第 30 卷　第 4 期　吉首大学学报(自然科学版)　Vol. 30　No. 4
2009 年 7 月　Journal of Jishou University (Natural Science Edition)　Jul. 2009

文章编号: 1007- 2985(2009)04- 0044- 04

基于扩展傅里叶变换的图像压缩*

李　涛, 张　波
(华南农业大学现代教育技术中心, 广东 广州　510642)

摘　要: 介绍了傅里叶变换的原理, 分析了傅里叶变换在声音和视频压缩方面的应用, 并且指出傅里叶变换在数字图像处理中的重要地位和作用.

关键词: 傅里叶变换; 信号; 图像压缩
中图分类号: TP342　文献标识码: A

首先从电子工程和符号处理的角度来讨论傅里叶变换的背景. 实现一个离散或连续信号从它的时间域描述到频率域描述的转换, 是傅里叶变换的一个基本性质, 这个基本性质被用来描述和分析不同频率对一个信号的影响. 更深一步, 将介绍傅里叶变换在图像压缩方面的应用.

1　连续和离散傅里叶变换

定义 1　一个连续信号是一个函数 $f: D \to \mathbf{R}^n$, 这里 $D \subseteq \mathbf{R}^m$ 并且 $m, n \in \mathbf{N}$.

一个离散信号是一个函数 $f: D \to \mathbf{R}^n$, 这里 $D \subseteq \mathbf{R}^m$ 并且 $m, n \in \mathbf{N}$.

另外, 如果 $f(D) \subseteq \mathbf{Z}^n$, 那么 f 被称作数字信号.

声音是有不同节拍和不同范围响度的连续信号. 它是信号 $f: \mathbf{R} \to \mathbf{R}$ 的一个实例. 在屏幕是灰色比例像素的情况下, 这个信号与每个点的强度值相关, 所以 $f: D \subseteq \mathbf{Z}^2 \to \mathbf{R}$. 当颜色是由 3 种基本颜色要素(RGB) 或 4 种基本颜色(CMYK) 表示时, 就在 $\mathbf{R}^3$ 或 $\mathbf{R}^4$ 上各自得到一个信号映射.[1-2]

一个离散的信号经常通过对一个连续的信号进行间隔取样而得到. 离散信号在例如生物工程、地震学、声学、声纳和雷达图像、语音通讯、数据通讯、电视卫星通讯、卫星图像和其他更多领域得到应用. 语音和电话信号是只有一维域信号的例子, 然而雷达图像、卫星图像是经过二维域处理的例子. 当对更复杂的问题建模时, 例如那些出现在地震学中的问题, 使用的域可能有很多维.

对信号进行一定的操作处理是重要的, 这些操作从信号中提取相关信息, 或者对信号进行转换, 以使从信号得出的信息更易于使用. 例如, 某人也许希望从一张心电图或脑电图的危险警报数据中提取一些重要的参数; 某人也许想对在电话信号中包含的数据进行压缩或识别(recognize) 与语音信号有关的话语.

一个通常的问题是从与例如电视传播或卫星图像相关的数据堆中提取有关信息. 另一个在信号传送中对信号进行处理的应用是设法消除由传播噪音、声音衰减或信道失真导致的信号冲突.

一个特别重要的例子是正弦(sine) 信号 $f: \mathbf{R} \to \mathbf{R}$, 其中 $f(t) = \sin t$, 而且它的复变量(complex variant)

* 收稿日期: 2009- 04- 14
基金项目: 广东省科技计划项目(0711150100136)
作者简介: 李　涛(1978-), 男, 湖北黄冈人, 华南农业大学现代教育技术中心实验师, 硕士研究生, 主要从事计算机网络与数据库技术、图像处理研究.

$f: \mathbf{R} \to \mathbf{R}^2 \approx \mathbf{C}, f(t) = e^{it} = \cos t + i\sin t$, 这里 $i = \sqrt{-1}$. 更一般地有信号 $f: \mathbf{R} \to \mathbf{C}, f(t) = a \cdot e^{ikt}$, a 为振幅(amplitude), 与信号的强度(intensity) 和频率(frequency) k 相关. 所有这些信号都是周期性信号的例子: 这里存在一个周期 $T \in \mathbf{R}_{>0}$, 对于所有 $t \in \mathbf{R}$, 有 $f(t+T) = f(t)$. 假设 $T = 2\pi$, 应用变换 $t \to 2\pi t/T$. 对于正弦(sinusoidal) 信号 $f(t) = a \cdot e^{ik}$, 最小周期 T 就是波长并且与通过 $T = 2\pi/K$ 算出的频率 k 相关. 同样, 一个离散信号也是周期性的, 对于所有的 $n \in \mathbf{Z}$, 若 $f(n+N) = f(n)$, 则周期为 $N \in \mathbf{N}_{>0}$.

另一个特别的以 2π 为周期的函数的例子是通过在一个有界限的时间间隔上定义的函数得到的, 将间隔标准化为 $[0, 2\pi]$, 并且使用周期率来扩展函数. 接下来, 在没有深入细究的情况下假设函数是足够光滑的, 这样就可以完美地定义所需要的积分[3-4].

定义 2　以 2π 为周期的信号 $f: \mathbf{R} \to \mathbf{C}$ 的(连续) 傅里叶变换为

$$f: \mathbf{Z} \to \mathbf{C}, f(k) = \int_0^{2\pi} f(t) e^{ikt} \, dt \qquad k \in \mathbf{Z},$$

其中 $i = \sqrt{-1} \in \mathbf{C}$.

接下来的转置公式根据函数 f 的傅里叶变换表示出了函数 f:

$$f(t) = \frac{1}{2\pi} \sum_{k \in \mathbf{Z}} f(k) e^{ikt}. \tag{1}$$

这对于所有的 $t \in \mathbf{R}$ 都有效; 对于 f 来说整个系列是均匀的聚合. 这个系列被称作 f 的傅里叶系列, 并且数字 $\beta_k = \frac{1}{2\pi} f(k) = \frac{1}{2\pi} \int_0^{2\pi} f(t) e^{-ikt} \, dt \ (k \in \mathbf{Z})$ 是 f 的傅里叶系数. 这个转置公式表明函数 f 是由它的傅里叶系数 $(\beta_k)_{k \in \mathbf{Z}}$ 的系列唯一决定的. 特殊函数 e^{ikt} $(k \in \mathbf{Z})$ 对于所有的复数矢量空间的 2π 周期函数是一个基础; 然而, 通常存在许多无限的非 0 系数.

通过给每个时间 $t \in [0, 2\pi]$ 的信号分配值 $f(t)$, 原始信号 f 在时间域上被描述, 而傅里叶变换 f 是对 f 在频率域上同意义的描述. 它将每个频率 k 和频率 k 引起的 $f(k)$ 联系起来, 也就是信号 $\exp(ikt)/2\pi$, f 由转置公式(1) 给出. 它将许多连续的 $f(t)$ 值压缩成许多具有可数性的值 $f(k)$. 对于 $k \in \mathbf{N}_{>0}$, 信号 $(f(k)\exp(ikt) + f(-k)\exp(-ikt))/2\pi$ 称作 f 的 k 次谐波(谐函数).

例 1　考虑由 $f(t) = \sin(t) + \sin(10t)/10$ 定义的以 2π 为周期的信号 $f: \mathbf{R} \to \mathbf{R}$, 可以认为 f 是由拥有大振幅的 $\sin(t)$ 的低频部分加上拥有小振幅的 $\sin(10t)/10$ 的高频部分所组成; 傅里叶变换将 f 分解成它的谐函数:

$$f(1) = -2\pi i = -f(-1), f(10) = -\frac{1}{5}\pi i = -f(-10),$$

而且在 $k \neq \pm 1, \pm 10$ 时 $f(k) = 0$. 因而 f 的第 1 个谐函数是 $(f(1)\exp(it) + f(-1)\exp(-it))/2\pi = \sin(t)$, 第 10 个谐函数是 $\sin(10t)/10$, 所有其他谐函数是 0.

离散傅里叶变换是连续傅里叶变换对于离散周期信号的类推.[5-6] 如果 $f: \mathbf{Z} \to \mathbf{C}$ 是一个周期为 $N \in \mathbf{N}_{>0}$ 的离散信号, 那么它的离散傅里叶变换 $f: \mathbf{Z} \to \mathbf{C}$ 由以下等式定义:

$$f(k) = \sum_{0 \leqslant n < N} f(n) e^{-2\pi ikn/N} = \sum_{0 \leqslant n < N} f(n) \omega^{kn} \qquad k \in \mathbf{Z},$$

这里 $\omega = e^{-2\pi ikn/N} \in \mathbf{C}$ 是一个 N 次本原单位根. 如果以一个连续的 2π 周期信号 g 开始并且在一般间隔点 $2\pi n/N$ $(0 \leqslant n < N)$ 来对它进行取样得到 f, 那么 f 的离散傅里叶变换仅仅是对于 g 的连续傅里叶变换积分定义的一般近似. 相对于连续的情况, f 又一次具有周期为 N 的周期性. 转置公式(1) 的类推为

$$f(n) = \frac{1}{N} \sum_{0 \leqslant n < N} f(k) e^{2\pi ikn/N} = \frac{1}{N} \sum_{0 \leqslant n < N} f(k) \omega^{-kn}. \tag{2}$$

如果联系一个次数不高于 n 的多项式 $f \in C[x]$, 一个周期为 n 的离散信号 $g: \mathbf{Z} \to \mathbf{C}$ 与 f 的系数形成一个映射 $\{0, \ldots, n-1\}$, 那么 $\mathrm{DFT}_\omega(f)$ $(\omega = \exp(-2\pi i/n))$ 是 g 的离散傅里叶变换.

在数字信号处理中, 例如声音和图像之类的连续信号要进行取样. 在数字硬件或软件系统上对于相应离散信号的分析和修正(例如噪音减少或对比扩大) 经常涉及大量数据的离散傅里叶变换, 这时, 快速傅里叶变换的使用是必不可少的.

2 傅里叶变换在声音和图像压缩方面的应用

傅里叶变换和它的逆可以被认为是在时间域和频率域之间表示方法的转换. 如果f 是一个(连续或离散)信号,$\hat{f}$ 是它的傅里叶变换, 那么$f(t)$ 是信号在时间t 处的值, 而$\hat{f}(k)$ 是频率k 对信号影响的结果. 两者都是一个信号等价的描述. 人们在接受声音和视频数据时, 通常倾向于在时间域的“缓慢”上来区别不同, 所以它们的高频率部分所起的作用是很小的. 这意味着在频率域上$|\hat{f}(k)|$ 的小值对应着k 的大值. 数据压缩的思想现在就是要抛弃$\hat{f}(k)$ 中“接近于0的值”并且仅仅存储剩下的部分. 相对于高频部分, 人类的耳朵和眼睛更乐意接受低频部分的错误, 听者和观察者很难注意到这种低频信息的丢失, 这样就可以得到相当可观的压缩比.

声音和视频数据是真值信号, 并且傅里叶变换的一个变量经常被用于再次映射真值信号到真值信号, 使$f:\{0, \cdots, N-1\} \to \mathbf{R}$是一个有限持续时间为$N \in \mathbf{N}$的离散真值信号, 例如经过取样的一段音乐或一行数字屏幕图像. 可以认为通过$f(n) = f(n \text{ rem } N)$ $(n \in \mathbf{Z})$, f 被扩展成了$\mathbf{Z}$上的一个周期信号. 如果使

$$\mathrm{DCT}(f)(k) = \frac{1}{\sqrt{N}} c(k) \sum_{0 \leqslant n < N} f(n) \cos \frac{\pi k(2n+1)}{2N} \qquad 0 \leqslant n < N,$$

$$\mathrm{IDCT}(f)(n) = \frac{1}{\sqrt{N}} \sum_{0 \leqslant k < N} c(k) f(k) \cos \frac{\pi k(2n+1)}{2N} \qquad 0 \leqslant n < N,$$

这里若$k = 0$则$c(k) = 1$, 否则$c(k) = \sqrt{2}$, 那么DCT 和IDCT 是有限持续时间N 上的真值信号到同类信号映射的逆操作. DCT(f) 是f 的离散余弦变换. 进一步研究显示, 计算这个变换或它的逆可以被减少到计算一个离散的傅里叶变换, 如果N 是2的幂次, 那么这个过程可以通过FFT 有效的完成, 在$\mathbf{R}$上要进行$O(N \log N)$ 次运算.

转置公式

$$f(n) = (\mathrm{IDCT} \circ \mathrm{DCT})(f)(n) = \frac{1}{\sqrt{N}} \sum_{0 \leqslant k < N} c(k) \mathrm{DCT}(f)(k) \cos \frac{\pi k(2n+1)}{2N} \qquad 0 \leqslant n < N,$$

显示了离散余弦变换导致原始信号f 作为一个周期信号γ_k 的线形混合体的表达式, 这里$\gamma_k(n) = \cos(\pi k(2n+1)/2N)$, 而且带有余数$c(k)\mathrm{DCT}(f)(k)/\sqrt{N}$.

一个可行的图像数据压缩算法按照如下过程工作. 对图像的每一行$f:\{0, \cdots, N-1\} \to \mathbf{R}$, 这里$f(n)$是那一行第$n$个像素的亮度, 在$0 \leqslant k < N$ 时计算DCT$(f)(k)$ (如果图像是彩色的, 那么, 例如对3种颜色红、绿、蓝的强度分别应用这个公式). 再选择一个量化参数$q \in \mathbf{R}_{\geqslant 1}$, 用$q$ 来划分DCT(f) 所有的值, 再近似约化为最接近的整数. 量子化的效果是选择了在绝对值上最接近于0的DCT(f) 的那些值(这些绝对值对于大k 来说, 一般是高频部分的情况). 因此q 的大值对应一个高的压缩率, 但是也对应比较差的图像质量. 最后, 一种对遗失数据进行处理的压缩技术的结合, 比如运行长度编码和Huffman编码, 被应用于已经被量子化的DCT(f) 值, 运行长度编码压缩每个连续的零系列为2个整数: 一个零标识运行的位置, 接着是运行的长度. 例如, 系列

$$1, 2, 3, 0, 0, 0, 0, 4, 0, 5, -6, 0, 0, 0, 1, 2$$

被压缩成

$$1, 2, 3, 0, 4, 4, 0, 1, 5, -6, 0, 3, 1, 2.$$

这样就减少了2个长度. 但是当0的出现次数为1的时候, 压缩后的数据实际上增加了数据的大小, 如上面的例子4和5之间只有单个0的情况就是这样. 为了重组图像, 在相反的方向上进行这项工作: 在对压缩数据解码后, 将这些数据乘以量化参数q 并且一行一行地应用离散余弦变换.

上面陈述的方法有一个缺点就是量子化会在图像的整个行上导致混乱, 这样一幅图像的局部结构(例如图像在某处亮度上的缓慢变化)就不能被量子化涉及到. 这个缺陷可以通过将图像划分为很多固定大小的更小的块来弥补, 再对每个单独的块分别使用上面介绍的压缩技术.

在JPEG 静态图像压缩标准中, 例如原始图像被划分为8×8 像素的正方形区域, 而且对每个正方形区域执行二维离散余弦变换(这是行式和列式一维离散余弦变换的混合体), 这样所有方行区域的DCT 系数就被量子化了, 然后运行长度编码和最终Huffman 编码. 一维行式离散余弦变换仅仅在水平独立性上

可以计算, 然而二维离散余弦变换同时适用于水平独立性和垂直独立性. 比起上面介绍的简单行式方法, 对图像进行分块处理并应用二维离散余弦变换, 就对图像的局部结构的变化提高了适应性, 从而可以显著地提高压缩率.

3　结语

傅里叶变换是数字图像处理技术的基础, 它通过在时空域和频率域来回切换图像, 对图像的信息特征进行分析和提取, 简化了计算工作量, 被喻为描述图像信息的第 2 种语言, 广泛应用于图像变换、图像编码与压缩、图像分割、图像重建中. 因此, 深入研究和掌握傅里叶变换及其扩展形式的特性, 是很有价值的.

参考文献:

[1] JOACHIM VON ZUR GATHEN, JURGEN GERHARD. Modern Computer Algebra [M]. Cambridge University Press, 1999.

[2] QUATIERI T F. Discrete-Time Speech Signal Processing: Principles and Practice [M]. Beijing: Publishing House of Electronics Industry, 2004.

[3] PFITZINGER H R. DFW-Based Spectral Smoothing for Concatenative Speech Synthesis [C]. Proc. ICSLP 2004, Korea, 2004: 1 397- 1 400.

[4] WOUTERS J, MACON M W. Control of Spectral Dynamics in Concatenative Speech Synthesis [J]. IEEE Trans. on Speech and Audio Processing, 2001, 9(1): 30- 38.

[5] KANG H, LIU Wen-ju. Sinusoidal All-Pole Modification Based Spectral Smoothing for Concatenative Speech Synthesis [C]. Natural Language Processing and Knowledge Engineering, IEEE NLP-KE' 05, 2005.

[6] WOUTERS J, MACON M W. Spectral Modification for Concatenative Speech Synthesis [C]. Acoustics, Speech, and Signal Processing 2000 ICASSP' 00, Proceedings 2000 IEEE International Conference, 2000, 2: 941- 944.

Image Compression Based on Extended Fourier Transform

LI Tao, ZHANG Bo

(Modern Education and Technology Center, South China Agriculture University, Guangzhou 510642, China)

Abstract: This paper explains the principium of Fourier transform, analyzes its application in audio and image compression, and presents its important status and function in numeral image disposal.

Key words: Fourier transform; signal; image compression

(责任编辑　向阳洁)