

文章编号:1001-9081(2006)10-2357-03

## 一种最优化链码指纹二值细化图像压缩编码

李超, 杜赓, 杨义先, 钮心忻

(北京邮电大学 信息安全中心, 北京 100876)

(chaolee@126.com)

**摘要:**提出了一种适合于对线状结构的条形纹线二值图像进行压缩的最优化 Freeman 链码压缩算法——Freeman 差分链码 Huffman 编码。与传统的 Freeman 链码相比,提出的压缩算法是基于 Freeman 链码、差分编码和 Huffman 编码的一种混和编码方式。通过理论分析和在指纹二值细化图上的实验结果证明,对于指纹二值细化图像,本算法优于现有的链码压缩二值图像的算法,针对于线状结构的条形纹线二值图像,本算法也优于其他压缩算法。其平均码长为 1.7651bits,低于 8 方向 Freeman 链码或者 Freeman 差分链码的 3bits 的平均码长。

**关键词:**Freeman 链码; 二值图像; 压缩编码

**中图分类号:**TP391.41 **文献标识码:**A

## Optimized chain code compression algorithm for fingerprint binary image

LI Chao, DU Geng, YANG Yi-xian, NIU Xin-xin

(Centre of Information Security, Beijing University of Posts and Telecommunications, Beijing 100876, China)

**Abstract:** An optimized compression algorithm for linear structure stripe image was discussed in this paper, which is Freeman differential chain code Huffman coding. Compared to the traditional Freeman chain code, the proposed one is a hybrid encoding method based on Freeman chain code, differential code and Huffman code. Theoretic analysis and results from experiments on fingerprint binary images show that this algorithm is superior to other binary image compression algorithms, especially to fingerprint binary image compression. The average code length of the proposed one in this paper is 1.7651bits, which is shorter than that of 8 orientations Freeman chain code or Freeman differential chain code whose average code length is 3 bits.

**Key words:** Freeman chain code; binary image; compression coding

## 0 引言

随着计算机的不断发展,计算机图像的使用越来越广泛。在图像数据通信和存储领域,条形纹线编码是一种十分重要的技术。目前,有多种线状结构的条形纹线图像,例如指纹、文字、印刷电路板、轮廓图等等,对这些图像的有效压缩,将可以节省大量存储和网络资源。

在二值图像的压缩编码方法中,比较流行的编码方法有游程编码(RLC)<sup>[1]</sup>、LZW 编码<sup>[1]</sup>、Freeman 链码(FCC)、算术编码和二进制小波图像编码<sup>[3]</sup>等等。在几何图像编码中,有效的编码率和可靠的图像再现都是十分重要的<sup>[4]</sup>。Freeman 提出的链码是一种简洁有效的针对曲线、字符这类线状结构图像的编码方法。FCC 对此类图像的编码效率较游程类编码效率高<sup>[5]</sup>,甚至一些利用小波变换的方法对此类线条纹进行有损压缩其编码效率也不比 FCC 的编码效率高<sup>[6]</sup>。由此可以看出 FCC 是对线条纹图像进行无损压缩的一种高效的编码方式。因此基于以上考虑,在对指纹的二值细化图库反复试验后,提出了一种最优化 Freeman 链码压缩算法——Freeman 差分链码 Huffman 编码,即 FCC、差分编码和 Huffman 编码的混和编码方式,简称为 FDHHC 编码(Freeman Differential Huffman Hybrid Coding)。

## 1 理论基础

### 1.1 Freeman 编码

在 FCC 中,沿着数字曲线移动,对数字曲线上的每个像素以 8 邻接的方式进行移动,每次移动的方向均记录在一个链表里,这样遍历曲线上每个像素,就形成一个矢量链或者 Freeman 链。矢量链中的每个矢量方向由数字集  $\{i | i = 1, 2, \dots, 7\}$  分别来表示,其方向分别为与 X 轴正方向  $45^\circ \times i$  的夹角<sup>[2]</sup>。链码编码是用 3bits 的数据来表示任意几何曲线上的像素(图 1)。本文中,我们对指纹二值细化图像进行特征统计,根据统计值对矢量链编码方法进一步改进优化,以提高对其进行压缩编码的效率。

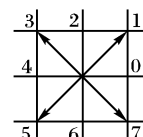


图 1 Freeman 链码

### 1.2 Huffman 编码

Huffman 编码是最常用的一种变长编码, Huffman 编码对于出现概率不同但固定的符号进行编码是最佳的,因此被称为最佳变长编码。其编码的基本原则是对出现概率较高的符

收稿日期:2006-04-05;修订日期:2006-06-09 基金项目:北京市自然科学基金资助项目(4042022;4062025)

作者简介:李超(1979-),男,山东金乡人,博士研究生,主要研究方向:信息安全、生物识别; 杨义先(1963-),男,四川盐亭人,教授,博士生导师,主要研究方向:网络信息安全、编码密码学、伪装式信息安全、应用数学等。

号用较短的编码表示,对出现概率较低的符号用较长的编码表示,以达到平均码长最短的目的。图 2 是 Huffman 编码的过程。

根据图 2 中的编码可以得到各个信源符号的编码:C0:0, C1:10, C2:11011, C3:110100, C5:110101, C6:1100, C7:111。

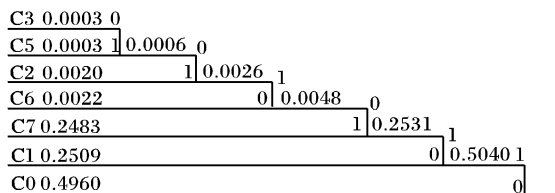


图 2 Huffman 编码

## 2 FDHHC 编码

### 2.1 Freeman 差分链码编码方法

研究指纹二值细化图时,不难发现指纹纹线基本是平滑连续的,用 Freeman 链码来表示指纹纹线,则当前编码和其后一个编码经常是相同或者相邻的。对 3 000 幅指纹二值细化图中相邻纹线点的方向变化进行了统计,结果见表 1。

表 1 变化角度值概率表 表 2 用链码差分编码表示的概率

角度差(度)	出现概率	差分链码编码	出现概率
0	0.4960	0	0.4960
45	0.2509	1	0.2509
90	0.0020	2	0.0020
135	0.0003	3	0.0003
180	0.0000	4	0.0000
225	0.0003	5	0.0003
270	0.0022	6	0.0022
315	0.2483	7	0.2483

由于指纹二值细化图中相邻的纹线点角度变化不大,因此借鉴差分编码的思想,结合 Freeman 链码和差分编码来表示各个角度差<sup>[7]</sup>。即如果指纹二值细化图中相邻纹线点的方向差是 0°,用 0 表示,如果方向差是 45°,用 1 表示,依此类推。这种编码方式我们称之为差分链码<sup>[8]</sup>。根据这种编码思想,表 1 的角度差可以变换为表 2 的编码。由于指纹纹线的特殊性,相邻两个像素之间的方向差不可能是 180°,故出现 180 方向差的概率为 0。

### 2.2 FDHHC 编码

由于表 2 中的链码概率分布区域较大,采用定长编码显然不是最佳编码,达不到最大压缩比。因此,我们考虑再对差分链码进行 Huffman 编码。差分链码 0、1、2、3、4、5、6、7 对应的 Huffman 编码的码值分别为:C0:0, C1:10, C2:11011, C3:110100, C5:110101, C6:1100, C7:111。具体编码过程见图 2。

这样经过 Huffman 编码的平均码长为:

$$\bar{K} = \sum_{i=0}^7 k_i \times p_i = 1.7651 \text{ bit}$$

其中  $k_i$  表示第  $i$  个码值的编码位数,  $p_i$  表示其出现的概率。这一编码方法我们称之为 FDHHC 编码,即 Freeman 差分链码 Huffman 混和编码。

### 2.3 FDHHC 在指纹二值细化图压缩中的应用

在指纹二值细化图的压缩中,我们考虑单独对每一条纹线进行压缩,依次压缩图像中所有的纹线,从而达到对整幅图像压缩的目的。下面讨论对纹线进行压缩的方法。首先找

到任意一条纹线的左端点,以其作为起始点顺序向右遍历纹线上的所有像素。这样只需要记住纹线第一个像素的位置就可以利用编码恢复出来整条纹线了。由于记录一个像素点的位置需要占用 4 字节的空间(640×640 的图像,纵横坐标各用 2 个字节表示),而一条纹线所有像素经过 Freeman 差分链码 Huffman 混和编码后的长度也仅仅只有几个字节,所以很好的处理每条纹线的起始点对压缩整幅图像十分重要。对于 640×640 的指纹图像,起始点位置最少可以用 20bits(纵横坐标分别用 10bits)表示。而 512×512 的指纹图像,起始点位置坐标最少可以用 18bits 来表示,所以根据图像大小合理选择坐标的表示方法能有效节省起始点的存储空间,整幅图像的压缩比也相应的增大。另外尽量减少纹线数量也是提高整幅图像压缩比的一个重要方法。其中对于分叉点的处理最为关键,如果遇到分叉就断开,无疑增加了纹线的数量,使得起始点增多,降低了图像的压缩比。如果利用编码表示纹线的分叉,就会增添编码的复杂度,而且编码效率不会明显提高甚至会降低,文献[9]详细的讨论了编码分叉点对纹线压缩的影响。在本文的算法中,遇到分叉点时,按照 0°~360°的顺序查找相邻像素作为该纹线上分叉点的下一个像素,另外的像素则按照其他纹线上的点来处理。这样就相当于延长了纹线的长度,减少了起始点的数量,从而节省了存储空间。

对于一条纹线,以图 3 中的指纹纹线为例说明其编码表示方法:



图 3 指纹纹线的一部分

1) 首先从左到右对其用 Freeman 链码表示: 010110101011011,其中第一个像素编码定义为 0;

2) 将上述 Freeman 链码转换成差分链码: 017107171710710;

3) 最后对差分链码进行 Huffman 编码: 01011110011110111100111101。

依次对一幅指纹二值细化图上所有的纹线遍历编码并记录其相应的起始位置,从而达到压缩整幅图像的目的。

## 3 几种压缩编码方式的比较

首先,比较一下 Freeman 链码、顶点链码和 FDHHC 编码的平均码长。8 方向 Freeman 差分链码有 8 个码值,如果对其进行等长编码,码长是 3bits,而进行 Huffman 变长编码,即 FDHHC 编码,则平均码长为 1.7651bits。顶点链码分别有 4 个和 3 个码值,使用 2 位二进制进行等长编码,码长是 2bits。由此可见,FDHHC 编码的平均码长最短。而表示一条数字曲线的链码平均位数等于平均码长乘链码的长度。因此表示指纹二值细化图这样的图像,在上述几种编码中,FDHHC 所用的平均码长最小,即压缩比最大。

然后,比较一下其他不定长编码,如游程编码 RLC、算术编码。由图 4,可以看出指纹二值细化图中背景是白色的,纹线信息仅仅是黑色像素部分。如果用游程编码进行对其进行压缩编码,其编码不仅要表示白色背景部分的信息,而且要表示纹线上的每个像素的信息,即黑色像素的信息。根据 RLC 的编码方法,通过比较可知其压缩指纹二值细化图的效果不如链码好,其原因是链码只记录了纹线信息,把其他部分都视为冗余信息,而 RLC 编码所记录的信息量要比链码大得多<sup>[4]</sup>。算术编码可以记录纹线信息,但是随着纹线的增长,

表示纹线的符号就越来越多,那么如果一条纹线上有 50 个像素,那么需要的表示符号也就有 50 个,那么用算术编码就会变得很困难了,因为那样需要很高的计算精确度,操作起来十分麻烦。实际上一条纹线一般都会有 50 以上的像素,多则达到 200 多个像素,那样用算术编码来表示纹线就十分困难。由此可以看出,FDHHC 编码和其他变长编码方式相比更适合于指纹二值细化图这样的线条纹图像压缩编码。



图 4 指纹二值细化图

FDHHC 编码和现有的其他链码一样存在一个误码的问题。如果一个在编码过程中,每个码值的错误都可能对其后的码值产生影响,这也是大多数压缩编码技术都存在的问题。因为在压缩信息中的每个码都与大量的图像信息想关联,所以一个码值出错其就会关联到很多信息。但是这个问题并不会影响压缩编码技术的应用,因为循环冗余校验技术 CRC 可以满足校验信息无误的传输。

#### 4 结语

本文提出了一种基于链码、差分编码和 Huffman 编码的混和编码方式——FDHHC 编码,FDHHC 编码与现有的几种链码比较结果说明,其平均码长更短,对指纹二值细化图的压缩比更大,与游程编码和算术编码相比其可操作性和压缩效

果也有很明显的优点。

鉴于 FDHHC 编码的优越性,对于文字、印刷电路板、轮廓图等线状结构的条形纹线图像的压缩编码,该编码方式是一个很好的选择。

#### 参考文献:

- [1] GONZALEZ RC, WOODS RE. 数字图像处理[M]. 第 2 版. 阮秋琦, 等译. 北京: 电子工业出版社, 2004.
- [2] FREEMAN H. Computer processing of line drawing image Comput [Z]. Surveys, 1974, 6: 57 - 97.
- [3] MULLER G, KLODIZ C, BONIS J, *et al.* Progressive transmission of line drawings using the wavelet transform[J]. IEEE Trans. on Image Processing, 1996, 5(4): 666 - 672.
- [4] 张宇, 尹昊晖, 张家谋. 图像质量客观测试的研究[J]. 北京邮电大学学报, 1999, 04: 39 - 40.
- [5] LIU K, PRASAD R. Comparing coding efficiency of vector chain coding and run-length coding for line drawings[J]. Communications, Speech and Vision, IEE Proceedings I, 1991, 138(5): 363 - 370.
- [6] SWANSON MD, TEWFIT AH. A binary wavelet decomposition of binary images[J]. IEEE Trans. On Image Processing, 1996, 5 (12): 1637 - 1650.
- [7] CHUNG J-W, MOON J-H, KIM J-K. Conditional differential chain coding for lossless representation of object contour[J]. Electronics Letters, 1998, 34(1): 55 - 56.
- [8] GOKMEN M, ERSOY I, JAIN AK. Compression of fingerprint images using hybrid image model[J]. Image Processing, 1996, 3: 395 - 398.
- [9] O'GORMAN L. Primitives chain code[J]. Acoustics, Speech, and Signal Processing, 1988, 2: 792 - 795.

(上接第 2356 页)

ENL 越大,表示越小受斑点的影响,相干斑抑制的能力越强。表 1 列出了各种情况下的 ENL。表 1 的数据也显示,本文方法得到的结果,要好于 PCA 方法。这与上面的主观评价结论是相同的。

表 1 各种情况下 ENL 的比较

原始极化 SAR 图像			PCA			ICA		
HH	HV	VV	PC1	PC2	PC3	IC1	IC2	IC3
4.9	7.8	3.8	16.1	4.6	7.7	21.0	9.8	10.5

#### 4 结语

本文将极化 SAR 图像相干斑的乘积模型,变换为应用 ICA 的信号独立加噪模型。并且将 HV/VV 的比值图像,也作为 ICA 的输入数据。利用 ICA 的分离性,得到了分别对应于 HH、HV 和 VV 极化的三幅降噪图像。实验表明,应用 ICA 算法可以有效地降低极化 SAR 图像的相干斑噪声,提高图像质量。随着研究的不断深入,相信 ICA 作为一种全新的信号处理方法,今后会在 SAR 图像处理等许多领域发挥更多的作用。

#### 参考文献:

- [1] PI Y-M, *et al.* Polarimetric Speckle Reduction Using Multi-Texture Maximum Likelihood Method [J]. IEEE-electronic Letter, UK, 2003, 39 (18): 1348 - 1349.
- [2] 杨竹青, 李勇, 胡德文. 独立成分分析方法综述[J]. 自动化学报, 2002, 28(5): 762 - 772.
- [3] FIORI S, PIAZZA F. A. General Class of  $\psi$ -APEX PCA Neural Al-

gorithms [J]. IEEE Transactions on Circuits and Systems-Part I, 2000, 47, (9): 1394 - 1398.

- [4] COSTA S, FIORI S. Image Compression Using Principal Component Neural Networks [J]. Image and Vision Computing Journal (special issue on Artificial Neural Network for Image Analysis and Computer Vision), 2001, 19(9 - 10): 649 - 668.
- [5] CHENG J, MILLER E. Model-based principal component techniques for detection of buried landmines in multiframe synthetic aperture radar images[A]. 2002 IEEE International Geoscience and Remote Sensing Symposium (IGARSS'02)[C]. 2002. 334 - 336.
- [6] CHEN CH, Zhan X. On the roles of independent component analysis in remote sensing[A]. Proc. of Progress in Electromagnetics Research Symposium (PIERS'2000) [C]. Cambridge (MA, USA), 2000.
- [7] CHITROUB S, HOUACINE A, SANSAL B. Statistical characterisation and modelling of SAR images[J]. Signal Processing, 2002, 82 (1): 69 - 92.
- [8] COMON P. Independence component analysis-a new concept? [J]. Signal Processing, 1994, 36: 287 - 314.
- [9] HYVÄRINEN A. Fast and Robust Fixed-Point Algorithms for Independent Component Analysis [J]. IEEE Transactions on Neural Networks, 1999, 10(3): 626 - 634.
- [10] OLIVER C, QUEGAN S. Understanding Synthetic Aperture Radar Images [M]. Artech-House, London, 1998.
- [11] CICHOCKI A, AMARI S. Adaptive Blind Signal and Image Processing[M]. John Wiley & Sons, New York, New revised and improved edition, 2003.