

卒業論文要旨

ウェーブレット変換を用いた音声通信の セキュリティ向上システムに関する研究

(知能情報システム学) 津田裕司

1. 序論

音声通信では第三者に知られては支障がある内容を対象とする場合がある。そこで、第三者に情報を与えないように音声情報の暗号化などの処理がなされるが、その場合データ量の増加(通信負荷の増加)や、送受信双方での処理時間の増加が問題となってくる。

そこで本研究では、ウェーブレット変換の多重解像度近似を用いて音声データをコード化することでデータ量を減らし、かつ、第三者がその内容を理解することを困難とする手法を提案し、その性能を評価した。

2. 開発環境

- ・OS:Windows XP ・使用したプログラム言語: Visual C++6.0
- ・Dell Optiplex 755 CPU:Core2 Duo 2.66GHz ・メインメモリ:1.96GB

3. 方法

3.1. ウェーブレット変換によるコーディング

通常、音声データ A のスケーリング係数列 $S(A)_n = \{x_1, x_2, x_3, \dots, x_n\}$ のヒストグラムをとると 0 を中心とした分布となる。それを5つの領域に分割し、それぞれの領域に $0 \sim 4$ の記号を割り当てる。本稿でのコーディングは、スケーリング係数列 $S(A)_n$ の各要素 x_n が当てはまる領域の記号 X_n を配列 $C(A)_n$ に与え、 $C(A)_n = \{X_1, X_2, X_3, \dots, X_n\}$ とすることである。

3.2. コードと平均値を使った音声の作成

3.1における各領域の平均値を $a_0 \sim a_4$ とし、 $a_3 = 0$ とする。また、スケーリング係数をコーディングして作成した配列 $C(A)_n$ の各要素を、各々当てはまる領域の平均値 $a_0 \sim a_4$ に置換したものを $S(A)'_n$ とする。この $S(A)'_n$ と全ての要素を 0 にしたウェーブレット係数を用いて逆ウェーブレット変換を行い、音声データ A' を作成する。 A が人間の音声のデータならば、通常、 A' を用いて音声を再生すれば音声データ A の内容が聞き取れる。

3.3. コードの選択的置換による音声の作成

A とは異なる音声データ B のスケーリング係数列を $S(B)_n = \{y_1, y_2, y_3, \dots, y_n\}$ とする。上記の A と同様にコーディングして、 $C(B)_n = \{Y_1, Y_2, Y_3, \dots, Y_n\}$ 、および各領域の平均値 $b_0 \sim b_4$ を作成する。 $C(A)_n$ と $C(B)_n$ を1番目の要素から順に比べていき、 $X_n \neq Y_n$ ならば、該当する Y_n の値をすべて X_n の値に置換して、要素置換後の配列 $C'(B)_n$ を作成する。その結果、 $C'(B)_n = C(A)_n$ となる。 $C'(B)_n$ の各要素を、各々当てはまる領域の平均値 $b_0 \sim b_4$ に置換して $S(B)'_n$ を作成し、全ての要素が 0 であるウェーブレット係数を用いて逆ウェーブレット変換を行って、音声データ B' を作成する。 A が人間の音声のデータならば、通常、 B' を用いて音声を再

生すれば音声データ A の内容が聞き取れる。 A の領域の平均値 $a_0 \sim a_4$ を B の領域の平均値 $b_0 \sim b_4$ に置き換えて、音声データ A と同じように聞こえるのは、音声データ A と B の主たる違いが振幅のみに生じることによると考えられる。

4. コード変換を用いた音声通信システム

4.1 送受信に必要とするもの

3.3節記載の手法を用いることで、音声データ A のデータを持っていなくても、音声データ B のデータと置換するコードの情報があれば、 A の内容を理解できる音声データを作成することができる。音声データのやり取りを行う発信者と受信者がどちらも音声データ B を所持していると、発信者が3.3節記載の手法で作成する(1)「スケーリング係数列 $C(B)_n$ のうち、置換すべき要素番号を表す数列」と(2)「置換後のコード列」の2つを受信者に発信すると、受信者がそれらを用いて音声データ B のデータを置換してやれば、再生すれば音声データ A と似たように聞こえる音声データ B' を作成することができる。

4.2 送受信するデータ

上記の(2)「置換後のコード列」は、領域番号である $0 \sim 4$ を要素とする配列になる。この配列の要素数は、スケーリング係数列の要素数を n とすると、3.3節のスケーリング係数の置換方法からして、 $0.8n$ 程度となる。

(1)「スケーリング係数 $C(B)_n$ のうち、置換すべき要素番号を表す数列」に関しては置換すべき要素番号のスケーリング係数 Y_n の絶対値の整数部を 256 で割った余りを送信するようにした。これにより、音声データ B 以外のデータを用いて A の音声に近いものを作ることは困難になる。また、要素数は(2)と同じであるので、データ量は $0.8n \times 8 = 6.4n$ bit程度となる。ただし、この方法では誤った位置でスケーリング係数を置換してしまうことがあるが、これまで作成された音声では、その現象による大きな影響はでていない。

5. 高周波ノイズの除去

本法で送信したデータを基に受信者側で作成した音声には高周波ノイズが含まれる。このため、フーリエ変換などを用いた高周波ノイズ除去を施すと、音声により明瞭に聞き取れる。

6. 結論と課題

ウェーブレット変換を用いた音声通信セキュリティ向上システムを提案し、その有効性を確認した。ウェーブレット変換を施す際の展開レベルを低く設定すればノイズが少なくなり、明瞭な音声となった。これは置換するデータ量が増えるからである。しかしながら、送信データ量が増えてしまう。処理時間は展開レベル1では発信側の処理が約 195 ms、受信側の処理が約 294 ms となり、展開レベル3では発信側の処理が約 67 ms、受信側の処理が約 288 ms となった。

今後の課題としては、処理の高速化、領域 $0 \sim 4$ を決定するパラメータの最適化や、効率的なノイズ除去法の検討、誤置換の低減などが挙げられる。