广西科学院学报
Journal of Guangxi Academy of Sciences

2007,23(4):238~241
Vol. 23,No. 4 November 2007

一种分组的信息隐藏算法*

An Algorithm of Hidden Information Based on Grouping

韦月琼,张显全,王继军,崔晨荣
WEI Yue-qiong,ZHANG Xian-quan,WANG Ji-jun,CUI Chen-rong

(广西师范大学计算机科学系,广西桂林 541004)
(Department of Computer Science,Guangxi Normal University,Guilin,Guangxi,541004,China)

摘要:将加密信息与载体图像转换成二进制序列,把加密信息每3个位分为一组,每组的加密信息与待嵌入图像低位的信息进行比较后,修改载体图像,嵌入加密信息;提取加密信息时,找出嵌入的位置,读取标志位,根据标志位取反与否读取隐藏信息。新算法在信息嵌入量不变的情况下,修改嵌入信息位少,载体图像变化小,安全性高。新算法实现简单,嵌入信息量大,隐藏效果好,具有一定的实用价值。

关键词:最低有效位 标志位 分组 取反

中图法分类号:TP301.6 **文献标识码**:A **文章编号**:1002-7378(2007)04-0238-04

Abstract:The secret information and the carrier image are turned to binary sequences, the secret sequence is divided into group, each of which contains three bits. With the results of comparing the encrypted information with the embedded image's low bits, the different bits of each group are changed and the information is embedded. When the secret is extracted, the flag bits are read after the finding of the embedded location, then the secret was found by the sign. With the same embedded quantity, the new algorithm obtains less altered bits, higher security and small changes in original image. Experiments show that this method has large hidden information capacity, good hidden effect and a certain practical value.

Key words:least significant bit,flag bit,grouping,negation

随着科学技术的迅猛发展,信息安全问题日显重要,这样也促进了信息隐藏技术[1]的发展。信息隐藏技术的主要思想是将加密信息隐藏在普通文件中,使之难以被其他人发现,从而达到信息在网络上安全传送的目的。隐写术是信息隐藏的一个重要分支,应用非常广泛。

在信息隐藏技术中,LSB算法[2]是最早出现的一种嵌入算法,该算法实现简单、嵌入速度快且嵌入量较大。文献[3]提出了一种索引数据链方法,寻找与加密信息最匹配的位置存放,实现对载体图象较少的改变;文献[4]先将加密信息置乱,后运用灰度融合和黑白图象集成的方法,再嵌入到某一个位平面中去,实现了提取加密信息不需要原始图象的嵌入;文献[5]则将混沌序列作为一种噪声的扩谱序列,把所要传送的加密信息以白噪声的扩谱信号调制隐藏在载体图像中进行安全通信;文献[6]将加密信息与第1~7中的某一位进行异或运算,得到的结果再存放到最低位,这种方法有一定的抗干扰性;文献[7]通过对调色板颜色分量进行三次排序,得到最接近的像素颜色集和颜色孤立的像素集,利用最低有效位与嵌入比特位的一致性进行颜色替换来实现对信息的隐藏;文献[8]利用图像的视觉特性,通过C均值聚类分析进行加密信息嵌入及对简单低位替换的优化调整,实现了信息的隐藏;文献[9]给出了对隐藏信息后的图像进行动态补偿的LSB信息隐藏方法,该方法隐秘性较好。

本文提出一种基于分组的信息隐藏算法。首先

收稿日期:2007-09-10
作者简介:韦月琼(1982-),女,硕士研究生,主要从事图象处理研究。
* 广西自然科学基金项目(0447035),广西教育厅项目(200607MS135),广西研究生教育创新计划项目(200610602812M37)资助。

是将加密信息变成二进制序列,其中每三位分为一组,同样地将载体图象的低位读成一个序列;然后将每一组的加密信息与待嵌入位的信息进行比较;最后根据匹配的情况修改载体图象,实现加密信息的嵌入。而提取加密信息时,找出嵌入的位置,先读取标志位,再根据标志取反与否读取隐藏信息。

1 LSB 算法原理

对于数字图象,灰度值是由多个比特平面来表示的。例如 8 位平面、16 位平面,其中每一个比特平面又可以表示成一个二值平面,称为位平面。例如 8 位位图,可以分解成 0～7 个位等平面,其中第 0 个位平面为最不重要位 LSB。从图 1 的位平面补图可以看出每个位平面对图象的影响,位平面越低,影响越小,很明显修改低 3 位的位平面对图像的影响较小,因此可以将信息隐藏在合适的低位平面中。

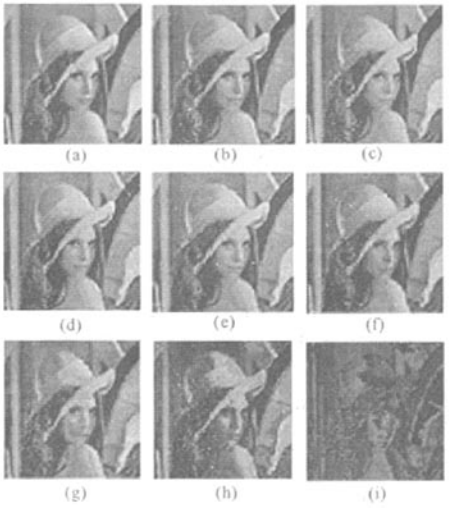

图 1 位平面补图

(a)原图;(b)位平面 0 补图;(c)位平面 1 补图;(d)位平面 2 补图;(e)位平面 3 补图;(f)位平面 4 补图;(g)5 位平面 5 补图;(h)位平面 6 补图;(i)位平面 7 补图。

大部分 LSB 算法是修改最低有效位来嵌入加密信息,类似于在原始信息上迭加了一小部分的噪声,对感官影响很小。例如 8 位位图,将信息嵌入在载体图象的最后一位,载体图象只改变 1/255,视觉上差别不大。而嵌入位置的选择又有顺序和随机两种,顺序方法容易被发现,目前所见的方法一般使用随机替换法,这样可以用密钥去控制嵌入的位置,以达到更好的效果。不同的 LSB 算法的嵌入位置和嵌入量的大小都不同,其安全性也主要依赖于这两个方面。当嵌入量增大时,LSB 算法的可靠性将随之减少,所以一般只修改最后一位,或者低两位,最多修改低四位。

2 连续三位嵌入算法

2.1 三位匹配运算

采用三位一组的方法来减少修改位。如果对应位相同,则嵌入加密信息相当于不修改载体信息低位,否则修改。那么相同位的数目越多,则修改位越少,对载体图象的破坏也越少。设 P_i、Q_j 为三位二进制有序数列,$P_i=\{a_{f(i)},a_{f(i)+1},a_{f(i)+2}\}$,$Q_j=\{b_{g(j)},b_{g(j)+1},b_{g(j)+2}\}$,定义 P_i、Q_j 的匹配数 $n(P_i,Q_j)=\sum_{k=0}^{2}a_{f(i)+k}\cdot b_{g(j)+k}$,其中·定义为:若 $a_m=b_n$ 则 $a_m\cdot b_n=1$,反之则 $a_m\cdot b_n=0$。

从匹配数的定义可知,$n(P_i,Q_j)$ 为两组三位有序数列对应位相同的个数,那么有 $0\leqslant n(P_i,Q_j)\leqslant 3$。令 $\overline{P}_i=\{\overline{a_{f(i)}},\overline{a_{f(i)+1}},\overline{a_{f(i)+2}}\}$,称 $\overline{P}_i$ 为 P_i 的取反。其中 $\overline{a_m}$ 为 a_m 的非,$\overline{P}_i$ 与 Q_j 的匹配数为 $n(\overline{P}_i,Q_j)$,由匹配数的定义,同理有 $0\leqslant n(\overline{P}_i,Q_j)\leqslant 3$。如果 $n(P_i,Q_j)=0$,P_i 与 Q_j 中对应的三位二进制数都不相同,那么 $\overline{P}_i$、Q_j 中的对应元素则会完全相同,因而有 $n(\overline{P}_i,Q_j)=3$。若 $a_m=b_n$,则 $\overline{a_m}\neq b_n$;若 $\overline{a_m}=b_n$,则 $a_m\neq b_n$,所以 $n(\overline{P}_i,Q_j)+n(P_i,Q_j)=3$。

要想得到两个二进制有序数列最大的匹配数,应该判断是否需要取反。令 P_i、Q_j 的最大匹配数 $\mathrm{num}(i)=\max(n(P_i,Q_j),n(\overline{P}_i,Q_j))$。如果 $n(P_i,Q_j)\geqslant 2$,$\mathrm{num}(i)=n(P_i,Q_j)$,$P_i$ 不需要取反;反之则将 P_i 取反。由 $n(\overline{P}_i,Q_j)+n(P_i,Q_j)=3$ 得到 $n(\overline{P}_i,Q_j)=3-n(P_i,Q_j)\geqslant 2$,这样保证 $\mathrm{num}(i)\geqslant 2$。如果将 P_i 隐藏于 Q_j 中,那么需要对与 Q_j 不相同的位进行修改。最大匹配数增大,修改的位数就会减少,根据上述最大匹配数定义可以知道最多只需要修改 1 位,即修改位比例不超过 $\frac{1}{3}$。P_i 是否取反,使用一个标志做记录,所以在 Q_j 中增加一个位用来做标志位,令 $Q'_j=\{b_{k(j)},b_{k(j)+1},b_{k(j)+2},b_{k(j)+3}\}$,则 P_i 与 Q'_j 的匹配数为 $n(P_i,Q'_j)=\sum_{k=0}^{2}a_{f(i)+k}\cdot b_{g(j)+1+k}$。对于标记位 $b_{g(j)}$,可用 0 来表示不取反,1 表示取反;当然也可以使用 1 表示不取反,0 表示取反。

例如 $P_i=\{0,0,1\}$ 与 $Q'_j=\{1,1,0,0\}$ 匹配,如果标志位为 1 表示取反、0 不取反,则匹配数如图 2 所示。Q'_j 中的第一位是标志位,$n(P_i,Q'_j)=1$,匹配见图 2(a);将 P_i 取反后(见图 2(b)),匹配数变为了

2，最大匹配数增大。

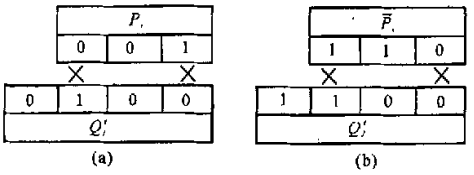

图2 匹配示意

(a)不取反，标志位存0；(b)取反，标志位存1。

2.2 位置选择

隐秘算法是将加密信息依次嵌入到选取的像素点的低位中，以达到隐藏的目的。为了提高安全性，一般选择嵌入的位置不能高度集中，如果那样会引起图象各部分统计特征的不一致，非常容易被发现，算法的安全性将会降低。本文采用均匀分组选取嵌入位置，让嵌入的加密信息能够分布均匀，假设已将加密信息转成二进制序列 $P=\{a_1,a_2,\cdots,a_M\}$，其中 M 为加密信息的个数，而取载体图象低位组成的序列为 $Q=\{b_1,b_2,\cdots,b_n\}$。将 P 分为 $s=\left[\frac{M}{3}\right]$ 组，即 $P_i=\{a_{3(i-1)+1},a_{3(i-1)+2},a_{3(i-1)+3}\}(i=1,2,\cdots,s)$；$Q$ 也分为 s 组，每组二进制序列的个数 $k=\left[\frac{N}{s}\right]$，且 $k\geqslant 4$，每组最前的第一位为标志位，后三位则作为嵌入位，即 $Q'_j=\{b_{k(j-1)+1},b_{k(j-1)+2},b_{k(j-1)+3},b_{k(j-1)+4}\}(j=1,2,\cdots,s)$，通过分组确定匹配关系。

2.3 嵌入过程

利用2.2中的分组方法，选取嵌入的位置 Q'_j，加密信息序列 P_i 要嵌入到载体图象中，首先与 Q'_i 的嵌入位进行匹配，如果匹配个数大于等于2，则 P_i 不需取反直接嵌入到 Q'_i 中，即修改 Q'_i 中与 P_i 不匹配的位，标志位存0；如果匹配个数小于2，将 Q'_i 中与 $\bar{P}_i$ 不匹配的位修改，并将标志位设为取反标志1，与此同时记录对标志位修改个数 t。如图3所示载体信息为0101……0100……，加密信息为111010……，匹配后根据上述原则进行嵌入操作。

第一组载体信息 Q'_1 中的第一个位置是标志位，111与101进行匹配，显然根据匹配个数为2可知加密信息不需要取反，标志位设为0，即原载体图象中的0不需修改。而数据101则应修改为111，即将不匹配位0替换为1。下一组 P_2 与 Q'_2 比较后易知应取反，标志位存1，取反后加密信息变成101，再与载体信息100比较，则需将最后的0替换为1，并且用 t 表示对标志位的修改个数，此时 $t=t+1$。同理对后面的信息如上述方法一样进行处理。将全部的加密信息嵌入完毕之后，统计对标志位的修改数 t，如果 t 大于标志位总数的一半，令1表示不取反、0表示取反，这样使得对标志位的修改数减少为 $1-t$，保证了修改位不大于总分组数的一半。

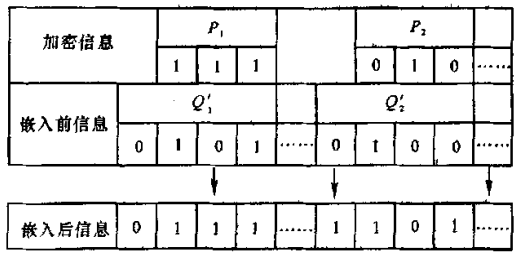

图3 嵌入加密信息

算法描述如下。

步骤1：$t=0$，标志1表示取反，0表示不取反；

步骤2：将加密信息转化为二进制序列，并将每组三位进行分组为 $P_i(i=1,2,\cdots,s)$，共有 s 组。读取载体图象的低位，分为 s 组每组 k 个二进制数，取前四位组成 $Q'_i(i=1,2,\cdots,s)$；

步骤3：对 P_i 与 Q'_i 进行匹配，如果匹配数 $\mathrm{num}(i)\geqslant 2$，标志位为0，对 Q'_i 中与 P_i 不匹配的位进行修改；否则标志位为1，对 Q'_i 中与 $\bar{P}_i$ 不匹配的位进行修改。若标志位被修改，则 $t=t+1$；

步骤4：如果 $t>\frac{s}{2}$，用标志1表示不取反，0表示取反，重新替换标志位 $b_{k(i-1)+1}(i=1,2,\cdots,s)$ 的值；

步骤5：保存得到加密后的图象。

提取加密信息是嵌入信息的逆过程，首先找出嵌入的位置，并排成一个序列；然后分组读取加密信息，取标志位的值，以此判断出加密信息是否取反，当计算出所有加密信息后，便可重构出原加密信息。

3 实例分析

采用大小为512×512的载体图象，256×256的lena图象作为加密信息来进行实验，结果如图4所示(图片按比例进行了缩放)，其中图4(c)是只在一个字节中嵌入一位加密信息的实验结果。从隐藏后的图象可以看出，嵌入加密信息后对载体图象的影响较小。

当加密信息大小为 m 时，运用该算法将信息隐藏到载体图象中，则对载体图象的修改小于等于 $\frac{m}{3}$，对标志位的最大修改可能是 $\frac{m}{6}$，即总的修改位不超过 $\frac{m}{2}$，使得载体图象嵌入加密信息后效果较好。图4(b)中加密信息总长为65536位，隐藏到载体图象

中总共修改了 25282 位,即修改了 38.577%,修改量较小。通过大量实验,加密信息与选择为嵌入位置的载体图象低位的匹配成功率一般都能达到 60%以上。加密信息恢复时,是一种无损还原,且时间复杂度较低。

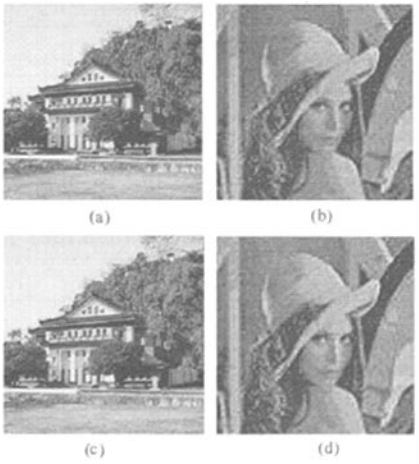

图 4 算法图

(a)载体图象;(b)加密图象;(c)隐藏结果;(d)无损还原。

4 结束语

本文通过分组,让加密信息与载体信息进行匹配,使得在嵌入量不变的情况下,对嵌入信息位的修改大大减少,提高了 LSB 算法的安全性,并且算法实现简单,嵌入与提取过程都不需要经过复杂的运算,速度很快,具有较好隐藏效果。LSB 算法是基于空域的隐藏算法,采用直接改变图像元素值,有一定的局限性,抗干扰能力较低,但是 LSB 算法的嵌入量是非常可观的。

参考文献:

[1] DAVID A UCSMITH Ed. Information hiding: proceedings of the second international workshop. Lecture Notes in Computer Science 1525[C]. Berlin: Springer-V erlag,1998.

[2] MITCHELL D SWANSON,MEI KOBAYASHI. Multimedia data embedding and watermarking technologies[J]. Proceedings of the IEEE,1998,86(6): 1064-1087.

[3] 靳战鹏,沈绪榜. 基于位平面的 LSB 图像隐藏算法分析及改进[J]. 计算机应用,2005,25(11):2541-2543.

[4] 尹德辉,李炳法,唐燕. 基于图像位平面级的信息隐藏算法的研究[J]. 四川大学学报:自然科学版,2006,43(6):1215-1219.

[5] 刘年生,郭东辉. 基于混沌加密的一种图像信息隐藏传送方法[J]. 计算机工程,2006,32(7):135-137.

[6] 姜吉涛,周雪芹,刘晓红. 一种基于 LSB 的数字图像隐藏的改进算法[J]. 山东理工大学学报:自然科学版,2007,20(3):66-68.

[7] 曾浩,范明钰,王光卫. 一种基于 RGB 分量排序的索引替代信息隐藏算法[J]. 计算机应用研究,2007,24(5):312-320.

[8] 刘建东,陈桂强,余有明,等. 基于视觉特性及低位替换优化的信息隐藏方法[J]. 计算机工程,2007,33(8):157-159.

[9] 罗向阳,陆佩忠,刘粉林. 一类可抵御 SPA 分析的动态补偿 LSB 信息隐藏方法[J]. 计算机学报,2007,30(3):463-467.

(责任编辑:尹 闯)

阿米洛利可治疗多发性硬化症

多发性硬化症是一种免疫系统错误攻击自身机体的自体免疫疾病,其发病的部分原因是神经细胞内积聚过多的钙,导致保护神经的髓鞘受到破坏。英国牛津大学研究人员对实验鼠的研究发现,阿米洛利可以阻止神经细胞内高剂量钙的积聚,从而防止神经组织恶化。该项研究负责人拉尔斯·富杰尔表示,如果临床试验证实阿米洛利对人体有效,多发性硬化症患者的治疗进程将加快。研究人员目前正在研究阿米洛利用于人体的合适剂量,并计划于 2008 年开始临床试验。因此,多年来用于治疗高血压的药物阿米洛利有望成为治疗多发性硬化症的关键药物。

(据科学网)