基于 N 阶 Bézier 曲线的多信息隐藏算法研究

周化灵，陈春玲
（南京邮电大学 计算机学院，江苏 南京 210003）

摘 要：由于信息安全涉及国家安全和经济利益等多个方面，因此有必要将秘密的信息隐藏在某些以明文传输的格式化数据中。文中简要地介绍了这个领域的总的概况，以及其中的一些算法；然后着重介绍了一种隐藏比比较高的算法：基于一阶 Bézier 曲线的信息隐藏算法，在分析该算法原理的基础上，对该算法做了进一步的推广，提出了 N 阶 Bézier 曲线的信息隐藏算法的相关定义。根据该算法实现了对多幅图像的隐藏和恢复，从而进一步提高了信息的安全性和隐藏比。

关键词：Bézier 曲线；信息隐藏；融合
中图分类号：TP301.6 **文献标识码**：A **文章编号**：1673－629X(2006)05－0085－03

Research on Multi－Information Concealment Algorithm Based on N－Bézier Curve

ZHOU Hua-ling, CHEN Chun-ling
(Computer Institute, Nanjing University of Posts and Telecommunications, Nanjing 210003, China)

Abstract: It is necessary to hide the secret information in some data transmitted obviously since information security is involved in national security, economy benefits and many other aspects. This paper concisely introduces the general situation in this field and some algorithm of the field, then emphasizes one algorithm which owns high hiding－ratio: concealment algorithm based on one－Bézier curve. On the basis of analyzing the principle of this algorithm, extends it and proposes some relative concepts about N－Bézier curve concealment algorithm. Based on implementation of hiding and resuming some images according to the algorithm, the security and hiding－ratio are further improved.

Key words: Bézier curve; information concealment; amalgamation

0 引 言

信息安全[1]是对信息、系统以及使用、存储和传输信息的硬件的保护。信息隐藏是信息安全领域中的一个重要分支，是按照预先给定的算法将秘密信息隐藏进入另一非机密的文件内容之中（称为宿主信息）。当今的时代是数字化和信息化的时代，信息隐藏技术将是未来信息对抗的焦点，是获取和破解信息的制高点，关系到上至国家安全下至商业机密乃至个人隐私等诸多方方面面，因此备受人们关注。

文中通过对图像间图像隐藏和恢复的分析，推广一阶 Bézier 曲线算法，引出 N 阶 Bézier 曲线算法作为信息融合的中心算法，在理论上推出了融合算法和恢复算法，并给出了算法的适用范围。

1 算法实现

Bézier 曲线是法国工程师 Pierre Bézier 在 20 世纪 70 年代设计雷诺汽车车身图案时提出的[2]，定义如下：

设在三维空间里假定有 $n+1$ 个控制点 $P_k, k=0,1,\cdots,n$，则 Bézier 曲线函数如下：

$$P(t)=\sum_{i=0}^{n}B_i^n(t)P_i, P_i\in R^3,$$
$$B_i^n(t)=C_n^i t^i(1-t)^{n-i}, 0\leqslant t\leqslant 1 \quad (1)$$

假如选取 $N+1$ 个控制点 $P_0,P_1,\cdots,P_N$，利用 N 阶 Bézier 曲线公式可以得到这 $N+1$ 个控制点的融合点 P，这个点 P 就是所需要的新的宿主信息。选取适当的参数，可以使其中一个点为显式点而其它 N 个点为隐藏点，但显式点不是可以任意指定的，这点下文中会讨论。

1.1 一阶 Bézier 曲线

1.1.1 对图像的融合算法

当式(1)中的 $n=1$ 时，式(1)变为，

$$P(t)=(1-t)P_0+tP_1 \quad (2)$$

其中 t 为融合参数，这时就变成一阶 Bézier 曲线。基于一阶 Bézier 曲线的信息隐藏算法，许多文献中都有描述（例如文献[3,4]），下面简单介绍一下算法思想。

设有两幅尺寸相同数字图像 $P_0(x,y)$，$P_1(x,y)$，在这两幅图中任意取出同位置的两像素 $P_0(x_0,y_0)$，$P_1(x_0,y_0)$，应用一阶 Bézier 曲线公式，可以计算出经过融

收稿日期：2005－08－25
作者简介：周化灵（1980－），男，安徽灵璧人，硕士研究生，研究方向为软件技术及在通信中的应用；陈春玲，副教授，研究方向为软件技术及在通信中的应用。

合后的像素 P:$P = (1-t) * P_0(x_0,y_0) + t * P_1(x_0,y_0)$,$t \in [0,1]$。融合系数 t 的取值决定了 $P_0(x,y)$,$P_1(x,y)$ 的融合效果,图 1 所示的五幅图像分别为原始图像和融合系数 t 分别为 0.032,0.5,0.973 时的融合图像。

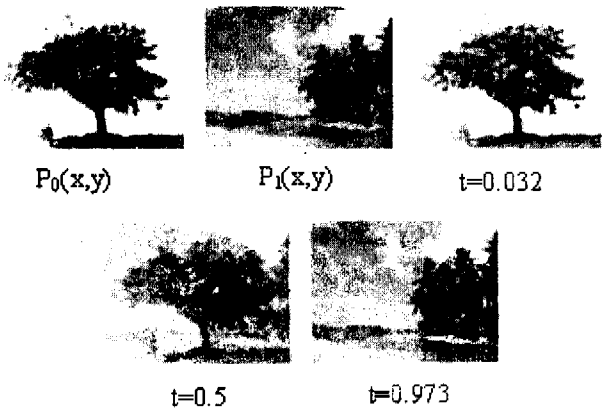

图 1 两幅原始图像和取参不同的三幅一阶 Bézier 曲线算法的融合图像

从上面几幅图可以看出:当 $t = 0.032$ 时,$F_0(x,y)$ 充当信息载体,$F_1(x,y)$ 隐藏其中;当 $t = 0.973$ 时,$F_0(x,y)$ 隐藏于 $F_1(x,y)$ 中。从画面上看含有融合信息的图像与原始图像没有什么不同。从融合算法可以根据融合后图像 $P_2(x,y)$ 和 $P_0(x,y)$,$P_1(x,y)$ 中任一幅图像就可计算出另一幅隐含的图像,像素恢复算法为:

$$rf_0 = \text{round}((P - t * P_1)/(1-t)) \quad (3)$$

$$rf_1 = \text{round}((P - (1-t) * P_0)/t) \quad (4)$$

round(x) 函数是取距结果最近的整数,与通常的四舍五入算法比较类似,融合系数 t 的取值不同,融合的效果不同。

1.1.2 对文件的融合算法

上文介绍的是将两幅等大 BMP 图像进行融合,如果想将任何形式的数据文件进行隐藏,可将该数据文件内容人为拼成一幅位图后再按式(1)操作即可。

对秘密文件信息按次序组成一长串比特流,然后将其划分为长度 $L(0 < L < 8)$ 的子串,按照从上到下,从左到右的顺序,将这些比特子串组成大小和宿主图像文件相等的"图像"。不足的位以补"0"计算。然后将这些比特子串转化为十进制序列(取值范围为 $[0,2^L - 1]$),对该序列和宿主文件信息应用曲线进行融合即可完成对秘密文件的隐藏,信息隐藏比为 $L/8$。

对于解码的过程,首先按式(4)得到人为构造的秘密图像序列 rf_1,对其进行纠错处理,并按下式修正。

$$rf_1 = \begin{cases} rf_1, & rf_1 \bmod 2^{8-L} \leqslant 2^{8-L-1} - 1 \\ rf_1 + 2^{8-L-1} - 1, & \text{otherwise} \end{cases} \quad (5)$$

从经过修正的 rf_1 中提取其高 L 位,即为恢复出的秘密信息比特子串前 L 比特。对宿主信息重复执行上述过程就可以提取出全部的以比特流形式的明文信息;然后只需按 8 比特/字节对上述比特流重新组帧即可将秘密信息完全恢复出来。

由算法引入的误差为[5]:

$$\begin{aligned} e &= rf_1 - f_1 = \text{round}((f_2 - \text{round}((1-t) \times f_0))/t) - f_1 \\ &= \text{round}(\text{round}(f_1 \times t)/t) - f_1 \\ &= \text{round}(f_1 \times t + e_1)/t - f_1 \\ &= \text{round}(e_1/t) = e_1/t + e_2 \end{aligned}$$

其中 e_1 为 round($f_1 \times t$) 引入的截断误差,e_2 为 e_1/t 引入的截断误差,则 $|e_1| \leqslant 0.5$,$|e_2| \leqslant 0.5$,于是有 $|e| \leqslant 0.5/t + 0.5$,在误差 $0.5/t + 0.5 \leqslant 2^{8-L-1}$ 时可以确保秘密信息无误地恢复出来。

对于原始图像,设其大小为 $M \times N$,根据式(2)可计算出其峰值信噪比 $\text{PSNR} = 10\lg[M \times N \times 255/D] \geqslant 20\lg[1/t]$,$D$ 为两原始图像的各个像素的差值的平方和。根据 PSNR 的经验值,$20\lg[1/t] \geqslant 28$,计算得出 $t \leqslant 0.03981$。再由 $0.5/t + 0.5 \leqslant 2^{8-L-1}$,可解得 $L \leqslant 3.1926$。因为 L 须取整数值,因此当 L 的取值是 1,2,3 时可以确保秘密信息完整无误地恢复出来。为了尽可能扩大宿主信息的隐藏能力,L 取 3,即每字节(8 比特)的宿主信息可用来隐藏 3 个比特的秘密信息,也就是说隐藏比可高达为 3/8[4]。$L = 3$ 时,t 的取值范围为 $[0.0345,0.0398]$。

2 N 阶 Bézier 曲线

一阶 Bézier 曲线算法只能把一幅图像融合在另一幅图像中或者把一个文件隐藏在另一个文件中,但是如果希望把多幅数字图像或文本信息隐藏起来,一阶 Bézier 曲线算法有明显的不足,因此下文引入 N 阶 Bézier 曲线的算法。

2.1 N 阶 Bézier 曲线的几个定义

定义 1 隐藏点:把经过 N 阶 Bézier 曲线算法后被隐藏在宿主信息里的点称为隐藏点。

定义 2 显式点:在 N 阶 Bézier 曲线算法中,当 $t \to 0$ 的 P_0 或者 P_1 时称为显式点。

从 Bézier 曲线的表达式可以看出,一个像素点并非一定是隐藏点或者显式点,当 $t \in (0.5 - \theta, 0.5 + \theta)$,$P_i C_n^i t^i (1-t)^{n-i} \to 0$,$i \neq 0$ 时,任何一个 P_i 都不是显式点,因为没有哪个点的系数趋向 1。可以说这时的宿主像素点是模糊的,宿主图像是 $N+1$ 幅图像按权值叠加。当 $t \to 0$ 时,这时只有 P_0 为显式点,同理当 $t \to 0$ 时,$P_i C_n^i t^i (1-t)^{n-i} \to 0$,$i \neq n$,这时只有 P_1 为显式点。所以 $N+1$ 幅数字图像中,只有 P_0 或 P_N 两幅图像可以是显式图像。

2.2 N 阶 Bézier 曲线融合算法

当式(1)中的 $n = N$ 时,式(1)变为:

$$P(t) = \sum_{i=0}^{N} P_i C_N^i t^i (1-t)^{N-i}, P_i \in R^3, 0 \leqslant t \leqslant 1 \quad (6)$$

仿照一阶算法,设有 $N+1$ 幅尺寸相同数字图像 P_0,$P_1,\cdots,P_N$,这两幅图中任意取出同位置的 $N+1$ 个像素 $p_1,p_2,\cdots,p_N$,应用 N 阶 Bézier 曲线公式,可以计算出经

过融合后的像素 p:$p=\sum_{i=0}^{N}p_iN!/[i!(N-i)!]t^i(1-t)^{N-i}$,$p_i\in R^3$,$0\leqslant t\leqslant 1$。同样,融合系数 t 的取值决定了这 $N+1$ 幅图像的融合效果。

图2所示的六幅图像分别为 $N=2$ 时的三幅原始图像 $P_0(x,y)$,$P_1(x,y)$,$P_2(x,y)$ 以及融合系数 t 分别为0.032,0.5,0.973时的融合图像。

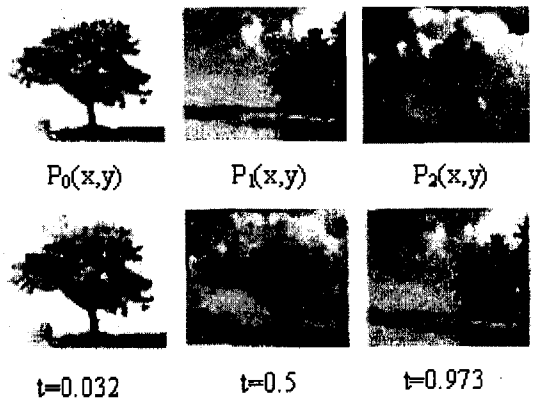

图2　三幅原始图像和取参不同的三幅2阶 Bézier 曲线算法的融合图像

根据融合后图像 p 和 $p_0,p_1,\cdots,p_N$ 中任 N 幅图像就可计算出另一幅隐含的图像,像素恢复算法为:

$$rf_k\text{round}((p-\sum_{i=0,i\neq k}^{N}P_iN!/[i!\quad(N-i)!]t^i(1-t)^{N-i})/(N!/[k!(n-k)!](1-t)^{N-k})) \qquad (7)$$

在实现 N 阶 Bézier 曲线融合算法的时候,可以按照下面的步骤实现 $N+1$ 幅图像(数据文件同理)的融合:

1)在 $N+1$ 幅图像集合中先选出要融合的两幅图像 P_i,P_j,根据需要的隐藏效果选取合适的融合参数 t_1,按照一阶 Bézier 曲线融合算法进行图像融合,融合后的图像为 P。

2)在上述图像集合中删除 P_i 和 P_j,把融合后的图像 P 加入到集合中作为待融合图像。

3)重复上述步骤直到得到最后所需要的宿主图像。

由式(3)和式(4)可知,在计算融合后的像素时 round()函数会引入取整误差。经过分析知道 N 阶 Bézier 曲线融合算法所产生的误差是一阶算法的 N 倍。对于图像,当 N 小于一个临界值时,这些引入的误差并不会影响画质,但是当 N 超过临界值时,将会影响画面质量,同时也会给原始图像的恢复带来困难。

3　结束语

一阶 Bézier 曲线融合算法实现起来较为简单,从可行性、可靠性和高信息隐藏比等几个关键指标来看,该算法做为信息隐藏的一种算法是可选的。但是由于该算法的简单性,所以破解者很容易破解出原始信息,而且该算法的信息隐藏比还可以进一步提高,所以文中提出 N 阶 Bézier 曲线融合加密算法。

信息隐藏是信息安全领域中的热门话题,且研究面相当广泛,像 LSB 算法,还有变换域的算法如 Arnold 变换、Hilbert 曲线变换、Fibonacci 变换等都已得到了广泛应用,而关于 Bézier 曲线的信息隐藏的研究却相对匮乏,由于它的诸多优势,相信它在信息的秘密传输领域会逐渐成为一个研究的热点。

参考文献:

[1] Whitman M E, Mattord H J. 信息安全原理[M]. 北京:清华大学出版社,2004.
[2] Hearn D. 计算机图形学[M]. 北京:电子工业出版社,2002.
[3] Fabien A, Petitcolas P, Anderson R J, et al. Information hiding - a survey[J]. Proc. of IEEE, 1999, 87(7): 1062 - 1078.
[4] 柳葆芳,平西建,邓宇虹. 基于融合的数据隐藏算法[J]. 电子学报,2001,29(11):1445 - 1448.
[5] 郎　锐. 基于一阶 Bézier 曲线的信息隐藏算法[J]. 电脑编程技巧与维护,2004(8):86 - 89.

(上接第3页)

行降维之后,再利用人工神经网络建立校正模型,可以比较准确地拟合出光谱数据与3种化学成分含量之间的关系,其平均拟合精度可以达到0.1%。

3　结　论

文中将主成分分析和 BP 神经网络方法相结合,对一组烟草样本光谱数据进行了拟合。实验结果表明,该方法对于某些烟草成分含量的近红外光谱数据分析能够达到比较高的精度,这在烟草行业具有实际应用价值。与多元线性回归、偏最小二乘等方法相比,神经网络方法的优势在于具有强大的非线性关联能力,可以有效解决光谱数据分析中的非线性校正问题。主成分分析作为一种线性降维方法,可以很好地解决光谱数据从高维空间向低维空间的映射问题,有效解决了神经网络方法在计算量大方面存在的限制。这二者的结合使用,对于近红外光谱分析技术的进一步应用和发展,具有重要的现实意义。

参考文献:

[1] 陆婉珍. 现代近红外光谱分析技术[M]. 北京:中国石化出版社,2000.
[2] 杨荣英. BP 神经网络主成分分析法在交通预测中的应用[J]. 武汉理工大学学报,2002(3):100 - 102.
[3] 王　旭. 人工神经元网络原理与应用[M]. 沈阳:东北大学出版社,2000.
[4] 魏广芬. 基于主成分分析和 BP 神经网络的气体识别方法研究[J]. 传感技术学报,2001(4):41 - 47.
[5] 成卫青. 应用主成分回归分析评估企业经济效益[J]. 微机发展,2003,13(7):29 - 31.