

# RoboCup サッカーにおける秘匿通信のための スペクトル拡散を用いた音声電子透かし法の提案

## A Digital Speech Watermarking Method Using Spread Spectrum for Secret Communication in RoboCup Soccer

坪倉和哉<sup>1</sup> 久保谷空史<sup>1</sup> 館拓磨<sup>1</sup> 小林邦和<sup>1</sup>

Kazuya TSUBOKURA<sup>1</sup>, Takashi KUBOYA<sup>1</sup>, Takuma TACHI<sup>1</sup>, Kunikazu KOBAYASHI<sup>1</sup>

<sup>1</sup> 愛知県立大学 情報科学部

<sup>1</sup> School of Information Science and Technology, Aichi Prefectural University

**Abstract:** 現在 RoboCup の標準プラットフォームリーグ (SPL) では、ロボット間のコミュニケーション手段として、主に無線通信が用いられている。しかし、人間のサッカーにより近づけていくためには、無線通信に代わるコミュニケーション方法の開発が必要となる。本研究では、無線通信に代わるものとして、音声を用いたコミュニケーションについて検討する。具体的には、自然言語の合成音声に対し、スペクトル拡散法を用いて秘匿データを埋め込む手法を提案する。そして、ロボット間の距離と透かし情報の大きさを変更した場合でのビット誤り率の比較検証を行った。

## 1 はじめに

RoboCup では 2050 年までに、サッカーワールドカップの世界チャンピオンに勝てるロボットサッカーチームを作ることを目指している [1]。RoboCup サッカーの標準プラットフォームリーグ (Standard Platform League: SPL) では、全チームが同一のロボットを用いてソフトウェアシステムの優劣を競うリーグである [2]。現在は、SoftBank Robotics 社のヒューマノイドロボット NAO が用いられている。

現在 RoboCup では、ロボット間コミュニケーションに無線通信を用いることが主流である。しかし、将来人間と対戦する上で、より人間と同じ条件にするためには、無線通信に代わるコミュニケーション方法が必要となる。とりわけ、SPL では、無線通信におけるパケット量を制限したり、ホイッスルの音源定位に関するテクニカルチャレンジが行われたりするなど、無線通信ではなく音を利用したロボット間コミュニケーション技術の開発が求められている。

一般的に、音を利用したコミュニケーションには、モジュール信号のような符号的な音や自然言語の発話が考えられる。前者は、人間には識別が困難である。エンターテインメント性を考慮すると、人間のサッカーに見られるような、プレイヤーと観客とのインタラクションは生まれにくく、観客を置き去りにしてしまいかねない。一方、自然言語の発話によるコミュニケーションは、観客にもロボット同士のコミュニケーションが理解でき、ロボットが何を考えて動いているかが推測できる。しか

し、単位時間あたりに伝えられる情報量が少ないということや、相手チームにも対話内容が伝わってしまうという危険性がある。

本研究では、この問題を解決するために、自然言語の発話の中に、秘匿データを埋め込む手法を提案する。具体的には、スペクトル拡散法を用いて合成音声にビット列を埋め込む。これにより、観客は合成音声からロボット同士の対話を観察でき、ロボットは埋め込まれたビット列から命令や意思疎通を行うことができる。さらに、相手チームからはビット列が読み取られないよう、秘匿性を担保することも可能となった。

## 2 先行研究

### 2.1 スペクトル拡散

スペクトル拡散 (Spread Spectrum : SS) 方式とは、信号の変調方式のことであり、携帯電話や無線 LAN などの無線通信や、音響電子透かしなどで用いられている [3][4][5]。SS 方式では、通常の情報伝送に必要な帯域幅を大幅に超えたより広い帯域幅を持つ信号を用いて通信を行っており、秘匿性や耐ノイズ性といった通信を行うにあたって非常に優れた特性を有している。

SS 方式には、直接拡散 (Direct Sequence : DS)/SS 方式や周波数ホッピング (Frequency Hopping : FH)/SS 方式などがある。DS/SS 方式では、信号強度を常に弱い状態で保てるため、FH/SS 方式よりも秘匿性が高いと言える。本研究の趣旨として、秘匿性がある方が望ま

しいと考えられることから,DS/SS 方式を用いた音声通信について考察する.

図 1 に DS/SS 方式の送信機, 受信機の構成におけるデータの流れを, 図 2 に各状態でのデータの様子を示す. 最初に, データを BPSK(Binary Phase Shift Keying) や QPSK(Quadrature Phase Shift Keying) を用いて変調する. 高速通信には向かないものの, 伝送品質が劣化しにくいことから, 本研究では BPSK 変調を用いることとする. 次に, 拡散系列を用いてデータの拡散を行い, 送信機から受信機へ拡散したデータを送信する. 受信機では, 受信したデータに対し逆拡散を行うことで, 拡散を行う前のデータ得ることができる. 最後に, 復調器でデータの復調を行うことで, データの復元ができる.

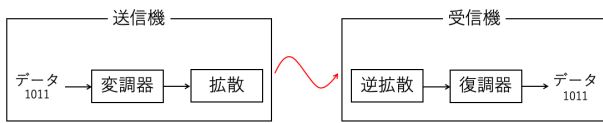


図 1: DS/SS 方式におけるデータの流れ

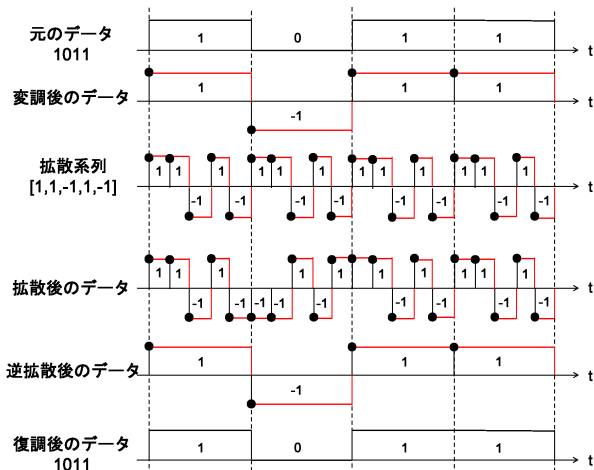


図 2: DS/SS 方式における各状態でのデータの様子

### 2.1.1 逆拡散

逆拡散では, 図 3 のように重み係数として拡散系列を与えた相関器を生成し, 受信信号との相関を求めている. 相関が高い程出力される値の絶対値が大きくなる. そのため, 図 4 のように相関器出力の各ピークを取り出すと拡散前の変調したデータを復元できる.

### 2.1.2 レイク合成

スペクトル拡散方式では, 伝送信号の波形が歪になる選択性フェージングを解決することは可能である. し

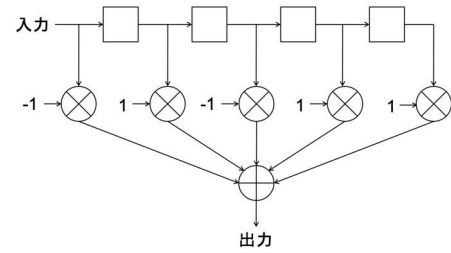


図 3: 相関器の入出力のイメージ

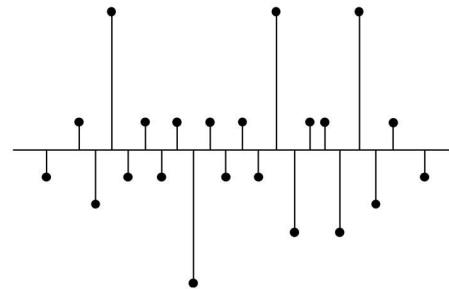


図 4: 相関器の出力例

かし, 伝送信号波形の振幅が小さくなるフラットフェージングの影響を受けてしまう. そのため, 必要となるのがレイク合成である [6]. レイク合成では, スペクトル拡散方式で分離されたマルチパス信号を SN 比が最大になるよう合成する. すると, 受信した伝送信号が大きくなり, フラットフェージングの影響を軽減できる. SN 比を最大にするには, 最大比合成を行う. 最大比合成とは, 各伝送経路に重み係数を掛け, 位相を揃えて合成する手法である. この重み係数をチャンネル係数と呼び, 1 を拡散して送信したものを受信機で相関器に通すことで得られる.

## 2.2 線形予測分析

線形予測分析とは, 現在の出力サンプル値  $x_t$  がそれ以前の  $N$  個のサンプル値の線形結合によって予測されるものとするモデルで, 声道フィルタの推定など音声分析に広く用いられている手法である [7][8].

声道フィルタ  $H(z)$  を,

$$H(z) = \frac{1}{1 + \sum_{n=1}^N \alpha_n z^{-n}}$$

と表される全極型のフィルタで近似することを考える.  $\alpha_n$  を線形予測係数と呼び, この値を調整することで, 声道フィルタを近似する.

### 3 提案手法

本研究では、送信したいビット列をスペクトル拡散させて、合成音声に埋め込んだ。この音声をロボットのスピーカから送信し、別のロボットで受信して得た音声からビット列を復元した。提案手法における音声生成からビット復元までの流れを図5に示す。この手法により、人間にも理解できる音声に、秘匿化された情報を埋め込むことができる。

#### 3.1 音声へのビット列の埋め込み

音声生成部分では、まず、送信したいビット列をBPSK変調する。例えば、送信したいデータが{1, 0, 1, 0}であれば、BPSKによる変調後は{1, -1, 1, -1}となる(0が-1に変換される)。これと、拡散系列をビット毎に掛け算し、拡散ビット列を得る。拡散ビット列の音圧レベルを調整するため、合成音声に重み係数を乗じてパワーを増幅した後、拡散ビット列と足し合わせ、再生する音声を得る。

#### 3.2 ビット列の復元

ビット復元部分では、まず、録音した音声にハイパスフィルタをかける。これにより、合成音声の影響を低減させるとともに、低周波ノイズが除去できる。次に、拡散ビット列を作成するときを使用した拡散系列を用いて、ハイパスフィルタをかけた録音音声の自己相関を求める。

実環境で録音すると、壁や障害物などに反射して遅延波が発生し、相関器の出力しまうため、レイク合成により、先行波と遅延波のピークの合成を行う。レイク合成後、BPSK復調を行い、ビット列を復元する。

## 4 実験

本研究では、前節の提案手法を用いて合成音にビット列を埋め込み、実環境にて送受信を行った。実験の評価のポイントは次の3点である。

- 送信側ロボットと受信側ロボット間の距離
- 合成音と埋め込みビット列のSN比 (Signal-to-Noise Ratio : SNR)
- ビット誤り率 (Bit Error Rate : BER)

### 4.1 実験条件

実験は、愛知県大学次世代ロボット研究所のアリーナ(室内)にて行った。実験中の環境音の大きさは40~45[dB]程度であり、換気扇の運転音のみが聞こえる状態であった。

音声の送信、受信には、現在、SPLで使用されているSoftBank Robotics社のNAO V6(図6, 7)を用いた。NAO V6のスペックを表1に示す[9]。送信には左右両方のスピーカから音声を出力し、ビット列の復元には4つのマイクのうち1つのマイクからの情報のみを使用した。なお、送信、受信時のサンプリング周波数は、NAOで録音可能な値である、48,000[Hz]とした。

項目	詳細
OS	NAOqi 2.8
高さ	57.4cm
スピーカ	左右 2つ
マイク	指向性マイク 4つ

音声を送信側NAOと受信側NAO間の距離は、1.0m, 3.0m, 5.0m, 7.0m, 9.0m, 10.8mの6パターンとし、それぞれのBERを比較した。NAOは、図8に示すように、それぞれが正面を向き合うように配置した。送信側NAOは直立しており、受信側NAOはしゃがんだ状態である。現在、SPLのフィールドの大きさは6.0m × 9.0mであり、10.8mまで誤りなく情報を伝送できれば、実用上、十分にフィールドをカバーできたといえる。

### 4.2 実験用音声の生成

まず、実験を行うために、送受信するための音声を作成した。合成音声は表2のように作成した。声帯音源は、基本周波数250[Hz]の音源に共振周波数1,000[Hz]の2次フィルタをかけ、ランダム雑音を足し合わせた。声道情報には、人間が発話した音声から線形予測分析(次数20)で抽出した線形予測係数を使用し、フレーム単位(フレーム長:1,024点)で声帯音源にフィルタリングした。また拡散ビット列は表3のように作成した。

項目	詳細
発話内容	“ガンバレーガンバレー”
発話時間	7.0[sec] (末尾0.2[sec]は無音区間)
サンプリング周波数	48,000[Hz]
基本周波数	250[Hz]

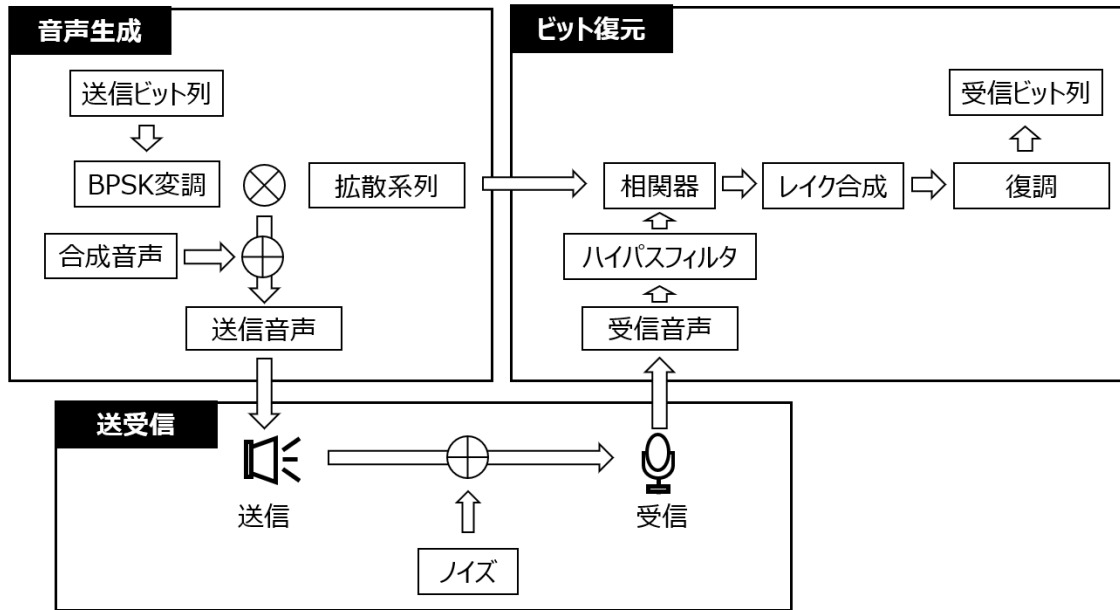


図 5: 提案手法における音声生成からビット復元までの流れ

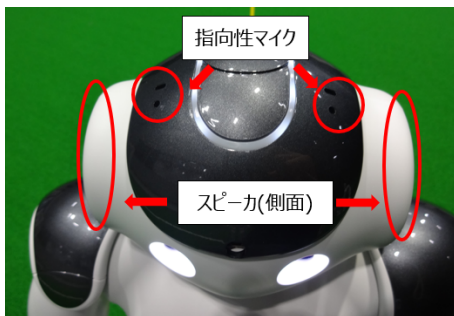


図 6: NAO の前頭部



図 7: NAO の後頭部

なお、拡散系列には、1 と-1 がランダムで現れる系列を使用した。

表 3: 作成した拡散ビット列

項目	詳細
送信ビット列	{1, 1, 0, 1, 0, 0, 1}
伝送速度	1[bps]
拡散系列長	48,000/bit
変調方式	BPSK

合成音声と拡散ビット列を足し合わせる際、そのまま足してしまうと、拡散ビット列の音の大きさが大きすぎるため、500、900、1,600 の 3 パターンで合成音声に重みづけを行った後、拡散ビット列と足し合わせて正規化をした。重み係数を乗じた合成音声と拡散ビット列との SNR を表 4 に示す。同表より、重み係数が大きくなると SNR も大きくなるので、生成音声のノイズ成分が

小さくなり、人間が聴取したときに合成音声聞き取りやすくなる。なお、SNR は合成音声を信号、拡散ビット列を雑音成分として算出した。

表 4: 重み係数と SNR の対応

重み係数	SNR[dB]
500	40.8
900	45.9
1,600	50.9

それぞれの重みで合成音声と足し合わせ、送受信したときの BER を観察した。なお、全体で 7 ビット送信することになるが、先頭 1 ビットはチャンネル係数を求めるためのテストデータとして用いるので、ビット誤り率は 6 ビット中の誤り率となる。



図 8: 実験中の NAO の配置

### 4.3 ビット列の復元

音声の送受信には NAO を用いたが、計算コストのため、ビット列の復元には計算機を使用した。ハイパスフィルタは、受信音声に表 5 の条件で処理を行った。レイク合成のチャンネル係数には、受信音声の初めの 1 秒の自己相関を用いた。

表 5: ハイパスフィルタの条件

項目	詳細
使用ソフトウェア	MATLAB R2019b
使用関数	highpass
通過帯域周波数	4,000[Hz]

### 4.4 結果と考察

合成音声の重み係数ごとに、ロボット間距離を変更したときの BER の推移を図 9 に示す。ただし、BER は、重み係数とロボット間距離の組合せ毎に 5 回送受信を行ったときの平均値である。重み係数が 500 のときは、すべてのロボット間距離で BER が 0[%] であった。

環境音が 40~45[dB] という実験環境では、重み係数が 500 のとき (SNR は 40.8[dB]), SPL フィールド全体に誤りなく情報を伝送できると思われる。重み係数が 900, 1600 のときは、拡散ビット列の音の大きさを弱めすぎてしまい、環境音と分離できなくなったものと考えられる。

## 5 おわりに

本研究では、合成音声にスペクトル拡散させたビット列を埋め込み、実環境で送受信した。環境音が 40~45[dB] であれば、SNR が 40.8[dB] の音声を用いて SPL フィールド全体に誤りなく情報を伝送できることがわかった。

今後の課題として、ノイズの少ない実験環境ではなく、より環境音が大きく、瞬間的なノイズも混在し得る

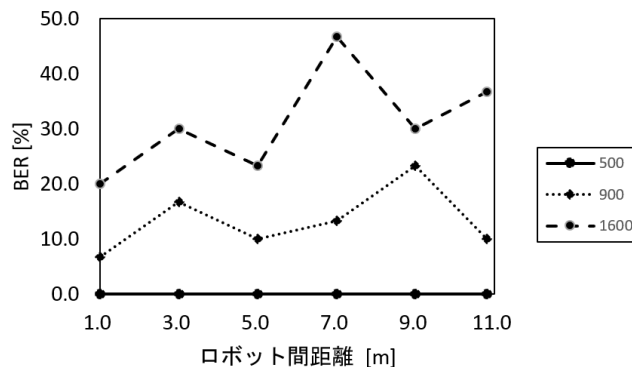


図 9: 重み係数毎のボット間距離に対応する BER の推移

状況での対処方法を検討していきたい。また、観客にとって違和感のない SNR の限界や、伝送速度を増加させる方法も探りたい。

## 謝辞

スペクトル拡散法について助言していただいた本学情報科学部准教授の神谷幸宏先生に感謝する。またロボカッププロジェクトチームのカメラアドラゴンズは、2019 年度本学から活動費の支援を受けている。

## 参考文献

- [1] ロボカップ日本委員会. “ロボカップとは”. <http://www.robocup.or.jp/robocup/>, (cited 2020-02-24).
- [2] RoboCup Standard Platform League. <https://spl.robocup.org/>, (cited 2020-02-24).
- [3] 神谷幸宏. MATLAB によるデジタル無線通信技術. コロナ社. 2008.
- [4] 神谷幸宏. RF ワールド No.31 特集 スペクトル拡散の技術. CQ 出版社. 2015. pp.46-56.
- [5] 竹花進悟, 近藤和弘, 中川清司. スペクトル拡散を用いた音響信号用電子透かしの基礎検討. 電気関係学会東北支部連合大会講演論文集. p.344. 2003.
- [6] 神谷幸宏. RF ワールド No.31 特集 スペクトル拡散の技術. CQ 出版社. 2015. pp.92-95.
- [7] 坂野秀樹. 考えて使いこなす音声のスペクトル分析 (やさしい解説). 日本音響学会誌. 2012, 68 巻, 4 号, pp.188-194.

- [8] 板橋秀一 編著. 音声工学. 森北出版. 2005. 256p.
- [9] SoftBank Robotics. “NAO - Documentation”.  
SOFTBANK ROBOTICS DOCUMENTATION.  
[http://doc.aldebaran.com/2-8/home\\_NAO.html](http://doc.aldebaran.com/2-8/home_NAO.html), (cited 2020-02-23).