

一种嵌入式头肩图像编码方法

卓 力, 沈兰荪, 张廷华

(北京工业大学信号与信息处理研究室, 北京 100022)

摘 要: 头肩序列图像是可视电话、会议电视等视频业务中常见的一种图像模式. 本文将 SPIHT 算法与人脸检测算法相结合, 提出一种嵌入式的头肩图像编码方法. 首先采用人脸检测技术确定头肩图像中的人脸区域; 然后产生人脸区的掩膜, 并对 LL 子带中人脸区掩膜内的小波系数进行定标, 以保证人脸区的小波系数能优先编码; 最后采用 SPIHT 算法对人脸区和背景区进行编码. 实验结果表明, 与原来的 SPIHT 方法相比, 这种方法可以保证人脸区域的重建质量好于背景区的重建图像. 同时压缩后的码流仍然具有嵌入的特性, 支持渐进传输.

关键词: 头肩图像; 人脸检测; SPIHT 算法; 小波变换

中图分类号: TN919. 81 **文献标识码:** A **文章编号:** 0372-2112 (2003) 12-1832-03

Embedded Head2Shoulder Image Coding Method

ZHUO Li, SHEN Lan2sun, ZHANG Yan2hua

(Signal & Information Processing Lab, Beijing Polytechnic University, Beijing 100022, China)

Abstract: In this paper, human face detection algorithm is combined with the set partitioning in hierarchical trees (SPIHT) algorithm, and an embedded head&shoulder image coding method is proposed. The human face detection algorithm is employed firstly to determine human face region in a head&shoulder image. The human face region mask is generated and the wavelet coefficients in the human face region mask of LL subband is scaled to ensure that they are encoded with higher priority. Finally, the SPIHT algorithm is implanted to encode human face region and background region. The experimental results show that the reconstructed quality of human face region is much better than that of background region. At the same time, the encoded bitstream is fully embedded and supports progressive transmission.

Key words: head shoulder images; human face detection; set partitioning in hierarchical trees (SPIHT) algorithm; wavelet transform

1 引言

头肩序列图像是可视电话、会议电视等视频业务中常见的一种图像模式. 在人机交互、虚拟现实、门卫安全中也有广泛的应用前景. 随着多媒体技术的发展, 人们不断在探讨新的头肩序列图像的编码方法.

基于小波的图像编码能充分利用人眼的视觉特性, 已经成为图像编码的一个研究热点. 在基于小波的图像编码方案中, Shapiro^[1]等提出的嵌入式零树小波 (EZW, Embedded Zerotree Wavelets) 方法最具代表性. 在 EZW 的基础上, Said 和 Pearlman 又提出了性能更好的多级树集合分裂算法^[2] (SPDHT, Set Partitioning In Hierarchical Trees). SPIHT 算法不仅能以较小的运算量实现对图像的高效压缩, 还可以得到完全嵌入的、具有连续可分级特性的输出码流, 便于实现精确的码率控制, 支持渐进传输. 在 SPIHT 算法的基础上, 近年来人们又提出了基于对象的、基于感兴趣区的等多种编码方法.

本文将 SPIHT 算法与人脸检测算法相结合, 提出一种新的嵌入式头肩序列图像编码方法. 首先采用人脸检测技术确定头肩图像中的人脸区域, 对人脸区内的小波系数进行特殊处理使其能够优先于背景区的其他系数进行编码, 然后再采

用 SPIHT 算法对人脸区和背景区进行编码. 这种方法可保证在任何码率下人脸区域的重建质量均好于背景区的重建图像. 同时压缩后的码流仍具有嵌入特性, 能够支持渐进传输.

2 头肩图像的编码

2.1 人脸区域的确定

头肩图像人脸区域的确定要采用人脸检测技术, 用以确定一幅图像中是否包含人脸, 以及人脸的个数、位置、区域大小等参数. 目前人们已经提出了多种人脸检测技术, 如模板匹配、基于人脸规则的检测、统计学习、基于彩色信息的检测、本征脸技术等. 虽然这些方法在各自的应用背景下取得了一定的效果, 但由于受到人脸姿势、光照、成像条件等限制, 目前还没有一种算法能够不加限制地应用于所有场合.

本文采用文献[6, 7]中的方法来确定图像中人脸区域. 这种方法对姿势、光照等具有很强的鲁棒性, 也不限制图像的背景, 特别适合于头肩序列图像的处理. 为了方便进一步的编码操作, 本文将人脸区划定为一个长方形的区域. 图 1 所示的是人脸检测的实验结果(图中白框仅用于表示人脸已被正确检测出来).



图1 人脸检测实验结果

21.2 人脸区掩膜的产生

若要对人脸区的小波系数进行特殊处理,则必须首先区分每个小波子带内用于重建人脸区和背景区的小波系数^[3,8].本文采用掩膜(Mask)来标记人脸区的小波系数.掩膜的确定取决于合成小波滤波器的长度以及小波分解的级数.以5-3小波滤波器为例,假设原始象素点为 $X(2n)$ 和 $X(2n+1)$,在第一级分解中低频和高频子带的象素点分别为 $l(n)$ 和 $h(n)$,则重建 $X(2n)$ 和 $X(2n+1)$ 的公式为:

$$X(2n) = l(n) - \frac{h(n-1) + h(n)}{4} \quad (1)$$

$$X(2n+1) = \frac{l(n) + h(n+1)}{2} + \frac{-h(n-1) + 6h(n) - h(n+1)}{8} \quad (2)$$

由于重建 $X(2n)$ 和 $X(2n+1)$ 时需要用到 $l(n)$, $l(n+1)$, $h(n-1)$, $h(n)$, $h(n+1)$ 等四个系数,那么这些系数就在掩膜内.每级分解后都可以确定掩膜的行和列,这样逐级分解下去,可以最终确定每个子带内人脸区的掩膜.图2中所示的是对头肩像进行三级小波分解后由空间象素域的人脸区推导各小波子带内人脸区掩膜的过程^[9].

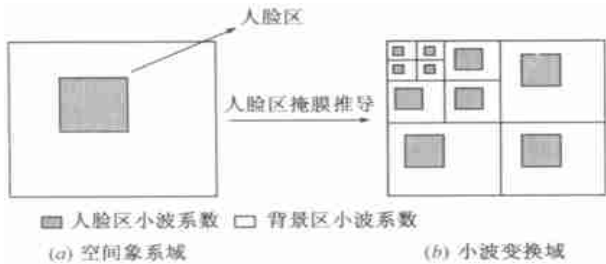


图2 人脸区掩膜的确定示意图

21.3 人脸区系数的定标

人脸区域的重建质量对于重建图像的主观感受有着非常重要的作用.在采用SPIHT算法对整幅图像进行编码之前,需要对人脸区掩膜内的小波系数进行定标(Scaling),以保证人脸区的小波系数能够进行优先编码.

SPIHT算法依据各个小波系数的重要性程度,采用逐次逼近的量化方法对系数进行编码.因此在某种程度上说,SPDHT的编码过程也是一种比特平面编码.如果将人脸区的小波系数所处的平面上移得足够多,使得人脸区的小波系数幅

值大于背景区的系数幅度值,则人脸区的小波系数就能够优先于背景区的系数被编码.考虑到LL子带的小波系数对重建图像的作用最大,我们仅将小波图像LL子带中人脸区掩膜内的小波系数进行比特平面上移.比特平面上移相当于使小波系数同时乘以2的幂次方,即

$$C(i, j) = C(i, j) @ 2^{\text{Shift_value}} \quad (2)$$

其中 $C(i, j) \in \text{LL}_{FR}$, LL_{FR} 表示LL子带中的人脸区, Shift_value 为上移的比特平面数.为了保证人脸区的小波系数能首先被编码,本文采用如下的方法确定 Shift_value :

$$\text{Shift_value} = N - M + 1 \quad (3)$$

其中 $N = \lceil \log_2[\max\{c(i, j)\}] \rceil + 8$, $c(i, j) \in \text{LL}_{BG}$, $M = \lceil \log_2[\min\{c(i, j)\}] \rceil + 8$, $c(i, j) \in \text{LL}_{FR}$, LL_{BG} 表示LL子带中的背景区.经过这样的定标操作,就可以保证LL子带中人脸区最小的小波系数大于背景区最大的小波系数,从而首先被编码.图3是定标示意图.

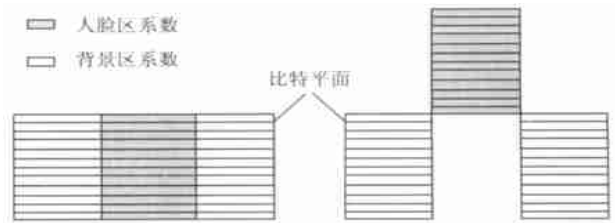


图3 LL子带中人脸区系数定标示意图

在编码端,先对LL子带中的人脸区小波系数进行比特平面上移,然后再采用SPIHT算法对整幅图像进行编码;而在解码端,先用SPIHT算法解码,再进行比特平面下移还原成原来的小波系数幅度值. Shift_value 和人脸区掩膜的起始点,掩膜大小等参数必须传送给解码器,这需要占用30个比特的,但这点开销对整个图像码流来说几乎没有什么影响.

3 实验结果

本文分别对Lena(512@512, 8bpp)和Barbara(576@720, 8bpp)两幅单色标准图像进行了实验,两幅图像具有不同的纹理特点.采用9-7小波滤波器分别对两幅图像进行5级和3级分解,图像边缘采用对称扩展.

图4所示是采用本文方法对Lena图像编码的结果.表1

所列是在不同编码速率下对 Lena 整幅图像和人脸区编码的结果,并将本文提出的方法和原来的 SPIHT 算法进行了比较. Barbara 图像人脸区和整幅图像的编码比较结果示于图 5 中.

从图 4.5 及表 1 均可以看出,原来的 SPIHT 算法对图像

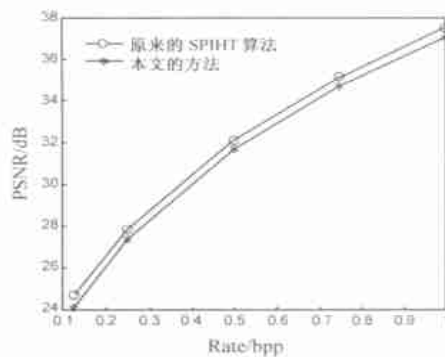
进行编码时,对人脸区和背景区等同对待,人脸区和背景区的重建质量近似,在极低码率时,人脸区的重建质量差.而在任何码率下,采用本文方法得到的人脸区重建质量均远远好于背景区的重建质量.



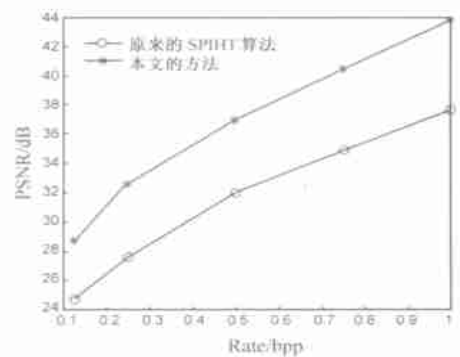
图 4 采用本文方法对 Lena 图像编码的结果

表 1 本文方法和 SPIHT 方法对 Lena 图像编码结果的比较

Rate/bpp	本文方法		原来的 SPIHT 方法	
	人脸区 /dB	整幅图像 /dB	人脸区 /dB	整幅图像 /dB
0.125	33.87	29.59	29.37	31.08
0.25	37.60	32.73	32.82	34.12
0.50	41.91	35.90	36.75	37.24
0.75	45.59	37.85	38.76	39.08
1.00	48.55	39.44	40.68	40.45



(a) 整幅图像编码结果比较



(b) 人脸区的编码结果比较

4 结论

本文提出的头肩图像编码方法可以保证在任何码率下人脸区的重建质量均好于背景区的重建质量.同时码流具有嵌入的特性,便于渐进传输.在无线、Internet 等带宽受限的网络中传输时,只要对人脸区的重要信息给予更高权级的抗误码保护,则可以保证网络状况恶化时重建图像有清晰的人脸.

参考文献:

- [1] J M Shapiro. Embedded image coding using zerotrees of wavelet coefficients[J]. IEEE Trans Signal Processing, 1993, 41(12): 3445- 3462.
- [2] Amir Said, William A Pearlman. A new fast and efficient image codec based on set partitioning in hierarchical trees[J]. IEEE Trans Circuits and Systems for Video Technology, 1996(16): 243- 250.
- [3] Keunzhyeong Park, Hyun Wook Park. Region of interest coding based on set partitioning in hierarchical trees[J]. IEEE Trans. Circuits and Systems for Video Technology, 2002, 12(2): 106- 113.
- [4] A Kassim, L Zhao. Rate-scalable object-based wavelet codec with implicit shape coding[J]. IEEE Trans Circuits and Systems for Video Technology, 2000, 10(7): 1068- 1079.
- [5] 沈兰荪,卓力,等. 视频编码与低速率传输. 北京: 电子工业出版社, 2001.
- [6] Shen Lansun, Wang Kongqiao, Xing Xin. Automatic human face detection and tracing in a complex background[J]. Chinese Journal of Elec-

tronics, 2000, 9(1): 65- 69.

- [7] Xing Xin, Shen Lansun, Jia Kebin. A scale and pose invariant algorithm for fast detecting human faces in a complex background[J]. Chinese Journal of Electronics, 2001, 10(2): 147- 152.
- [8] Charilaos Christopoulos, et al. Efficient methods for encoding regions of interest in the updating JPEG2000 Still Image Coding Standard[J]. IEEE Signal Processing Letters, 2000, 7(9): 247- 249.
- [9] Raphael Grosbois, Diego Santacruz. New approach to JPEG 2000 compliant region of interest coding[A]. Proc. SPIE. 2001[C]. CA USA: San Diego, 2001(1): 9.

作者简介:



卓力 女, 1971 年生于江苏徐州, 1992 年和 1998 年分别获得电子科技大学和东南大学学士和硕士学位, 现为北京工业大学讲师、博士生, 近年来发表论文多篇, 主要研究方向为视频编码、无线 IP 视频传输等. E-mail: siplsvr@bjpu. edu. cn.

沈兰荪 男, 1938 年生于江苏苏州, 1961 年毕业于北京邮电学院通信工程系, 现为北京工业大学教授、博士生导师, IEEE 学会高级会员, 研究领域为图像处理、图像传输、图像压缩和光谱信号的检测等, 发表学术论文二百余篇, 出版专著十部. E-mail: slx@bjpu. edu. cn.