

视频监控应用中图像压缩算法的研究

镇之皓,郭前岗,周西峰

(南京邮电大学 自动化学院,江苏 南京 210046)

摘要:基于 DCT 的 JPEG 图像压缩算法是视频监控应用中的核心技术。虽然传统的 DCT 其画质有保障,但其复杂度较大,不利于视频实时传输。所以,文中主要研究的目的就是降低其算法复杂度,并且保证传输图像的质量。基于图像能量都集中于 DCT 系数的低频部分这一特性,文中提出局部截取法,可以快速得到图像的重要信息,在没有明显影响图像质量的情况下,使得二维 8×8 DCT 的计算量大幅降低。由实验论证,在 PSNR 没有明显下降的情况下,算法复杂度明显降低。

关键词:离散余弦变换;JPEG 图像压缩算法;峰值信噪比

中图分类号:TP301.6

文献标识码:A

文章编号:1673-629X(2014)04-0084-04

doi:10.3969/j.issn.1673-629X.2014.04.021

Research on Image Compression Algorithm of Video Surveillance Applications

ZHEN Zhi-hao, GUO Qian-gang, ZHOU Xi-feng

(College of Automation, Nanjing University of Posts and Telecommunications, Nanjing 210046, China)

Abstract: DCT-based JPEG image compression algorithm is the core technology of video surveillance applications. Picture quality of conventional DCT is guaranteed, but its complexity is greater, which is not conducive to real-time video transmission. Therefore, the main purpose is to reduce the complexity of the algorithm, and to ensure the transmission quality of the image. Based on this feature that image energy is concentrated in the low-frequency portion of the DCT coefficient, propose a partial intercept method which makes the calculation of 8×8 DCT be significantly reduced, and you can quickly get important information about the image. The results of experiments show that in the case of no significant decrease in PSNR, the algorithm complexity is significantly reduced.

Key words: Discrete Cosine Transform (DCT); JPEG image compression algorithm; PSNR

0 引言

21 世纪以来,视频监控系统在诸多领域有着广泛的应用。随着图像与视频处理技术、网络技术和自动控制技术的发展,视频监控系统已从早期模拟监控过渡到数字化网络监控。数字视频监控系统是以数字视频的压缩、传输、存储和播放为核心,并采用先进的数字图像压缩、编码、解码和传输技术,从而实现可视化监控。现在常用的视频核心编解码技术有 M-JPEG, MPEG 和 H.264^[1]等算法,这里考虑到视频监控要求压缩图像质量较好,传输时画质较高,因此视频监控系统中选用 M-JPEG 算法更为客观。M-JPEG 算法是一种运动静止图像压缩技术,它是从静态图像 JPEG 压缩算法发展而来,可顺序的从视频中每一帧图像进

行压缩编码,即每一帧图像都被看做一个静止独立的图像进行处理。而离散余弦变换通常被认为是对语音和图像信号进行变换的最佳方式。文中即对基于 DCT 的 JPEG 图像压缩算法展开研究。

1 JPEG 图像压缩算法的现状分析

JPEG 全称 Joint Photographic Experts Group,即联合图像专家小组,是最常用的一种图像文件格式。JPEG 压缩算法主要通过选择性丢弃一些冗余信息,将本来存储空间很大的图片压缩成存储空间较小的图片。原来图片中一些重复或者人眼不容易察觉出来的信息将被丢弃,易造成图像数据损伤,因此是一种有损压缩格式,图像解压缩后会比原来图像质量有所下降。

收稿日期:2013-06-04

修回日期:2013-09-12

网络出版时间:2014-01-28

基金项目:国家自然科学基金资助项目(61105082)

作者简介:镇之皓(1987-),男,江苏如东人,硕士研究生,主要从事基于 ARM 的 WEB 远程视频监控系统的研究;郭前岗,教授,博士,主要从事电力电子变换控制与新能源发电技术的研究;周西峰,副教授,主要从事计算机监测控制技术的研

网络出版地址: <http://www.cnki.net/kcms/detail/61.1450.TP.20140128.1144.033.html>

采用基于 DCT 的压缩算法,其压缩比通常是 10:1 到 40:1 之间,而采用小波变换的压缩算法,其压缩比率可以达到 200:1。通过 Raph Levien 对比测试报告可以看出:

基于 DCT 的压缩算法在中低度压缩比率下表现更为出色;而基于小波变换的压缩算法在高压缩比下的表现更加令人满意。但是在平时的嵌入式应用和手持应用中,基于 DCT 的压缩算法更加合适这些低复杂度的运算。

2 基于 DCT 的压缩算法

2.1 JPEG 算法流程

JPEG 算法首先对图像进行分块处理,即分成 8 * 8 大小的像素块,然后对每一个像素分块进行离散余弦变换,其变换后的系数基本不相关,且系数矩阵的能量都集中在低频区域,即系数矩阵的左上角,通过保留低频系数,去除高频系数,来达到压缩的目的;其次再对矩阵系数进行量化和 Zig-Zag 模块变换;最后进行编码,实现图像的压缩,如图 1 所示。

DCT变换模块

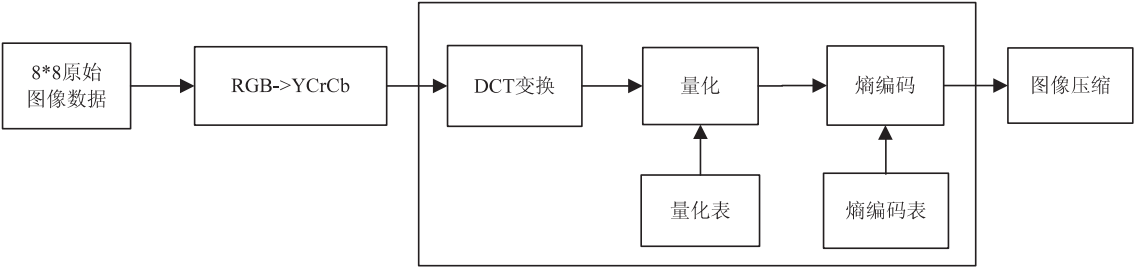


图 1 JPEG 图像压缩算法流程图

2.2 色彩空间的转换和采样模块

一般显示系统中采用 RGB (红绿蓝),然后通过这三个基本色彩混合成其他的色彩。而 YCbCr 更适用于图形压缩,其中 Y 表示色彩亮度,Cb、Cr 表示两个色差。采用 RGB 的图像信息比较大,且人的视网膜对于红绿蓝敏感程度不同。但是,人对于光线的亮度却具有最强的分辨能力,一般人对色彩的分辨力只有对光线分辨力的四分之一。所以首先得将显示系统采用的 RGB 转换成适合图形压缩的 YCbCr 模块^[2]。

RGB 到 YCbCr 的转换公式:

$$\begin{cases} Y = 0.299R + 0.587G + 0.114B \\ Cb = -0.169R - 0.331G + 0.5B + 128 \\ Cr = 0.5R - 0.419G - 0.081B + 128 \end{cases} \quad (1)$$

YCbCr 到 RGB 的逆变换公式:

$$\begin{cases} R = Y + 1.402(Cr - 128) \\ G = Y - 0.334(Cb - 128) - 0.714(Cr - 128) \\ B = Y + 1.722(Cb - 128) \end{cases} \quad (2)$$

由多个亮度分量子块和两个色度分量子块构成的一个完整编码单元称之为最小编码单元。由于后面处理图像时需要将其分成 8 * 8 的子块,所以在分割图像时,需要对初始图像进行填充,使其宽度和高度为 8 的倍数。

2.3 离散余弦变换(DCT)

在压缩之前,首先对初始图像进行预处理,将彩色图像转换成灰度图像,再把图像分割成 8 * 8 的像素块,图像数据值必须减去 128,是因为 DCT 转换公

式所接受的数字范围是在-128 到+127 之间,其相邻元素之间存在很强的相关性,而 DCT^[3]就是消除这些相关性。经过变换之后,作用域之内的总能量不变,但能量会重新分布,且在空间域中,能量分布具有一定随机性。

DCT 的变换公式:

$$F(u,v) = \frac{1}{4}c(u)c(v) \sum_{x=0}^7 \sum_{y=0}^7 f(x,y) \cos(\frac{2x+1}{16}u\pi) \cos(\frac{2y+1}{16}v\pi) \quad (3)$$

DCT 的逆变换公式(4)所示:

$$f(x,y) = \frac{1}{4} \sum_{u=0}^7 \sum_{v=0}^7 c(u)c(v) F(u,v) \cos(\frac{2x+1}{16}u\pi) \cos(\frac{2y+1}{16}v\pi) \quad (4)$$

式中 $x,y,u,v = 0,1,\cdots,7$;

$$c(u) = c(v) = \begin{cases} \frac{1}{\sqrt{2}}, u = 0, v = 0 \\ 1, \text{其他} \end{cases}$$

其中,x,y 代表图像数据矩阵内某个数值的坐标位置;f(x,y) 代表图像数据矩阵内的某个数值;u,v 代表 DCT 变换后矩阵内某个数值的坐标位置;F(u,v) 代表 DCT 变换后矩阵内的某个数值。经过 DCT 变换后的矩阵数据自然数为频率系数,这些系数以 F(0,0) 的值最大,称为 DC,其余的 63 个频率系数则多半是一些接近于 0 的正负浮点数,一概称之为 AC。而 f(x,y) 经 DCT 变换之后,F(0,0) 是直流系数,其他为交流系数。

DCT 变换后的 64 个 DCT 频率系数与 DCT 变换

前的 64 个像素块相对应,DCT 变换过程的前后都是 64 个点,说明这个过程只是一个没有压缩作用的无损变换过程。而且变换后,单独一个图像的全部 DCT 系数块的频谱几乎都集中在最左上角的系数块中,即低频分量集中在系数模块的左上角,高频分量集中在系数模块的右下角,其中低频分量包含了图像的主要信息,而高频分量是人眼不敏感的部分,所以高频部分就不显得那么重要了,这样就为之后的量化打好了基础。因此,可以看出 DCT 本身并不是用来压缩图像的,而是很好地把图像的能量集中在一起,这样便于以后图像能更好地进行压缩。DCT 的过程是可逆的,所以不经过量化,还是可以无损还原的。

2.4 量化与 Zig-Zag 变换模块

量化过程是对 DCT 系数的一个优化过程,是对 DCT 系数进行压缩的关键一步。量化过程是利用了人眼对高频部分不敏感的特性,去除掉图像中不重要的细节,降低 DCT 系数精度。其过程实际上是把频率领域上每个成分,除以一个对应于该成分的常数,且接着对得到的结果进行四舍五入,最后取最接近的整数。这个过程经常会把很多高频率的成分四舍五入而接近 0,剩下的系数很多会变成小的正或负数。整个量化过程的目的是减小非“0”系数的幅度,以及增加“0”值系数的数目。量化是整个过程中的主要有损运算,是图像质量下降的最主要的原因。

由于人眼对亮度信号比对色差信号更敏感,因此在量化过程中使用了亮度量化表,用 8 * 8 模块系数除以量化表(如图 2 所示)中的量化系数,最终达到保持低频,抑制高频的目的。

16	11	10	16	24	40	51	61
12	12	14	19	26	58	60	55
14	13	16	24	40	57	69	56
14	17	22	29	51	87	80	62
18	22	37	56	68	109	103	77
24	35	55	64	81	104	120	101
49	64	78	87	103	121	120	101
72	92	95	98	112	100	103	99

图 2 亮度量化表

量化后的数据,有一个很大的特点,就是直流分量相对于交流分量来说要大,而且交流分量中含有大量的 0。然后按图 3 进行“Z”形(Zig-Zag)^[4]扫描,这样可以使得相同或相邻频率系数在一堆序列中保持相邻的位置,最后转换成一维排列方式。这样通过“Z”形扫描后,这些非零的 DCT 系数就集中在一维数组的前面,而后面接着就是一串量化为零的 DCT 系数。这就为熵编码提供了条件。

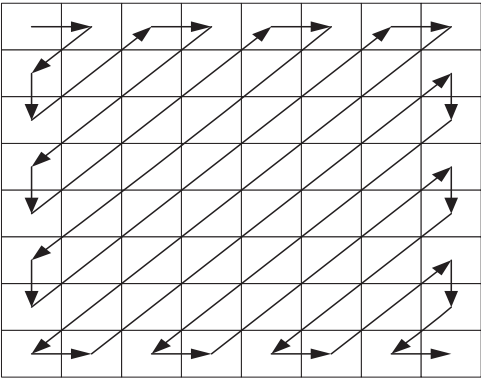


图 3 Zig-Zag 扫描示意图

2.5 熵编码模块

熵编码^[5]也就是只有非零系数被编码。而非零系数的编码又由两部分组成:其中前一部分是非零系数前连续零系数的数量,称为游程;而后一部分是非零系数的符号和幅值。由于之前的“Z”形扫描使得连续出现的零比较多,游程编码的效率自然就高。编码的时候一般有 4 个霍夫曼^[6]表,参照这些表,在熵编码的基础之上发现各个数值间的规律就可以进行编码了。

3 基于 DCT 的优化与仿真实验

3.1 传统 DCT 算法

首先,如果按照传统的二维 DCT^[7-8]算法来计算 8 * 8 像素,则获得一个 DCT 系数需要 8 * 8 = 64 次乘法和 8 * 8 = 64 次加法;而完成整个 8 * 8 像素的 DCT 系数则需要 8 * 8 * 8 = 4 096 次乘法,同样也需要 4 096 次加法。因此,可以看出其复杂度是相当得大。如果我们把二维的 DCT 算法转换成一维来计算,则需要 1 024 次乘法和 1 024 次加法,其计算量还是比较大。所以传统的压缩算法,图像在压缩和解压缩的过程中,都耗费了比较多的时间,这直接导致了视频监控系统有一定的滞后性。

3.2 对传统 DCT 的局部截取法

DCT 的局部截取法^[9]是基于 ANN^[10]算法,对传统 DCT 算法进行架构修改,即第一级为 4 个蝶形运算,第二级为乘法累加运算。对于一个 8 * 8 像素,获取其中一个 DCT 系数则只需要 5 次乘法和 29 次加法,而要获得整个 DCT 系数则需要 80 次乘法和 464 次加法。

由于图像的能量集中在低频部分,经过 DCT 变换后,其低频分量都集中在系数块的左上角部分。因此,利用这一特性,对 8 * 8 的系数块矩阵进行先行后列的 DCT 变换。在行变换中,由于低频分量集中在左上角,则每一行只取前 4 位,而后 4 位对于 DCT 变换没有影响,即忽略不计。这样对于一个 8 *

8 像素,其中一个点的 DCT 运算的行变换只需要 5 次乘法和 25 次加法。而对于完成整个 8 * 8 像素的 DCT 系数的行变换需要 40 次乘法和 200 次加法。同样,DCT 列变换的计算过程中也只取前 4 列系数。则第一列只需要计算前 4 个系数,第二列只需要计算前 3 个系数,第三列只需要计算前 2 个系数,第四列只需要计算第一个系数。所以完成整个 8 * 8 像素的 DCT 系数只需要 52 次乘法和 272 次加法。

通过上述算法修改计算量,得出其计算量的对比列表如表 1 所示。

表 1 三种压缩算法计算量对比

压缩算法	加法次数	乘法次数
传统 DCT 压缩算法	4 096	4 096
基于 ANN 压缩算法	464	80
基于 DCT 局部截取法	272	52

从表中可以明显看出,基于 DCT 低频系数中在左上角这一特性的算法,其运算量远小于前两种算法,可以给图像压缩和解压缩的过程节省不少时间,这样有利于视频监控的实时性。

3.3 基于 DCT 局部截取法的仿真实验

首先,实验都是基于 MATLAB7.1 这个软件平台。MATLAB 具有较强的数学运算能力,对图像处理^[11]也有较高的集成性。在其函数库中使用由 dctmtx 返回的 DCT 变换矩阵,此种函数适合较小的输入方阵。由于 DCT 系数为 8 * 8 的矩阵,则其函数调用为 $D = \text{dctmtx}(8)$,其中 D 为 DCT 变换矩阵。然后利用 blkproc 函数进行分块操作。其函数调用为 $B = \text{blkproc}(A, [m, n], \text{fun}, P_1, P_2, \dots)$,其中 A 为原始信号矩阵, $[m, n]$ 为分块大小,而 fun 就是压缩构造准则,后面的 P_i 是 fun 函数的参数。分块完成后利用公式(5)进行 DCT 转换。

$$T = H F H^T$$

(5)

其中, H 是 dctmtx 生成的 DCT 变换矩阵; F 为 8 * 8 图像块; T 是 DCT 变换的结果。

然后对其系数矩阵进行“Z”形扫描,去掉后面的一串零,将非零数字列为一维数组,最后进行熵编码,图像最终压缩完成。

而衡量一个图像压缩的画质最客观的依据就是 PSNR^[12],即峰值信噪比。而 PSNR 的计算公式如式(6)所示。通常情况下,PSNR 的值越大,代表图像失真越少;PSNR 的值越小,则代表图像失真比较严重。

$$\left\{ \begin{aligned} \text{PSNR} &= 10 \times \log \left(\frac{255^2}{\text{MSE}} \right) \\ \text{MSE} &= \frac{\sum_{n=1}^{\text{Framesize}} (1^n - P^n)}{\text{Framesize}} \end{aligned} \right.$$

(6)

仿真实验中,先将彩色图像转换成灰度图像,即 flower.bmp 作为源图像。其中传统 DCT.jpg 和 DCT 局部截取法.jpg 分别采用原始 DCT 运算和改进之后的 DCT 局部截取法运算。2 种图像压缩算法 PSNR 值的变化如表 2 所示。从表中可以看出,基于 DCT 局部截取法与传统 DCT 压缩算法的 PSNR 值相差 6 dB。压缩图像和源图像对比如图 4,可以看出传统 DCT 算法压缩后的图像跟源图像没有明显的损失,而经过 DCT 截取法获得的图像较源图像,其细节方面稍有损失,压缩效果比较满意。

表 2 两种图像压缩算法 PSNR 值的变化

实验对比	传统 DCT 压缩算法	基于 DCT 局部截取法
PSNR/dB	30.478 2	24.627 4

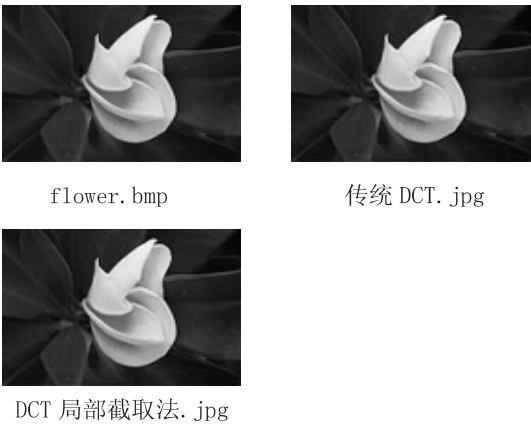


图 4 两种压缩算法与源图像对比

4 结束语

文中重点研究了基于 DCT 的图像压缩算法,并对传统的 DCT 算法进行局部截取,使得其改进的 DCT 算法复杂度降低,在保证图像压缩后画质没有明显下降的情况下,使 DCT 算法速度加快,这样就可以给图像压缩和解压缩的过程节省不少时间,有利于视频监控的实时性;其次对传统算法和优化算法在压缩之后的图像画质进行对比分析,优化压缩算法的画质比较理想。这样总体性能更适合视频监控的传输和画质要求。

基于视频监控的图像压缩算法还有很大的发展空间,文中还没有对实现高压缩比且画面失真较少的算法做研究。所以既保证高压缩比又保持画质的压缩算法是今后重点研究的问题。

参考文献:

[1] 刘 贞,张润彤. 视频压缩方法的文献综述[C]//Proceedings of the 2011 international conference on information, services and management engineering. [s. l.]: [s.],

表 1 基于 Corel 图像集的分类精度值 %

图像类别	DD-SVM	MILES	文中方法
Africa	66.70	68.8	78.0
Beach	68.4	66.0	81.0
Buildings	74.3	75.7	80.9
Buses	90.3	90.3	93.8
Dinosaurs	99.7	99.7	92.4
Elephants	76.0	77.7	80.1
Flowers	88.3	96.4	92.3
Horses	94.3	95.0	93.6
Mountains	70.3	71.0	78.0
Food	87.0	85.4	88.2
平均值	81.63	82.6	85.83

注:加粗的为在各类上得到的最好分类精度。

局部极大值来构建特征空间,然后采用 TSVM 训练分类器。文中算法特点在于运用 DD 算法能够找出样本集中最特殊的示例,运用直推式支持向量机能够有效利用未标记的样本来训练分类器,在基于 Corel 图像数据库的实验结果表明,文中算法是一种有效的图像分类算法。

参考文献:

[1] Dietterich T G, Lathrop R H, Lozano-Pérez T. Solving the multiple-instance problem with axis-parallel rectangles[J]. Artificial intelligence, 1997, 89(1/2): 31-71.

[2] Maron O, Lozano-Pérez T. A framework for multiple-instance learning[C]//Proceedings of the 1997 conference on advances in neural information processing systems 10. Cambridge, MA, SA: MIT Press, 1998: 570-576.

[3] Zhang Q, Goldman S A. EM-DD: An improved multiple-instance learning technique[C]//Proc of advances in neural information processing systems 14. Cambridge, MA, USA: MIT

Press, 2002: 1073-1080.

[4] Wang J, Zucker J D. Solving the multiple-instance problem: A lazy learning approach[C]//Proceedings of the 17th international conference on machine learning. San Francisco, CA: [s. n.], 2000: 1119-1125.

[5] Ruffo G. Learning single and multiple instance decision trees for computer security applications[D]. Torino, Italy: University of Turin, 2000.

[6] Andrews S, Hofmann T, Tsochantaridis I. Multiple instance learning with generalized support vector machines[C]//Proceedings of the 18th national conference on artificial intelligence. Edmonton, Canada: AAAI Press, 2002: 943-944.

[7] Chen Yixin, Wang J Z. Image categorization by learning and reasoning with regions[J]. Journal of machine learning research, 2004, 5(8): 913-939.

[8] Chen Yixin, Bi Jinbo, Wang J Z. MILES: Multiple-instance learning via embedded instance selection[J]. IEEE trans on pattern analysis and machine intelligence, 2006, 28(12): 1931-1947.

[9] Rahmani R, Goldman S A. MISSL: Multiple instance semi-supervised learning[C]//Proc of the 23rd int conf on machine learning. Pittsburgh: [s. n.], 2006: 705-712.

[10] Zhou Z H, Xu J M. On the relation between multi-instance learning and semi-supervised learning[C]//Proc of the 24th ICML. Corvalis: [s. n.], 2007: 1167-1174.

[11] Wang C H, Zhang L, Zhang H J. Graph-based multiple-instance learning for object-based image retrieval[C]//Proc of the 1st ACM int conf on multimedia information retrieval. Vancouver: [s. n.], 2008: 156-163.

[12] Joachims T. Transductive inference for text classification using support vector machines[C]//Proc of the 16th international conference on machine learning. San Francisco, USA: [s. n.], 1999: 200-209.

(上接第 87 页)

n.], 2011.

[2] 王镇道, 陈迪平, 文康益. JPEG 图像压缩算法的 IP 核设计[J]. 计算机应用, 2005, 25(5): 1076-1077.

[3] 潘 榕, 刘 昱, 侯正信, 等. 基于局部 DCT 系数的图像压缩感知编码与重构[J]. 自动化学报, 2011(6): 674-681.

[4] 万 源, 吴传生. MPEG-4 视频 DCT 量化模块的改进方法[J]. 计算机工程与应用, 2007, 43(12): 42-44.

[5] 安会乐. JPEG/MJPEG 中 Huffman 编解码的 IP 设计[D]. 西安: 西安电子科技大学, 2007.

[6] 李 薇, 胡伟文, 沈 静. 霍夫曼编码和游程编码在图像编码中的应用[J]. 舰船电子工程, 2010, 30(7): 67-69.

[7] 沈 洁, 杜宇人, 殷玲玲, 等. 基于 DCT 变换的图像压缩技术研究[J]. 信息技术, 2006(10): 133-134.

[8] 武 瑛. DCT 变换在图像压缩中的应用[J]. 计算机与现代化, 2013(4): 103-106.

[9] 张 羽, 梁德群. 基于图像分割改进的 DCT 压缩算法[C]//第十二届全国图象图形学学术会议论文集. 出版地不详: 中国图象图形学学会, 2005.

[10] Zhao Ming, Liu Zhibin, Ren Baosheng, et al. Improved ANN algorithm based on the change of search direction[C]//Proceedings of the 2010 international conference on computational and information sciences. [s. l.]: [s. n.], 2010.

[11] 涂望明, 魏友国, 施少敏. MATLAB 在数字图像处理中的应用[J]. 微计算机信息, 2007, 23(6): 299-300.

[12] 沙 亮, 王贵锦, 赵安邦, 等. H. 264 编码下基于 DCT 系数量化值的 PSNR 估计方法[J]. 清华大学学报(自然科学版), 2008(4): 510-513.

视频监控应用中图像压缩算法的研究

作者：[镇之皓](#)，[郭前岗](#)，[周西峰](#)，[ZHEN Zhi-hao](#)，[GUO Qian-gang](#)，[ZHOU Xi-feng](#)
作者单位：[南京邮电大学 自动化学院, 江苏 南京, 210046](#)
刊名：[计算机技术与发展](#)

英文刊名：[Computer Technology and Development](#)

年，卷(期)：2014(4)

本文链接：http://d.wanfangdata.com.cn/Periodical_wjfz201404021.aspx