

文章编号:1001-9081(2010)01-0270-04

## 基于非线性变换法语谱图的伪彩色显示

张 静,余小清,许雪琼,常辽豫,万旺根,李昌莲

(上海大学 通信与信息工程学院,上海 200072)

(yxq@staff.shu.edu.cn)

**摘 要:**鉴于人眼对彩色的分辨率远远高于灰度分辨率的特点,提出了一种利用 C# 程序实现伪彩色显示的新方法,该方法通过非线性映射变换函数,建立频谱幅度与颜色的对应关系,从而实现伪彩色显示,并将其用于语谱图显示及语音信号处理界面设计中;此外,还将语音播放功能加入语谱图显示过程中,实现了“可视”语言的播放。该方法编程简单,界面设计灵活,对设备的要求较低,在普通 PC 机上即可实现。

**关键词:**语谱图;伪彩色;灰度;调色板;映射变换

**中图分类号:** TP274 **文献标志码:** A

### Speech spectrogram display in pseudo-color based on non-linear transform method

ZHANG Jing, YU Xiao-qing, XU Xue-qiong, CHANG Liao-yu, WAN Wang-gen, LI Chang-lian

(School of Communication and Information Engineering, Shanghai University, Shanghai 200072, China)

**Abstract:** Concerning that the human eye to color resolution is much higher than gray-scale difference, this paper proposed a new method to show the spectrogram by C# procedure, which focused on the mapping transformation function and built up the relationship between spectral magnitude and color, to make pseudo-color display come true, and also used it for speech spectrogram display and audio signal processing interface design. Meanwhile in order to achieve better analysis and learning effects, the authors added audio play function so as to watch the spectrogram while listening to the corresponding audio. This method, with simple programming, flexible interface design and less demand on the equipment, can be realized on ordinary PC.

**Key words:** speech spectrogram; pseudo-color; gray-scale; palette; mapping transformation

## 0 引言

语音的时域分析及频域分析是语音分析的两种主要方法,而这两种单独分析的方法均有其局限性:在语音的时域分析中,对语音信号的频率特征没有直观的了解;而频域分析出的特征又没有语音信号随时间的变化关系<sup>[1]</sup>。因此从事语音研究的人们致力于研究语音的时间依赖于傅里叶分析的方法。这种时间依赖于傅里叶分析的显示图形称做语谱图,早在 1941 年由贝尔实验室的研究人员发明,得到语谱的灰度图像。但是这种灰度图像的分辨率和视觉效果不尽如人意,随着技术的发展,人们尝试语谱图的彩色显示的实现。本文利用 c# 编制程序,设计界面,采用非线性映射变换函数实现语谱图的伪彩色显示,同时为了达到更好的分析、学习效果,不仅设置了对照实验组,还添加了语音播放功能,可以做到边观测语谱图边听相应的语音。该方法简单,对设备要求低,显示效果良好。语谱图综合了频谱图与时域波形的优点,充分展示了语音频谱随时间的变化情况,除此之外,语谱图还呈现出丰富的纹理特征,因此在聋人语训、声纹鉴定、车型识别、数字音频水印等方面具有广泛的研究意义。

## 1 语谱图产生原理

由于语音信号是一种短时平稳信号,可以在每个时刻用其附近的短时段语音信号分析得到一种频谱,将语音信号连续地进行这种频谱分析,可以得到一种二维图谱,这种二维图谱可以表达三维信息,它的横坐标表示时间,纵坐标表示频率,每个像素的灰度值大小反映相应时刻和相应频率的能量密度,这种时频图称为语谱图。语谱图依据时间分辨率及频谱分辨率可以分为窄带语谱图和宽带语谱图两种,窄带语谱图用于获得较高的频率分辨率,而宽带语谱图用于获得较高的时间分辨率<sup>[2]</sup>。语谱图的产生基于短时傅里叶变换<sup>[3]</sup>,定义如下:

$$X_n(e^{j\omega}) = \sum_{m=-\infty}^{\infty} x(m)w(n-m)e^{-j\omega m} \quad (1)$$

其中: $n$  为时域采样点序号; $w(n)$  是窗函数,在求短时频谱时通常采用的是汉明窗;一方面是因为汉明窗在频率范围中的分辨率高;另一方面它的旁瓣衰减大,具有频谱泄露少的优点。

类似地还可以定义离散的短时傅里叶变换,令  $\omega = 2\pi k/N (0 \leq k \leq N-1)$ , 有:

**收稿日期:** 2009-07-07; **修回日期:** 2009-08-19。 **基金项目:** 国家自然科学基金资助项目(60872115);上海市科委国际合作基金资助项目(075107035);上海市教委电路与系统重点学科资助项目(J50104);上海市重点学科和科委重点实验室资助项目(S30108)。

**作者简介:** 张静(1984-),女,山东聊城人,硕士研究生,主要研究方向:语音信号处理;余小清(1958-),女,上海人,副教授,主要研究方向:语音信号处理、音频信息检索;许雪琼(1984-),女,福建泉州人,硕士研究生,主要研究方向:音频分类检索;常辽豫(1982-),男,河南漯河人,硕士研究生,主要研究方向:压缩域音频的分类检索;万旺根(1961-),男,上海人,教授,主要研究方向:数字媒体、虚拟现实、音频分类检索;李昌莲(1985-),女,江苏南通人,硕士研究生,主要研究方向:音频场景分析。

$$X_n(k) = X_n(e^{2\pi kj/N}) = \sum_{m=-\infty}^{\infty} x(m)w(n-m)e^{-2\pi kj/N} \quad (2)$$

由功率谱函数(能量密度谱函数)的定义,短时功率谱和短时傅里叶变换有如下关系:

$$S_n(e^{j\omega}) = X_n(e^{j\omega})X_n^*(e^{j\omega}) = |X_n(e^{j\omega})|^2 \quad (3)$$

或:

$$S_n(k) = X_n(k)X_n^*(k) = |X_n(k)|^2 \quad (4)$$

不难证明  $S_n(e^{j\omega})$  是  $x(n)$  的短时自相关函数的傅里叶变换:

$$S_n(e^{j\omega}) = \sum_{k=-\infty}^{\infty} R_n(k)e^{j\omega k} \quad (5)$$

其中  $R_n(k) = \sum_{m=-\infty}^{\infty} x(m)w(n-m)x(m+k)w(n-m-k)$ 。

用时间  $n$  作为横坐标,  $k$  作纵坐标, 将  $S_n(k)$  的值表示为灰度级所构成的二维图像就是语谱图, 通过变换  $10 \lg(S_n(k))$ , 能得到语谱图的 dB 表示<sup>[7]</sup>。

## 2 伪彩色映射变换方法

伪彩色处理是将灰度图像线性或非线性映射到一定的调色板上, 以达到彩色显示效果的处理方式。由于人眼对彩色的分辨率远远高于对灰度差的分辨率, 所以可用于识别灰度差较小的像素。通过伪彩色处理, 可以得到具有较高分辨率和较好视觉效果图像。图像伪彩色处理的关键在于调色板的编码方法, 在 RGB 空间常用两种方法进行调色板的编码, 即: 灰度分割法和灰度级-彩色变换法。灰度分割法形成的调色板色彩给人的感觉不是连续的, 因此在 RGB 空间一般都采用灰度级-彩色变换法进行调色板编码, 即: 对灰度图像中的每一个像素点, 取得该点的灰度值, 通过红、绿、蓝变换函数, 产生 RGB 色彩空间的三个分量, 从而只要保证变换函数连续, 则调色板彩色编码就是连续的<sup>[8]</sup>。因此灰度级-彩色变换法的关键在于红、绿、蓝变换函数的构造。

### 2.1 传统的伪彩色映射变换

传统的映射变换函数<sup>[9]</sup>如图1所示。

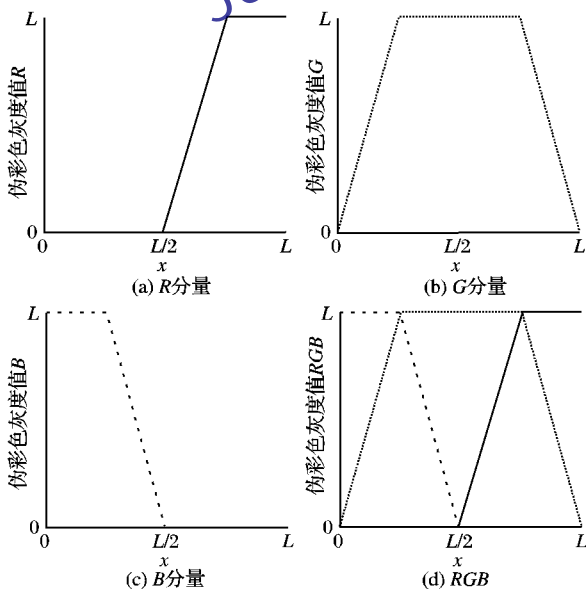


图1 典型的线性映射变换函数曲线

图1中  $R$ 、 $G$ 、 $B$  分别表示伪彩色灰度值,  $x$  表示原始灰度图像的灰度值, 前三幅图像分别为红、绿、蓝三种变换函数, 最后一幅是把三种变换画在一张图上, 这种映射关系可以使不

同的灰度对应不同色彩, 呈现一一映射关系。灰度最大值  $L=255$ , 将低于  $L/4$  的所有灰度映射成饱和度线性增加的绿色, 将  $L/4$  到  $L/2$  之间的所有灰度映射成饱和度线性减少的蓝色, 在  $L/2$  到  $3L/4$  之间的灰度映射成饱和度线性增加的红色, 在  $3L/4$  到  $L$  之间的灰度映射成饱和度减少的绿色, 并且只有在  $0$ 、 $L/2$ 、 $L$  处出现纯基色(分别表示  $B$ 、 $G$ 、 $R$ ), 其他灰度将合成多种不同的颜色; 但是这种线性变换存在一定得局限性<sup>[5]</sup>, 表现为: 1) 低阶灰度转换的绿色与高阶灰度转换的绿色容易混淆; 2) 由于在  $0$ 、 $L/2$ 、 $L$  处存在纯基色, 其周围一定范围内的彩色太暗, 从而无法区分开。

### 2.2 非线性伪彩色映射变换

鉴于传统方法的局限性, 利用函数关系式(6)、式(7)和式(8), 本文提出一种非一一映射的方法。

$$R(x) = \begin{cases} 0, & x \leq l \\ 0, & l < x \leq 2l \\ (x/l) - 2, & 2l < x \leq 3l \\ 4 - (x/l), & x > 3l \end{cases} \quad (6)$$

$$G(x) = \begin{cases} 0, & x \leq l \\ (x/l) - 1, & l < x \leq 2l \\ 3 - (x/l), & 2l < x \leq 3l \\ 0, & x > 3l \end{cases} \quad (7)$$

$$B(x) = \begin{cases} x/l, & x \leq l \\ 2 - (x/l), & l < x \leq 2l \\ 3 - (x/l), & 2l < x \leq 3l \\ (x/l) - 3, & x > 3l \end{cases} \quad (8)$$

式中  $l = L/4$  ( $L = 255$ ), 可以得到归一化的非线性映射变换函数曲线, 如图2所示。

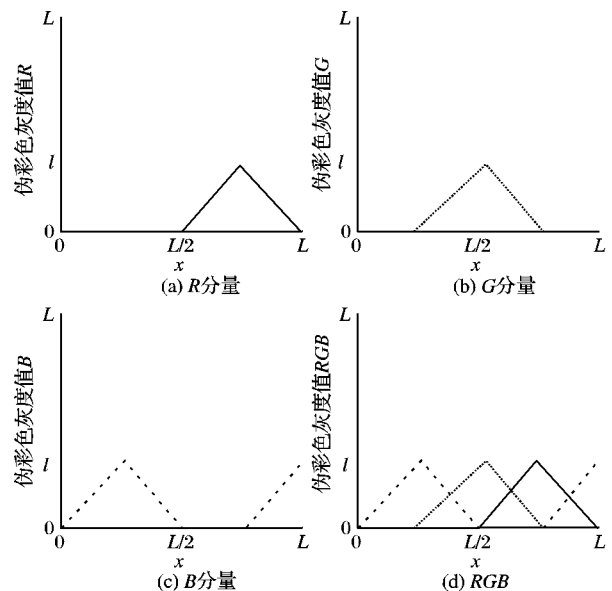


图2 归一化的非线性映射变换函数曲线

图2中  $L = 255$ ,  $x$  为初始灰度值,  $R$ 、 $G$ 、 $B$  分别表示变换后归一化的伪彩色灰度值。为了达到更好的显示效果, 通过函数关系式(9), 可以使 RGB 三分量的伪彩色灰度值分布在  $0 \sim 255$  灰度范围内。

$$\begin{cases} R' = \sqrt{R} \times 255 \\ G' = \sqrt{G} \times 255 \\ B' = \sqrt{B} \times 255 \end{cases} \quad (9)$$

式中  $R$ 、 $G$ 、 $B$  分别为图2中归一化伪彩色灰度值,  $R'$ 、 $G'$ 、 $B'$  为

最终所得的伪彩色灰度值。将 256 个灰度级等分成四个子区间,在这些子区间内映射变换后的  $R'$ 、 $G'$ 、 $B'$  分别由最小值线性变换为最大值,再变换为最小值,此过程如图 3 所示。

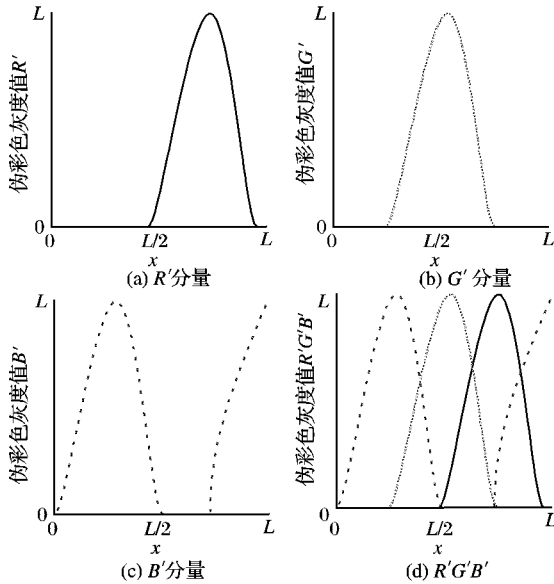


图3 由灰度到彩色的非线性映射变换函数曲线

在图3的变换过程中,不同的灰度值可能会映射为同种颜色,但它使得图像的对比度大大增强,从而使得原灰度图像内部被掩盖的细节信息变得清晰可见<sup>[6]</sup>。从图中也可反映出颜色的渐变过程,灰度值为0时对应黑色(能量最低),继而随着灰度值的增大依次由蓝色( $L/4$ )渐变为绿色( $L/2$ ),再渐变为红色( $3L/4$ ),按照原理最后可能会呈现紫色(即:红色与蓝色的混合色),与后面语谱图显示的实验结果相对应。由于语音能量主要集中在低频区域,随着频率的升高,能量值减小,最后趋于零,反映在语谱图中,频率轴自上而下灰度值逐渐增大,颜色的渐变过程与图3一致。

### 3 C#中语谱图的伪彩色显示

伪彩色语谱图的形成总体上可以分为两部分:像素矩阵  $\text{PixelMatrix}[,]$  的形成和伪彩色处理,下面逐一进行说明。

#### 3.1 像素矩阵的形成

像素矩阵  $\text{PixelMatrix}[,]$  的形成,流程如图4所示。

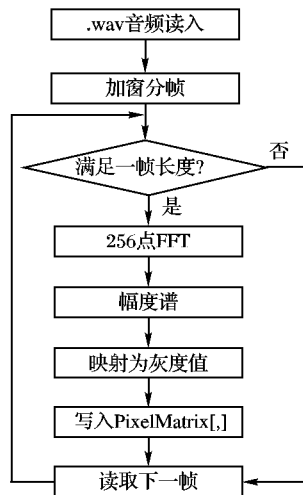


图4 像素矩阵  $\text{PixelMatrix}[,]$  形成

图4中取每帧.wav音频数据进行256点FFT变换,将所得的频谱幅度值归一化后可映射为像素灰度值,这一过程用

到的映射变换函数如式(10)所示:

$$\text{retVal} = \frac{20\lg(\text{Mag}) + 100}{100} \times 255 \quad (10)$$

其中:Mag表示频谱幅度;retVal是由幅度值映射得到的像素灰度值。最后将每帧得到的灰度值按列写入  $\text{PixelMatrix}[,]$ ,即得像素矩阵。

#### 3.2 伪彩色处理

要显示的语谱图非常精细复杂,若将其直接绘制在屏幕上,需要很长时间,在此利用c#中将图像用于双倍缓冲的图形编程技巧<sup>[9]</sup>,先把图形绘制在一块虚拟的“画布”上,然后通过PictureBox再将图像呈现在屏幕上。流程如图5所示。

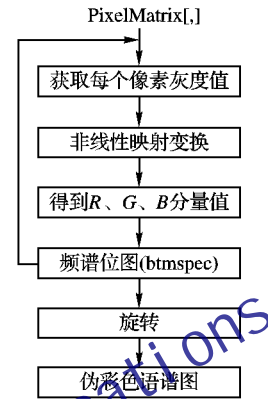


图5 伪彩色处理

首先获取每个像素的灰度值,即:  $\text{PixelMatrix}[,]$  每个元素值,通过图3所示的非线性映射变换函数得到  $R$ 、 $G$ 、 $B$  分量,从而为每个像素着色,然后将其写入创建好的频谱位图中,因图像的二维方向跟数组的二维方向是转置关系,要将所得的频谱位图进行 RotateNoneFlipY 方式的变换;最后通过 PictureBox 将频谱位图呈现在屏幕上,即可得到伪彩色语谱图。

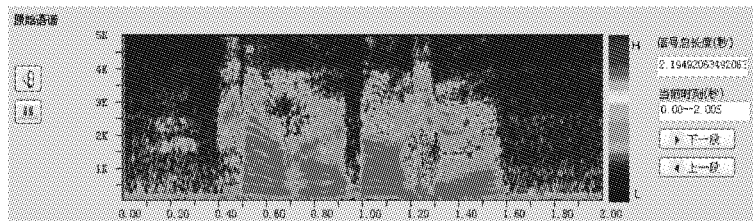
### 4 语谱图显示结果分析

与利用传统的映射变换函数(图1所示)相比,本文非线性映射变换方法得到的语谱图可以更精确地展现所选时间段的音频信号的频谱特征,下面以长度约为2.19s的女声“上海大学”为例加以说明,实验取帧长为256点,帧移为128点,每次显示的时间为2s,语音信号的采样频率为11.025kHz,图6为传统方法与本文方法所得语谱图对比,其中(a)是利用传统方法得到的伪彩色语谱图,(b)由本文非线性映射变换方法得到。

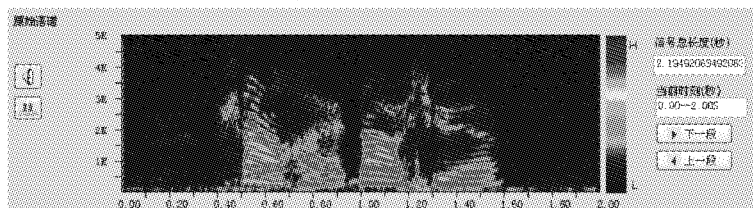
图6(a)利用传统线性变换方法,会出现灰度值较小的像素被灰度值较大的像素所掩盖以及不同灰度值映射成的同种颜色连为一体等现象,致使所得语谱图纹理细节特征模糊不清,而从图6(b)可以清晰地看出横杠、竖直条、乱纹及颜色不同,能够很好地反映出语谱图所要表达的丰富信息,其中横杠是明显的一条条横方向的条纹,条纹的地方实际是颜色深的点聚集的地方,随时间延续形成条纹,表示语音中频率值为该点横坐标值的能量较强,被称为共振峰,从横条对应的频率和宽度可以确定相应的共振峰频率和带宽,在一个语音段的语谱图中有没有横杠出现是判断它是否是浊音的重要标志,一般的,对于语音识别常取前三个共振峰,语音合成需取五个;竖直条(又叫冲直条),与时间轴垂直,每个竖直条相当于一个基音,条纹的起点相当于声门脉冲的起点,条纹之间的距离表示基音周期,条纹越密表示基音频率越高;乱纹往往表示清音、擦音或送气声;语谱图中颜色的显示与色度条相对应,越接近上端(红色),能量值越高,越接近下端(蓝色),能量值越



低,从语谱图上可以看出能量高的值(如:一条条横方向的条纹)往往出现在低频区域,符合语音信号的特点,高频端大约在 800 Hz 以上按 6 dB/倍频跌落,所以语音信号处理是常常要进行预加重,来提升高频部分,使信号的频谱变得平坦。



(a) 传统映射变换方法



(b) 本文非线性映射变换方法

图6 传统方法与本文方法所得语谱图对比

此外在利用 C#设计语谱图显示系统过程中还加入语音播放功能,以便观察音频信号语谱结构的同时,可以听其声音,感受不同频率结构对发声的影响,从而同时从听觉和视觉两方面来了解不同的音频信号所具有的特点,是音频信号分析的有效工具。

## 5 结语

本文采用灰度级-彩色变换法,通过一种非线性映射变

换函数,建立起频谱幅度与颜色的对应关系,从而实现语谱图的伪彩色显示;同时,为了达到更好的分析学习效果,还添加了语音播放功能,能够做到观测语谱图的同时,播放相应的语音,实现“可视语言”在 PC 机上的快速、精确显示。实验结果表明,与传统线性变换方法相比,本文非线性变换方法得到的语谱图显示效果良好,同时声音播放清楚,为进一步的音频分析做好了铺垫,具有广泛的应用意义。

## 参考文献:

- [1] LU WENKAI, ZHANG QIANG. Deconvolutive short-time Fourier transform spectrogram [J]. Signal Processing Letters, 2009, 16(7): 576-579.
- [2] MORRIS L R. A PC-based digital speech spectrograph [J]. IEEE Micro. 1988, 8(6): 68-85.
- [3] 魏松,李琦,赵仁才.基于短时傅里叶变换语言信号分析算法[J].电子测量技术,2006,29(1): 16-17.
- [4] 韩纪庆,张磊,郑铁然.语音信号处理[M].北京:清华大学出版社,2005.
- [5] 雷小丽,党群.一种新的非线性变换法实现图像增强的方法[J].光子学报,2007,36(6): 346-348.
- [6] 李金秀,李均利.肺癌图像的伪彩色处理[J].宁波大学学报:理工版,2007,20(1): 55-58.
- [7] 李富强,万红,黄俊杰.基于 Matlab 的语谱图显示与分析[J].微计算机信息,2005,10(21): 172-174.
- [8] 刘维,刘纪元,黄海宁,等.声纳图像伪彩色处理的调色板连续色编码方法[J].系统仿真学报,2005,17(11): 1724-1731.
- [9] WATSON K. C#入门经典[M].4版.齐立波,译.北京:清华大学出版社,2008.
- [2] ZHU B B, SWANSON M D, TEWFIK A H. When seeing isn't believing multimedia authentication technologies[J]. IEEE Signal Processing Magazine, 2004, 21(2): 40-49.
- [3] FRIDRICH J, GOLJAN M. Images with self-correcting capabilities [C]// ICIP 99: International Conference on Image Processing. Washington, DC: IEEE, 1999, 3: 792-796.
- [4] 张鸿宾,杨成.图像的自嵌入及篡改的检测和恢复算法[J].电子学报,2004, 32(2): 196-199.
- [5] LIN P L, HSIEH C K, HUANG P W. A hierarchical digital watermarking method for image tamper detection and recovery [J]. Pattern Recognition, 2005, 38 (12): 2519-2529.
- [6] 和红杰,张家树,陈帆.一种高定位精度的可恢复水印算法[J].中国科学: E 辑,2008,38(4): 533-552.
- [7] 刘泉,江雪梅.用于图像篡改定位和恢复的分层半脆弱数字水印算法[J].通信学报,2007, 28(7): 104-110.
- [8] LEE T Y, LIN S D. Dual watermark for image tamper detection and recovery [J]. Pattern Recognition, 2008, 41(11): 3497-3506.
- [9] CHANG C, FAN Y H, TAI W L. Four-scanning attack on hierarchical digital watermarking method for image tamper detection and recovery [J]. Pattern Recognition, 2008, 41(2): 654-661.
- [10] LIU S H, YAO H X, GAO W, et al. An image fragile watermark scheme based on chaotic image pattern and pixel-pairs [J]. Applied Mathematics and Computation, 2007, 185(2): 869-882.
- [11] 和红杰,张家树,田蕾.能区分图像或水印篡改的脆弱水印方案[J].电子学报,2005,33(9): 1557-1561.

(上接第 206 页)

图6中文献[8]和本文的篡改恢复图像的峰值信噪比分别为 15.43 dB 和 33.13 dB。图6(a)不存在非零点(文献[8]没有检测到篡改块,不执行篡改恢复操作),说明文献[8]不能抵抗内容篡改攻击。从图5、6可看出本文算法能有效地抵抗内容篡改攻击,且具有理想的篡改检测和篡改恢复效果。

## 3 结语

本文利用混沌对初值的敏感性和伪随机性,提出一种用于图像篡改检测和恢复的安全双水印算法。与文献[8]的算法相比,本文的突出特点是水印的生成和嵌入方法不同:

- 1) 利用混沌映射加密恢复特征生成图像块的恢复双水印;
- 2) 利用混沌映射并结合图像块内容和嵌入在低位恢复水印生成认证水印;
- 3) 水印嵌入分恢复水印嵌入和认证水印嵌入两个阶段。

此外,本文还从理论和实验两方面验证了本文算法的不可见性,并利用统计实验分析了本文算法抵抗“字典搜索”攻击和“内容篡改”攻击的能力。理论分析和实验结果表明该算法不仅具有好的篡改恢复质量和不可见性,而且有效地提高了自恢复双水印算法的安全性。

## 参考文献:

- [1] 和红杰,张家树.基于混沌置乱的分块自嵌入水印算法[J].通信学报,2006, 27(7): 80-87.