,v}^{+}=\frac{1}{4}\left(M_{u,v}^{\rightarrow}+M_{u,v}^{\leftarrow}+M_{u,v}^{\downarrow}+M_{u,v}^{\uparrow}\right) \quad (7)$$

$$F_{u,v}^{\times}=\frac{1}{4}\left(M_{u,v}^{\searrow}+M_{u,v}^{\nwarrow}+M_{u,v}^{\swarrow}+M_{u,v}^{\nearrow}\right) \quad (8)$$

- 以上两类特征分别有k个，$k=(2T+1)^2$，这$2k$维特征成为一阶SPAM特征，记为$\mathbf{F}^{1\mathrm{st}}$

3-5 SPAM特征的分析（2阶特征提取）

- 对8个方向统计相邻像素差值二阶转移概率，例如在从左向右水平方向统计$M_{u,v,w}^{\rightarrow} = \Pr\left(D_{i,j+2}^{\rightarrow} = u | D_{i,j+1}^{\rightarrow} = v, D_{i,j}^{\rightarrow} = w\right)$
- 统计范围由$u, v, w \in [-T, T]$限定的截断区间限定；如果$\Pr(D_{i,j+1}^{\rightarrow} = v, D_{i,j}^{\rightarrow} = w) = 0$，则$M_{u,v,w}^{\rightarrow} = 0$
- 类似地得到$M_{u,v,w}^{\leftarrow}$，$M_{u,v,w}^{\downarrow}$，$M_{u,v,w}^{\uparrow}$，$M_{u,v,w}^{\searrow}$，$M_{u,v,w}^{\nwarrow}$，$M_{u,v,w}^{\swarrow}$，$M_{u,v,w}^{\nearrow}$。对相同一组$u, v, w \in [-T, T]$进行公式（7）、（8）类似的合并，但此时$k = (2T+1)^3$，这$2k$维特征成为二阶SPAM特征，记为$\mathbf{F}^{2\mathrm{nd}}$
- 在T. Pevny等人的实验方案中，对一阶特征$T = 4$或者$T = 8$，这样$\mathbf{F}^{1\mathrm{st}}$的维度是$2k = 162$或者578，对二阶特征$T = 3$，因此$\mathbf{F}^{2\mathrm{nd}}$的维度是$2k = 686$

3-6 SPAM分类器与实验

- 实验采用的分类器是具有非线性高斯核$K(\boldsymbol{x},\boldsymbol{z})=\exp(-\gamma\|\boldsymbol{x}-\boldsymbol{z}\|^2)$的SVM，分别采用以上三组特征构建了3个分类器。为了获得好的效果，每个实验对惩罚参数C与核参数γ按照以下网格进行了搜索：

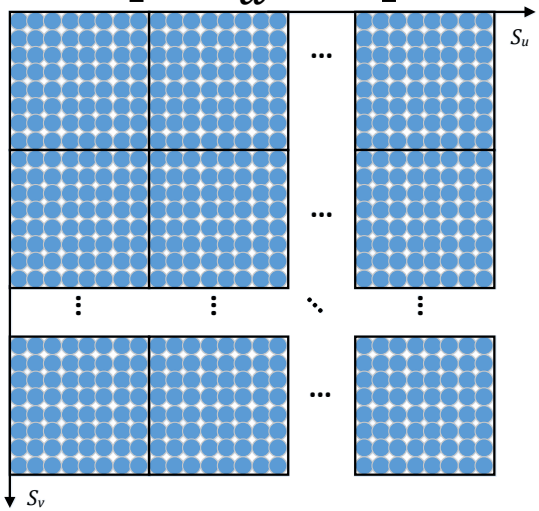
$$C\in\{0.001,0.01,\cdots,10000\}$$

$\gamma\in\left\{2^i\middle|i\in\{-d-3,\cdots,-d+3\}\right\}$，$d$为对特征维数求对数$\log_2$

- SPAM可以对LSBM等空间域隐写取得了非常好的分析效果，典型地，在LSBM嵌入率为0.25bpp时，基于约10000对训练样本，以上三个分类器在BOWS2图像库上的错误率仅分别为9.8%、12.3%与5.5%
- 由于JPEG隐写也扰动了空间域，SPAM分析方法也对这些隐写有效，基于二阶SPAM特征的方法相比当时最好的联合校准特征，在分析F5、MB、MME、PQ等JPEG隐写中，多数情况下效果仅仅稍差，但是对PQt有超出

4-1 Markov特征分析（原理）

- 基于JPEG图像的隐写往往在JPEG编码域中进行，由于JPEG系数也有比较强的相关性，因此，可以通过建立相应的Markov模型考察这种相关性的变化，借此判定隐写的存在
- 基于以上考虑，Y. Shi等人提出了基于Markov过程模型的JPEG图像隐写分析方法
- 由于JPEG编码的单元是图像的空间域 8×8 分块，因此，最后的DCT变化也是基于分块的，因此，整个DCT量化系数按照分块为单位按照空间域顺序排列。记S_u, S_v分别表示二维DCT分块量化系数阵列的水平与垂直尺寸，$F(u,v), u \in [0, S_u - 1], v \in [0, S_v - 1]$表示阵列中的系数绝对值



4-2 Markov特征分析（现象观察）

- 以下定义的相邻系数差值阵列在统计上能够较好的表达相邻DCT量化系数的相关性：

$$F_h(u,v)= F(u,v) - F(u+1,v)$$

$$F_v(u,v)= F(u,v) - F(u,v+1)$$

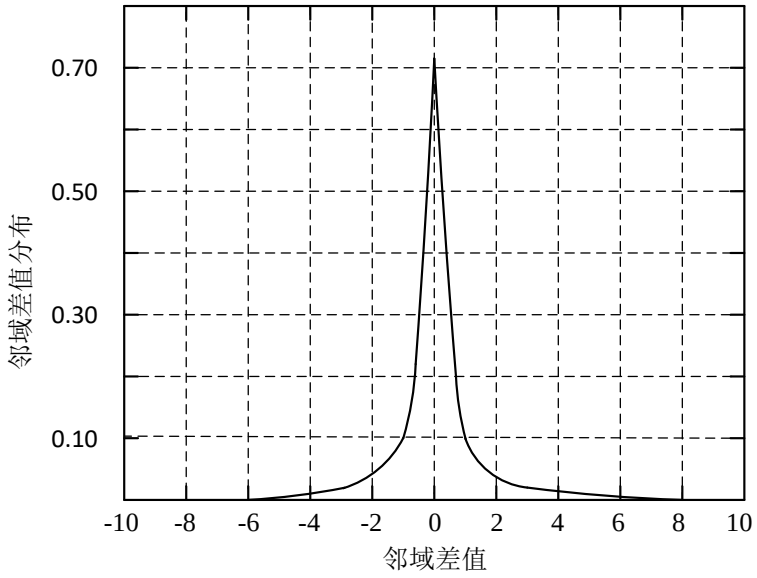
$$F_d(u,v)= F(u,v) - F(u+1,v+1)$$

$$F_m(u,v)= F(u+1,v) - F(u,v+1)$$

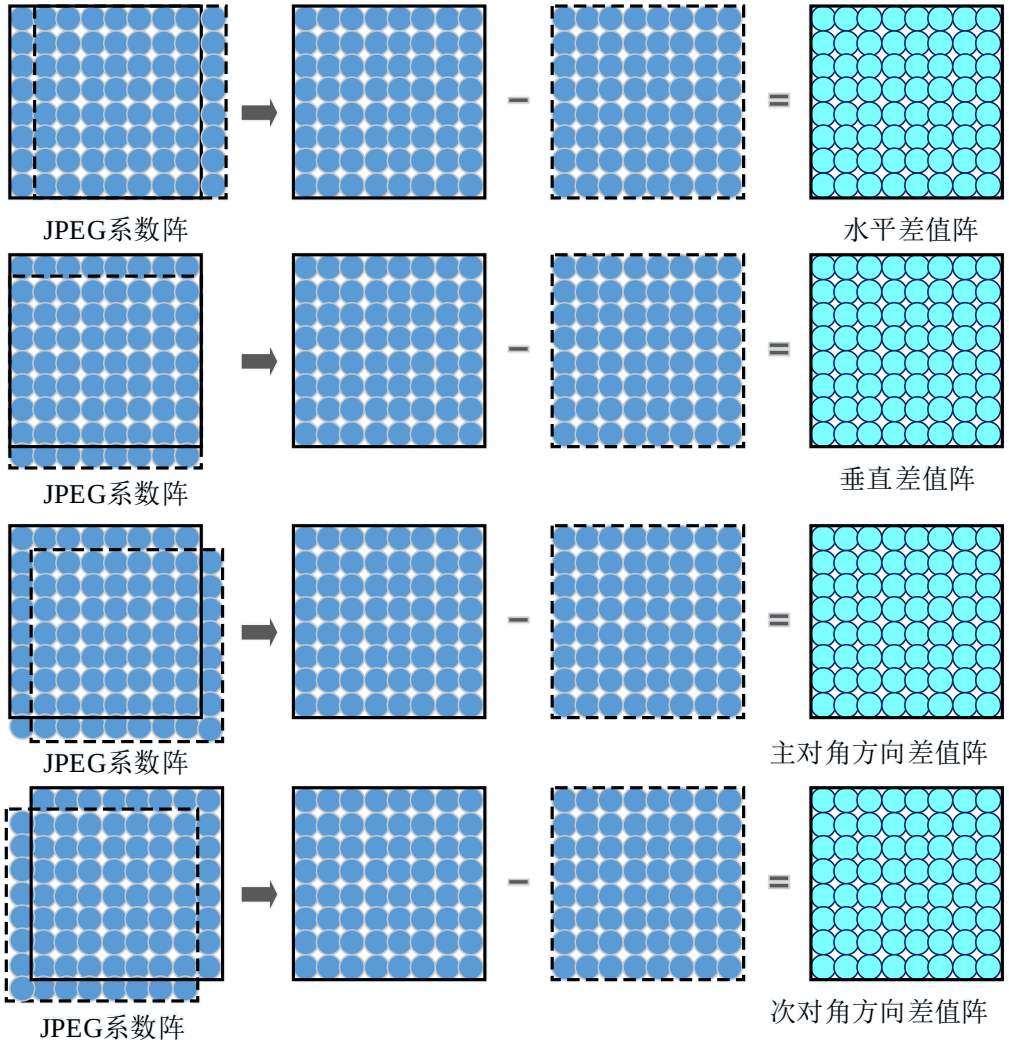
- 其中，$F_h(u,v), F_v(u,v), F_d(u,v), F_m(u,v)$分别表示水平、垂直、主对角、次对角方向上的相邻系数差值矩阵（下页图），注意此时$u \in [0, S_u - 2], v \in [0, S_v - 2]$

4-3 Markov特征分析（相邻系数差值矩阵的相关性）

$$F_h(u,v) = F(u,v) - F(u+1,v)$$
$$F_v(u,v) = F(u,v) - F(u,v+1)$$
$$F_d(u,v) = F(u,v) - F(u+1,v+1)$$
$$F_m(u,v) = F(u+1,v) - F(u,v+1)$$



相邻系数差值的分布



SKLOIS
信息安全国家重点实验室

4-4 Markov特征分析（特征提取）

- 由于差值分布最密在0附近以及主要分布在[-4,4]之间，相邻系数差值的相关性可以基于Markov过程建立以下转移概率表达：

$$M_h(a,b)=\Pr\left(F_h(u+1,v)=b\middle|\,F_h(u,v)=a\right)=\frac{\sum_{u,v}\delta(F_h(u,v)=a,F_h(u+1,v)=b)}{\sum_{u,v}\delta(F_h(u,v)=a)}$$

$$M_v(a,b)=\Pr\left(F_v(u,v+1)=b\middle|\,F_v(u,v)=a\right)=\frac{\sum_{u,v}\delta(F_v(u,v)=a,F_v(u,v+1)=b)}{\sum_{u,v}\delta(F_v(u,v)=a)}$$

$$M_d(a,b)=\Pr\left(F_d(u+1,v+1)=b\middle|\,F_d(u,v)=a\right)=\frac{\sum_{u,v}\delta(F_d(u,v)=a,F_d(u+1,v+1)=b)}{\sum_{u,v}\delta(F_d(u,v)=a)}$$

$$M_m(a,b)=\Pr\left(F_m(u,v+1)=b\middle|\,F_d(u+1,v)=a\right)=\frac{\sum_{u,v}\delta(F_d(u+1,v)=a,F_d(u,v+1)=b)}{\sum_{u,v}\delta(F_d(u+1,v)=a)}$$

- 其中，$a,b\in[-T,\cdots,T]$，$u\in[0,S_u-2],v\in[0,S_v-2]$，$\delta$ 为计数函数，当括弧里的条件成立输出1，否则输出0

4-5 Markov特征分析（分类器与实验）

- 将以上转移概率作为特征向量。由于每组转移概率有$(2T+1)^2$个，因此特征总维数为$4\times(2T+1)^2$，在实验中，取$T=4$，因此这个方法实际使用了$4\times(2\times4+1)^2=324$维特征
- 在分类器构造中，该方法采用了线性核的SVM，对OutGuess、F5、MB等JPEG隐写取得了较高的准确率，典型地，在采用7千对图像样本进行训练，在0.2bpnc的嵌入率下，对它们的检测准确率分别达到95.5%、87.0%与97.3%

5-1 融合校准特征的分析（原理）

- 在通用隐写分析方法发展初期，特征向量主要基于一组原理近似的提取方法获得，例如，提出了基于小波高阶特征、基于JPEG域Markov特征、基于DCT系数的分布特征等的隐写分析方法
- T. Pevny等人提出的融合校准特征（Merged Calibrated Features）是通用隐写分析法中出现的一组重要特征，它的设计思想体现了将各类互补的分析特征进行融合的思想，以及体现了部分隐写分析特征需要通过校准等技术抑制载体内容干扰的思想，在二分类与多分类隐写分析中取得了非常好的效果
  - **融合校准特征组合。通过校准改造了DCT特征与Markov两组特征**
  - **大量采用了校准技术。**以下用 J_1 代表待检测的样本，J_2 代表其校准后的版本（一般是在空间域去掉4行4列后得到的JPEG文件），前面已多次说明，校准版本的很多统计特性接近原始的载体，因此，两个版本的特征差值 $F(J_1) - F(J_2)$ 往往能更好地刻画隐写，更好压制载体内容干扰

5-2 融合校准特征提取（第一组：扩展DCT特征）-1

- 记分块DCT系数为$d_{ij}(k), i,j=1,\cdots,8, k=1,\cdots,n_B$，$n_B$为块数，系数均考虑亮度分量。第一组特征分为6个部分，一共193维：
  - 总校准直方图采样。记总直方图为H，它是全部分块上所有DCT系数的直方图，记H_l为该直方图上值为l的数值（出现概率），则取以下11个特征进入特征向量：$H_l(J_1) - H_l(J_2),\ l \in [-5,5]$

$$H_l = \frac{1}{64n_B}\sum_{i,j=1}^{8}\sum_{k=1}^{n_B}\delta\left(l, d_{ij}(k)\right)$$

  - AC分量（Mode）直方图校准差采样。记$L=\{(i,j)\}=\{(1,2),(2,1),(3,1),(2,2),(1,3)\}$，对$(i,j)\in L, l\in[-5,5]$，$h_l^{ij}$是分块DCT变换系数$(i,j)$频率分量上$l$值系数出现的概率，将以下$5\times 11$个特征纳入特征向量：$h_l^{ij}(J_1) - h_l^{ij}(J_2),\ l\in[-5,5],(i,j)\in L$

$$h_l^{ij} = \frac{1}{n_B}\sum_{k=1}^{n_B}\delta\left(l, d_{ij}(k)\right)$$

$\delta(x,y)=1$仅当$x=y$，否则$\delta(x,y)=0$

5-3 融合校准特征提取（第一组：扩展DCT特征）-2

- 接前页：
  - AC分量数值频次校准差采样。记$L=\{(i,j)\}=\{(2,1),(3,1),(4,1),(1,2),(2,2),(3,2),(1,3),(2,3),(1,4)\}$，对$(i,j)\in L$，$l\in[-5,5]$，将以下$11\times 9$个特征纳入特征向量：$g_{ij}^{d}(J_1)-g_{ij}^{d}(J_2)$，$d\in[-5,5]$，$(i,j)\in L$。由于由于频率分量标号之间的关系，，这类特征也被称为对偶直方图（Dual Histogram）特征

$$g_{ij}^{l}=\frac{1}{n_B(l)}\sum_{k=1}^{n_B}\delta\left(l,d_{ij}(k)\right)$$

$$n_B(l)=\sum_{i,j}\sum_{k=1}^{n_B}\delta\left(l,d_{ij}(k)\right)$$

5-4 融合校准特征提取（第一组：扩展DCT特征）-3

- 接前页：
  - 平均相邻块JPEG系数变化程度。定义：

$$V=\frac{\sum_{i,j=1}^{8}\sum_{k=1}^{|\mathbf{I}_r|-1}\left|d_{ij}(\mathbf{I}_r(k))-d_{ij}(\mathbf{I}_r(k+1))\right|+\sum_{i,j=1}^{8}\sum_{k=1}^{|\mathbf{I}_c|-1}\left|d_{ij}(\mathbf{I}_c(k))-d_{ij}(\mathbf{I}_c(k+1))\right|}{|\mathbf{I}_r|+|\mathbf{I}_c|}$$

$\mathbf{I}_r$与$\mathbf{I}_c$分别表示按行序与列序的分块标号集合，$\mathbf{I}_r(k)$是其中第k个分块，$|\mathbf{I}_r|+|\mathbf{I}_c|$表示行块与列块的数量和，这样，特征维度为1

  - 块效应。块效应反映分块之间相关性的变化，从幅度与能量变化上分别定义为$B_\alpha=\frac{\sum_{i=1}^{\lfloor(M-1)/8\rfloor}\sum_{j=1}^{N}\left|c_{8i,j}-c_{8i+1,j}\right|^{\alpha}+\sum_{j=1}^{\lfloor(N-1)/8\rfloor}\sum_{i=1}^{M}\left|c_{i,8j}-c_{i,8j+1}\right|^{\alpha}}{N\lfloor(M-1)/8\rfloor+M\lfloor(N-1)/8\rfloor}$，$M,N$是图像的行列尺寸，$c_{i,j}$是图像空间域像素的灰度分量；$\alpha=1,2$，因此特征是2维

  - 邻块同频DCT系数的共生矩阵。定义以下块间系数联合分布的表达形式：

$$C_{st}=\frac{\sum_{i,j=1}^{8}\sum_{k=1}^{|\mathbf{I}_r|-1}\delta\left(s,d_{ij}(\mathbf{I}_r(k))\right)\delta\left(t,d_{ij}(\mathbf{I}_r(k+1))\right)+\sum_{i,j=1}^{8}\sum_{k=1}^{|\mathbf{I}_c|-1}\delta\left(s,d_{ij}(\mathbf{I}_c(k))\right)\delta\left(t,d_{ij}(\mathbf{I}_c(k+1))\right)}{|\mathbf{I}_r|+|\mathbf{I}_c|}$$

计算以下25维特征并加入特征向量：

$$C_{st}(J_1)-C_{st}(J_2),\quad (s,t)\in[-2,2]\times[-2,2]$$

5-5 融合校准特征提取（第二组：Markov特征）

- 融合校准特征中包含的第二组特征来自前面介绍的JPEG系数Markov特征
- Markov特征在联合校准特征中是经过校准后的，即$M^{(c)} = M(J_1) - M(J_2)$，并且对其4个方向上的特征组进行了合并：$\bar{M} = (M_h^{(c)} + M_v^{(c)} + M_d^{(c)} + M_m^{(c)})/4$，因此，特征维度由原来的324维下降为 $324/4 = 81$ 维

5-6 融合校准特征提取（分类器与实验）

- 合并以上两组特征，得到融合校准特征的特征维度为193+81=274 维，被称为CC-Pev-274特征
- 融合校准特征分析方法采用了高斯核的SVM，T. Pevny等人实现了二分类器与多分类器两类版本
- 二分类器构造和训练方法与之前介绍的SPAM分析方法类似，多分类器采用了构造C_M^2个二分类器进行两两分类的方法
- 在T. Pevny等人的实验中，基于以上特征向量的二分类器由3400对样本训练，采用2500个图像进行测试。实验显示，融合校准特征相比之前的特征（23维DCT特征、324维Markov特征），分别有显著的提高
- 融合校准特征分析方法也能够有效对F5、JP Hide & Seek、MB1、OutGuess、Steghide等JPEG隐写进行有效的多分类识别。

5-7 融合校准特征提取（主要2分类实验结果）

cover vs.	Message length	Detection accuracy DCT	Markov	Merged
F5	100%	99.49%	99.80%	**99.92%**
	50%	98.80%	99.20%	**99.84%**
	25%	84.54%	86.94%	**98.36%**
	cover	**99.80%**	91.53%	99.64%
JP Hide&Seek	100%	**99.88%**	98.08%	99.52%
	50%	98.56%	84.38%	**99.60%**
	25%	86.46%	27.16%	**92.01%**
	cover	99.32%	97.00%	**99.56%**
MB1	100%	99.64%	**99.96%**	**99.96%**
	50%	98.92%	**99.96%**	99.92%
	25%	86.94%	99.72%	**99.72%**
	cover	97.72%	97.20%	**99.88%**
MB2	30%	92.29%	99.92%	**100.00%**
	cover	98.92%	98.48%	**99.92%**
OutGuess	100%	99.92%	99.92%	**100.00%**
	50%	99.64%	99.68%	**99.96%**
	25%	98.36%	97.84%	**99.48%**
	cover	99.48%	98.04%	**99.76%**
Steghide	100%	99.84%	99.96%	**100.00%**
	50%	99.48%	99.92%	**99.92%**
	25%	90.93%	98.88%	**99.32%**
	cover	97.40%	98.00%	**99.92%**

5-8 融合校准特征提取（主要多分类实验结果）

	Classified as						
Embedding algorithm	Cover	F5	JP Hide&Seek	MB1	MB2	OutGuess	Steghide
F5 100%	0.00%	99.52%	0.04%	0.08%	0.04%	0.08%	0.24%
JP Hide&Seek 100%	0.32%	0.00%	99.64%	0.00%	0.00%	0.04%	0.00%
MB1 100%	0.00%	0.00%	0.04%	98.76%	0.44%	0.04%	0.72%
OutGuess 100%	0.00%	0.04%	0.04%	0.08%	0.00%	99.64%	0.20%
Steghide 100%	0.00%	0.00%	0.04%	0.12%	0.08%	0.44%	99.32%
F5 50%	0.16%	99.36%	0.00%	0.00%	0.04%	0.24%	0.20%
JP Hide&Seek 50%	0.28%	0.04%	99.60%	0.00%	0.00%	0.08%	0.00%
MB1 50%	0.00%	0.00%	0.04%	97.04%	1.36%	0.08%	1.48%
OutGuess 50%	0.04%	0.08%	0.00%	0.20%	0.12%	99.28%	0.28%
Steghide 50%	0.04%	0.00%	0.00%	0.36%	0.12%	0.76%	98.72%
MB2 30%	0.00%	0.04%	0.04%	1.08%	98.48%	0.00%	0.36%
F5 25%	1.84%	97.12%	0.20%	0.00%	0.16%	0.36%	0.32%
JP Hide&Seek 25%	8.23%	0.32%	91.45%	0.00%	0.00%	0.00%	0.00%
MB1 25%	0.12%	0.12%	0.04%	90.10%	1.92%	0.36%	7.35%
OutGuess 25%	0.52%	0.28%	0.04%	0.20%	0.08%	98.08%	0.80%
Steghide 25%	0.60%	0.04%	0.00%	0.76%	0.20%	1.44%	96.96%
Cover	99.16%	0.24%	0.44%	0.00%	0.08%	0.08%	0.00%

6 文献阅读推荐

- [1] 教材第8章
- [2] Y. Q. Shi, C. H. Chen, W. Chen. A Markov process based approach to effective attacking JPEG steganography. In: Proc. Information Hiding 2006, LNCS, vol. 4437, pp.249-264, Springer, 2007
- [3] T. Pevny, J. Fridrich. Merging Markov and DCT features for multi-class JPEG steganalysis. In Proc. SPIE Electronic Imaging, Security, Steganography, and Watermarking of Multimedia Contents IX, vol. 6505, pp. 3-1 - 3-14, 2007
- [4] T. Pevny, P. Bas, J. Fridrich. Steganalysis by subtractive pixel adjacency matrix, IEEE Trans. Information Forensics and Security, 5(2), pp. 215--224, June 2010
- [5] 模式识别教材
  - S. Theodoridis, K. Koutroumbas. Pattern Recognition (3rd Edition)（模式识别），Elsevier, 2006（机械工业出版社，影印版）
  - N. Cristianini, J. Shawe-Taylor. An Introduction to Support Vector Machines and Other Kernel-based Learning Methods（支持向量机导论）. Cambridge: Cambridge University Press, 2000（机械工业出版社，影印版）

谢谢！

中国科学院信息工程研究所
INSITUTE OF INFORMATION ENGNEERING CAS

SKLOIS
信息安全国家重点实验室