

情報システム工学科 平成18年度後期 「自主課題研究」

研究テーマ：音声圧縮データの特性調査

名列番号：017 小林 幸司

平成19年2月13日

1 研究課題

音声圧縮には、人間の聴覚特性である「最小可聴限界」と「マスキング効果」を利用し、聞き取りづらい部分の情報を減らすことで大幅にデータを小さくする非可逆圧縮を行うものが広く用いられている。本課題では、MP3（MPEG-1 Audio Layer-3）による非可逆圧縮によって、エンコード／デコード後のデータが元データと比較してどのように変化しているかを調査する。また、MP3のビットレートや音声データに含まれる音の種類によってどのような違いがでるかも調べる。

2 研究方法

音の種類が違うWAVEファイルをいくつか用意し、ビットレート別（32、64、128）にエンコードしてMP3ファイルを作る。MP3ファイルをデコードしてWAVEファイルに戻し、圧縮前のWAVEファイルと比較する。

＜比較方法＞

WAVEファイルの音声データ部分を読みとり、1024点毎にフーリエ変換して振幅スペクトルを求めるプログラムをC言語で作成した。このプログラムによって得られる周波数特性のグラフの見比べと、実際の耳での聞き比べによって比較を行う。

3 結果と考察

ビットレートを変化させたところ、32kbps、64kbpsでは人間の可聴範囲である1000Hz～20kHzの周波数成分を大幅に削っており、実際に耳で聞いて分かる程に音質は低

下していた。ビットレートによってかなり音質には差があるため、MP3によるエンコードの際にはビットレートの設定は重要である。音の種類による比較では、楽器の音のみのWAVEファイルより、楽器の音と歌声の入ったWAVEファイルの方が圧縮前後の差が大きいことがグラフからわかった。このことより、様々な種類の周波数成分を含む音声データの方が圧縮による影響を受けやすいと言える。

4 今後の課題と感想

非可逆圧縮によって音声データの周波数成分が削られていることがわかったが、実行結果のグラフを見る限りでは人間がほとんど聞き取れないはずの低周波成分、高周波成分がほぼそのまま残っており、MP3の圧縮の原理である人間の聴覚特性とは矛盾しているように見える。これについてはプログラムの見直しやMP3の仕組みに関するさらに深い調査が必要であると思われる。また、今回は3種類しか元データとなるWAVEファイルを用意していなかったもので、もっと様々な種類の音による比較ができればよかった。自分の情報通信やC言語に関する理解不足によって研究の進行が遅れたために課題が多く残る結果となったが、今回の研究で理解を深めることができたので有意義だったと思う。