

基于视频会议系统的 DCT 和 Inverse-DCT 算法研究及实现

闫晶, 宋文爱

(中北大学 仪器科学与动态测试教育部重点实验室, 山西 太原 030051)

摘要: 视频图像压缩是视频会议系统能否提供良好性能的关键技术之一。要想最大限度地利用有限的网络带宽, 在视频图像传输前, 必须在不严重降低图像质量的前提下, 采用合适的压缩编码算法对视频图像进行压缩。离散余弦变换 DCT (Discrete Cosine Transform) 被认为是优秀的视频图像压缩算法之一。首先, 以视频会议系统应用为背景, 简单介绍了 DCT 和反向离散余弦变换 IDCT (Inverse-DCT) 算法、公式及其相关; 其次, 以 VC++ 为工具, 以个人笔记本电脑为平台, 实现了 DCT 和 IDCT 算法, 给出并且分析了部分应用程序代码; 最后, 做了视频图像压缩的仿真实验。

关键词: 视频会议系统; 视频图像压缩; 离散余弦变换; 反向离散余弦变换

中图分类号: TP312

文献标识码: A

文章编号: 1674-7720(2013)09-0005-03

Research and implementation of algorithm for DCT and IDCT based on video-conference system

Yan Jing, Song Wen'ai

(Key Laboratory of Instrumentation Science & Dynamic Measurement of Ministry of Education, NUC, Taiyuan 030051, China)

Abstract: The video image compression is one of the key technologies which decides whether the video-conference system can provide good performance. In order to maximize the utilization of limited network bandwidth, we should adopt an appropriate algorithm of compression code to compress video image on condition that the image quality can not be degraded severely before video image transmission. DCT (Discrete Cosine Transform) is considered as one of the outstanding algorithms for video image compression. Firstly, it made a brief introduction of the algorithm and formula for DCT and IDCT (Inverse-DCT) and their related information with the application of video-conference system as the background. Secondly, it implemented the DCT algorithm and IDCT algorithm by using VC++ as the tool, with a personal notebook computer as a platform, and it gave a part of program code used with introductions and made simulation experiment of video image compression which gave some useful datum. Lastly, it draws a conclusion that the DCT and IDCT have important practical significance to realization, development and innovation of video-conference system and its technique in engineering.

Key words: video-conference system; image compression of video; DCT; Inverse-DCT

视频会议系统是网络技术、计算机技术和数字信号处理技术相互融合的产物, 它为处于不同地点的用户提供一个虚拟的环境, 用户可以通过视频、音频等多媒体手段进行交流和协作^[1-2], 而视频图像压缩是视频会议系统能否提供良好性能的关键技术之一。

视频图像压缩的目的就是要减少原始数据的数据量, 即以尽量少的比特数表示原来的图像, 利于存储和

传输, 同时又要能保持复原图像的质量, 满足规定的要求。DCT 是视频会议系统理想的视频图像压缩方法之一。与之相对应, 在接收端要对 DCT 后的图像数据进行 IDCT (不影响视频会议进行时) 以近似恢复原始图像。以软件视频会议系统 (即 PC+Internet 模式) 的应用为研究背景, 进行了 DCT 和 IDCT 算法的研究和实现, 这对于视频会议系统及其技术的发展、创新有重要的实际意义。

1 VC++实现 DCT 和 IDCT

已有的各种成熟的视频压缩标准如 JPEG、MPEG、H.26X 以及 HDTV 等都无一例外地采用基于 DCT/IDCT 的压缩编码^[3]。DCT 的基本思路是:首先在时间域将原始图像分成 N (水平像素数) $\times N$ (垂直像素数)像素子块,根据需要可选择 4×4 、 8×8 、 16×16 等块,这些取样的像素块代表了原图像帧各像素的灰度值,然后对 $N\times N$ 像素块逐一进行单独的 DCT,以便将取样块由时间域转换为频率域的 DCT 系数块,利于后续的量化及编码,实现图像信息的压缩。

接下来所做的基于视频会议系统的视频图像压缩算法研究、实现及仿真实验分别选取 $N=4$ 、 8 和 16 等。本文以 VC++实现 DCT 和 IDCT 算法所采用的软硬件环境如下:

(1)硬件环境:个人笔记本电脑(华硕)。配置如下:

①CPU: Intel Core 2 i5-2450 M(主频:2.5 GHz,双核);

②内存:三星(容量:4 GB,类型:DDR3 1 333 MHz);

③独立显卡:显卡芯片 NVIDIA GeForce GT 520M,显存容量/类型/位宽:1 GB/GDDR3/64 bit。

(2)软件环境:Windows 7, Visual C++ 6.0。

2 离散余弦变换(DCT)

视频压缩编码算法中帧内和帧间编码的核心算法都是 DCT,它有效地削减了视频序列中帧内图像和预测残差信号空域冗余,且它的性能最接近理论上最优的 K-L 变换^[4],有相对简单的实现方法,因而有着广泛的应用。二维 $N\times N$ 点 DCT(离散余弦正变换)公式为^[5-7]:

$$F(u, v) = \frac{2}{N} C(u) C(v) \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x, y) \cos \frac{(2x+1)u\pi}{2N} \cos \frac{(2y+1)v\pi}{2N} \quad (1)$$

在式(1)中:

(1)空域变量取值范围为 $x, y=0, 1, \dots, N-1$; 频域变量取值范围为 $u, v=0, 1, \dots, N-1$;

(2)当 $u=0, v=0$ 时, $C(u)=C(v)=1/\sqrt{2}$; 当 $u, v=1, 2, \dots, N-1$ 时, $C(u)=C(v)=1$;

(3) $f(x, y)$ 是 $N\times N$ 像素子块中某点的像素值, $F(u, v)$ 是 $N\times N$ 像素子块在频域中某点的 DCT 系数值;

(4) $F(u, v) = \frac{1}{N} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x, y)$, 为 $N\times N$ 像素子块 DCT 后的直流(DC)系数值。

以 VC++为工具, C++实现 DCT 算法, 所用部分程序代码如下:

```
for (u=0; u<rows; u++) {
    for (v=0; v<cols; v++) {
        for (x=0; x<rows; x++) {
            for (y=0; y<cols; y++) {
                f[u*cols+v]+=
```

```
temp[x*cols+y]*cos((2*x+1)*u*PI/
(2*cols))*cos((2*y+1)*v*PI/(2*rows));}
if (u==0&&v==0) f[u*cols+v]=one_div_sqrtcolrow*f
[u*cols+v];
else if (v!=0||u!=0) f[u*cols+v]=sqrt_two_colrow*f
[u*cols+v];
else f[u*cols+v]=two_div_sqrtcolrow*f[u*cols+v];
}
}
```

其中:

(1) $N\times N$ 像素子块中, rows 为水平像素数; cols 为垂直像素数, 分别取 4、8 及 16 等;

(2) #define PI 3.1415926;

(3) double one_div_sqrtcolrow=1.0/sqrt((double)(cols*rows));

double sqrt_two_colrow=sqrt(2.0/(cols*rows));

double two_div_sqrtcolrow=2.0/sqrt((double)(cols*rows));

原始时域 256 级(像素取值范围 0~255), 灰度图像如图 1 所示。分辨率大小为: 256×256 。以 VC++为工具, 对原始时域图像进行离散余弦变换, 并对得到的 DCT 系数进行量化取整, 然后绘制出频域图像, 分辨率大小为: 256×256 。如图 2(a)、(b)所示。

仿真实验选取 $N=128$ 和 $N=256$ (即以 1/4 帧和整帧原始灰度图像为像素子块) 作为特例实现了 DCT, 分辨率大小为: 256×256 。频域图像如图 2(c)、(d)所示。



图 1 原始时域图像

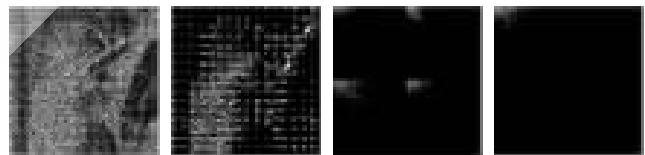


图 2 DCT 后频域图像

DCT 是一种实数域空间变换, 不能直接对图像产生压缩作用。它的最大特点是对于一般的图像都能够将像素块的能量集中于像素块左上角的少数低频 DCT 系数上, 这样就可能只编码和传输少数系数而不严重影响图像质量, 为压缩打下了基础。图 2(c)、(d) 的白点区域即为低频 DCT 系数区域, 也是像素块的能量集中区。在图 2(a)、(b)、(c)、(d) 中白点区域图像的像素值在 200~255 之间, 其余黑色区域图像的像素值基本为 0。参考图 1, 由图 2(b)、(c) 可知原始时域图像在较暗(灰度值较小)的部分含有较多的能量。

3 反向离散余弦变换 (Inverse-DCT)

二维 $N \times N$ 点 IDCT (离散余弦逆变换) 公式为^[5-7]:

$$f(x,y) = \frac{2}{N} \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} C(u)C(v)F(u,v) \cos \frac{(2x+1)u\pi}{2N} \cos \frac{(2y+1)v\pi}{2N} \quad (2)$$

在公式(2)中:

(1) 空域变量取值范围为 $x, y=0, 1, \dots, N-1$; 频域变量取值范围为 $u, v=0, 1, \dots, N-1$;

(2) 当 $u=0, v=0$ 时, $C(u)=C(v)=1/\sqrt{2}$; 当 $u, v=1, 2, \dots, N-1$ 时, $C(u)=C(v)=1$;

(3) $f(x,y)$ 是 $N \times N$ 像素子块中某点的像素值, 是 $N \times N$ 像素子块在频域中某点的 DCT 系数值。

以 VC++ 为工具, C++ 实现 IDCT 算法, 所用部分程序代码如下:

```
for (x=0; x<cols; x++){
for (y=0; y<rows; y++){
for (u=0; u<cols; u++){
for (v=0; v<rows; v++){
if (u==0&&v==0)
F[x*rows+y]+=one_div_sqrtcolrow*temp
[u*rows+v]*
cos((2*x+1)*u*PI/(2*
cols))*cos((2*y+1)*v*PI/(2*rows));
else if (u==0 || v==0)
F[x*rows+y]+=sqrt_two_colrow*temp
[u*rows+v]*
cos((2*x+1)*u*PI/(2*cols))*cos((2*y+1)*v
*PI/(2*rows));
else
F[x*rows+y]+=two_div_sqrtcolrow*temp[u*
rows+v]*
cos((2*x+1)*u*PI/(2*cols))*cos((2*y+1)*v
*PI/(2*rows));
}
}
}
}
```

其中:

(1) 与 DCT 对应, 在 IDCT 中: rows 代表行元素的个数, cols 代表列元素的个数;

(2) #define PI 3.141 592 6;

(3) double one_div_sqrtcolrow=1.0/sqrt ((double) (cols*rows));

double sqrt_two_colrow=sqrt (2.0/(cols*rows));

double two_div_sqrtcolrow =2.0/sqrt ((double) (cols*rows));

以 VC++ 为工具, 对频率域图像数据 DCT 后的系数

进行 IDCT, 可以近似恢复出时域图像数据, 并且用得到的图像数据绘制出时域图像, 分辨率大小为: 256×256, 如图 3 所示。经过变换以后恢复的图像与原图像相比较, 在肉眼上看不出差别, 相当逼真。但当 $N=16$ 时, 恢复出的图像(图 3)有较强烈的方块效应。



图 3 对应图 2(b), IDCT 后时域图像

4 实验总结

在 DCT 及 IDCT 测试实验中, 像素子块 N 的取值分别选为 4、8、16、256, 测试项目包括 DCT 及 IDCT 算法耗时、第一像素子块的直流 (DC) 系数值和均方误差^[8] (MSE), 实验结果如表 1 所示。

表 1 实验结果

| 像素子块 N 的取值 | 算法耗时/ms | | 第一像素子块的 直流 (DC) 系数值 | 均方误差 (MSE) |
|-----------------|----------|----------|------------------------|---------------|
| | DCT | IDCT | | |
| 4 | 124.67 | 140.33 | 538.000 | 2.292 |
| 8 | 441.67 | 457.33 | 1 055.750 | 3.819 |
| 16 | 1 736.67 | 1 742.33 | 2 101.250 | 5.444 |
| 256 | 437 822 | 440 417 | 25 261.871 | 13.157 |

随着 N 取值的增大, DCT 及 IDCT 算法耗时及均方误差也在增加, 本文所做的是浮点 DCT 和浮点 IDCT, 均方误差反映的是 IDCT 后的时域图像数据与原始时域图像数据的误差, N 越大, 变换前后的误差也越大。

当 $N=4$ 时, 算法耗时最短, 即算法复杂度最小, 但由图 2 可知存在传输数据较多的问题, 因此为提高压缩率和传输效率, 在选取 N 时还应该综合考虑计算机软硬件配置及网络传输环境。

DCT 变换的特点是变换后图像大部分能量集中在左上角, 即直流系数 (DC) 附近, 因为左上角反应原图像低频部分数据, 右下角反应原图像高频部分数据, 而图像的能量通常集中在低频部分, 这由表 2、表 3 可知。当 $N=8$ 时, 选取原始图像第一个 8×8 像素子块的数据如表 2 所示, DCT 后的频域数据如表 3 所示, 左上角第一个数据为直流系数: 1 055.75, 是 64 个数据中最大的一个。

表 2 原始时域第一像素子块数据

| | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|
| 137 | 136 | 133 | 136 | 138 | 134 | 134 | 132 |
| 137 | 136 | 133 | 136 | 138 | 134 | 134 | 132 |
| 138 | 133 | 134 | 134 | 136 | 132 | 130 | 130 |
| 133 | 133 | 133 | 130 | 134 | 133 | 128 | 125 |
| 129 | 133 | 130 | 130 | 133 | 131 | 132 | 128 |
| 131 | 133 | 130 | 122 | 132 | 131 | 130 | 130 |
| 131 | 130 | 130 | 130 | 132 | 131 | 128 | 130 |
| 131 | 132 | 130 | 130 | 131 | 131 | 130 | 128 |

表 3 DCT 后的系数值 (对应表 2)

| | | | | | | | |
|----------|-------|-------|-------|-------|-------|-------|-------|
| 1 055.75 | 7.20 | -3.13 | 7.26 | 0.25 | -4.28 | -1.68 | 3.06 |
| 14.85 | 2.24 | -1.68 | 0.88 | 2.59 | 1.45 | -1.16 | -0.01 |
| 5.66 | -0.71 | 0.41 | -0.52 | 2.10 | 1.17 | -0.79 | -1.29 |
| -1.02 | -2.66 | 1.87 | 0.03 | -1.39 | -1.90 | -1.22 | 0.77 |
| -1.25 | 0.25 | -1.89 | 0.04 | -1.59 | 0.71 | -1.32 | -0.90 |
| -1.55 | 1.39 | -0.81 | 0.27 | 0.92 | -0.07 | 2.20 | 0.002 |
| -0.72 | 0.61 | 2.04 | 0.48 | -0.49 | -0.73 | 0.66 | 1.34 |
| 1.14 | -0.67 | -0.84 | -1.28 | 1.83 | 1.54 | -0.68 | -0.78 |

将表 3 (DCT 后的系数值) 中的数据进行量化、取整等操作得表 4, 方法为:

```
unsigned char tempchar;
tempchar=int(src[i*cx+j]/len*255);
```

len 是表 3 中 64 个数据中的最大值与最小值差值的绝对值。

将表 3 中的数据进行 IDCT, 得表 5 中的数据, 对比表 2 和表 5 可知变换前后的图像数据存在一定的误差, 而且随着 N 的增大, 误差也越大。

表 4 DCT 系数量化、取整等后的数值(对应表 3)

| | | | | | | | |
|-----|---|---|---|---|-----|---|---|
| 253 | 1 | 0 | 1 | 0 | 255 | 0 | 0 |
| 3 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

表 5 IDCT 后的数据(对应表 3)

| | | | | | | | |
|-------|-------|-------|-------|-------|-------|-------|-------|
| 136.8 | 136.1 | 133.5 | 135.4 | 137.8 | 134.3 | 133.9 | 132.1 |
| 136.8 | 136.1 | 133.5 | 135.4 | 137.8 | 134.3 | 133.9 | 132.1 |
| 137.1 | 133.5 | 133.7 | 133.5 | 135.9 | 132.4 | 130.6 | 130.2 |
| 132.9 | 132.8 | 132.4 | 130.0 | 133.8 | 132.5 | 128.6 | 126.0 |
| 129.9 | 132.7 | 130.1 | 129.9 | 133.0 | 131.0 | 131.3 | 128.0 |
| 131.1 | 132.5 | 129.9 | 124.0 | 132.0 | 130.7 | 129.6 | 129.2 |
| 131.2 | 130.4 | 130.0 | 129.8 | 132.2 | 130.8 | 128.3 | 129.3 |
| 131.2 | 131.9 | 130.0 | 129.8 | 131.5 | 130.9 | 129.8 | 127.9 |

本文所做的 DCT 和 IDCT 算法的研究和实现是以软件视频会议系统的应用为背景的, 是视频会议系统视频图像压缩的一个仿真实验, 离具体的实际工程应用还有

不小的差距。但对于视频会议系统及其技术的实现、发展和创新有着重要的实际意义。

参考文献

- [1] 陈婷. 视频会议系统技术浅析及其应用[J]. 科技信息, 2009(30): 222-223.
- [2] 杨高波, 韩龙. 视频会议系统综述[J]. 电视技术, 2004(9): 60-62.
- [3] 山洪刚, 郑南宁, 杨国安, 等. 一种应用于 8×8 二维 DCT/IDCT 的高效结构[J]. 半导体技术, 2002, 27(6): 13-17.
- [4] 韦长江, 郝鹏威, 石青云. 基于整型 DCT 变换的图像编码研究[J]. 中国图象图形学报, 2002, 7(3): 287-291.
- [5] 陈宝远, 吴孟泽, 张清喜, 等. H.264 标准中整数 DCT 图像压缩算法的优化及实现[J]. 哈尔滨理工大学学报, 2009, 14(5): 27-31.
- [6] 马媛媛, 杨峰, 信科, 等. 基于 DCT 的 JPEG 图像压缩的研究[J]. 计算机技术与发展, 2011, 21(8): 133-136.
- [7] 王军敏, 杜豪杰. 基于 DCT 的图像压缩算法及快速实现[J]. 平顶山学院学报, 2012, 27(2): 65-69.
- [8] 禹蒲阳. 基于 DCT 和 RLE 的灰度图像压缩方法[J]. 湖南科技大学学报(自然科学版), 2010, 25(2): 89-92.

(收稿日期: 2013-01-10)

作者简介:

闫晶, 男, 1982 年生, 硕士研究生, 主要研究方向: 信号与信息处理(图形图像处理方向)。

宋文爱, 女, 1964 年生, 教授, 主要研究方向: 测试计量及仪器、无损检测以及信号与信息处理。