

技術解説 第1回

『音声コミュニケーションを支える技術(1)』

従来、音声ネットワークはデータネットワークとは独立に構築されてきました。しかし近年登場したVoIP(Voice over IP)技術を用いることでIPネットワークを介した音声通話の実現が可能となり、両者を統合したネットワークを構築出来るようになってきています。

本技術解説ではまずVoIPを実現する4つの要素技術を解説します。

今回はこのうち音声コーデックについて取り上げます。

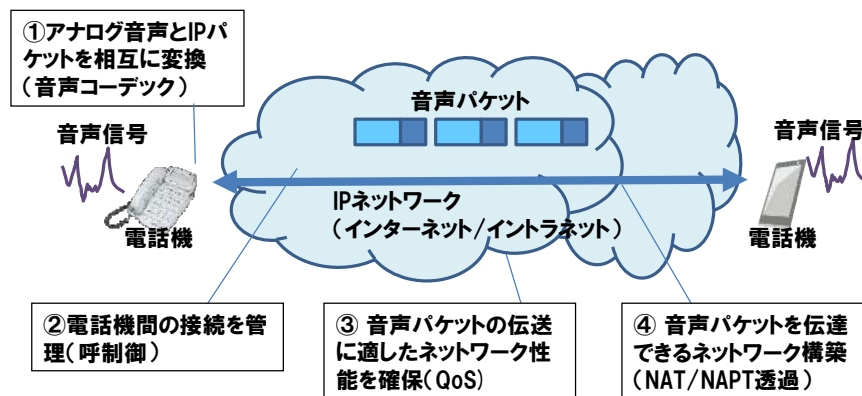
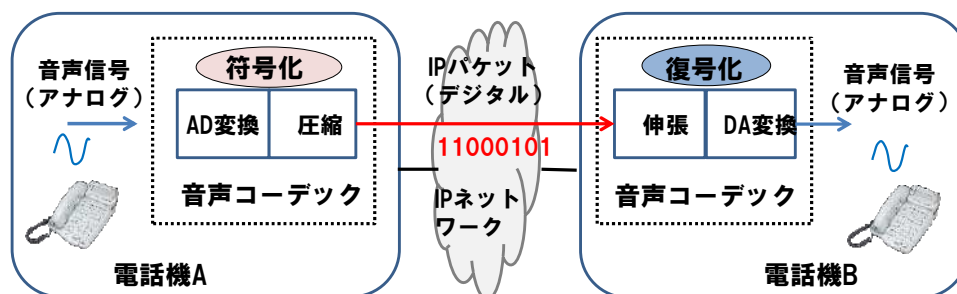


図 1: VoIP を実現する要素技術

音声コーデック技術

● 音声コーデックとは

IPネットワークを介して音声信号を電話機Aから電話機Bに送信するシステムにおける音声コーデックの模式図を示します。



IPネットワークはデジタル信号しか扱えません。アナログ信号である音声信号はデジタル化する必要があります。

図 2: 音声コーデック

電話機 A は音声信号の送り側です。IP ネットワークではデジタル信号しか扱えませんので、IP ネットワークに音声信号を流すには、送り側ではアナログの音声信号をデジタル化する必要があります。また伝送路を効率良く利用するためにデジタル信号は圧縮されます。

電話機 B は音声信号の受け側です。音声信号は電話機 A でデジタル化され、さらに圧縮されたデジタルデータとなっていますので、これを元の形に復元する必要があります。すなわち、受け取ったデジタルデータを伸張し、デジタル信号のアナログ信号化を行います。

このアナログの音声信号からデジタル信号に変換する機能 (AD 変換) と、デジタル信号のサイズを小さくする機能 (圧縮) をまとめてエンコード (符号化) といいます。逆に圧縮されたデジタルデータを元のデジタル信号に戻す機能 (伸張) と、デジタル信号からアナログ信号に変換する機能 (DA 変換) をまとめてデコード (復号化) といいます。

これらの機能をハードウェアやソフトウェアで実現したモジュールを「音声コーデック」と呼びます。コーデック (codec) とは coder (符号器) と decoder (復号器) の合成語です。

電話機 A で入力された音声信号 (アナログ信号) は、音声コーデックによって符号化処理と圧縮処理が行われ、IP データ化されます。IP データ化された音声情報は RTP (Real-time Transport Protocol) というリアルタイムデータを伝送するプロトコルで IP ネットワーク上に流れます。電話機 B では同じ音声コーデックによって伸張処理と復号化処理が行われ、電話機 A で入力された音声信号が再現され届けられることとなります。

●音声コーデック G.711 と G.729a

音声コーデックはいくつかの方式が ITU-T で勧告されており、代表的な方式は G.711 と G.729a があります。

G.711 は電話回線等でも使用されている、最も一般的な音声コーデックです。G.711 によってデジタル化された信号のビットレートは 64kbps と大きめですが、音質が良好なので広帯域な IP ネットワークでは良く用いられる方式です。

電話回線における音声周波数は最高で 3.4kHz です。原信号の 2 倍以上の周波数で標本化 (サンプリング) を行うと原信号を劣化なく再現できることから、G.711 ではサンプリング周波数 8kHz で標本化を行います。そしてサンプリングしたそれぞれの信号強度を 8bit、256 段階で表して量子化 (デジタル化) します。

復号はこの逆の操作をおこない、デジタル信号からアナログ信号へ戻します。G.711 では音声圧縮処理を行っていないため、原音の劣化を抑えて再現できます。

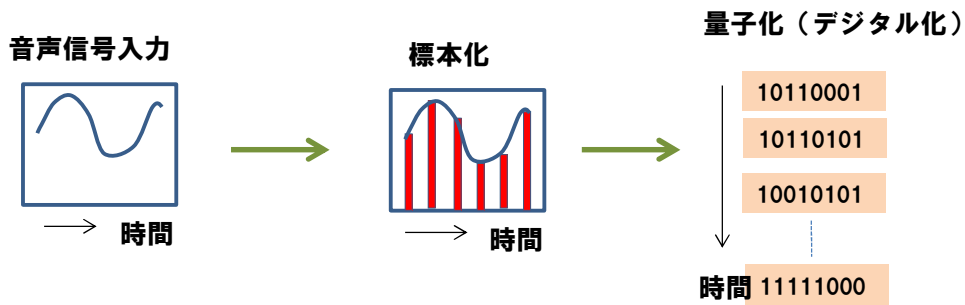


図 3: G.711 による符号化

G.729a は音声圧縮率が高く、IP ネットワークを効率良く利用できる音声コーデックであり、回線が細い携帯電話でも利用されています。音声信号そのものを符号化した情報を伝送するのではなく、あらかじめ波形パターンをコードブックというものに登録をしておき、波形パターン番号と実際の音声に近づけるための音響特性 (フィルタ係数) を伝送します。復号された音質は G.711 に比べると劣りますが、ビットレートは 8kbps と小さくなります。

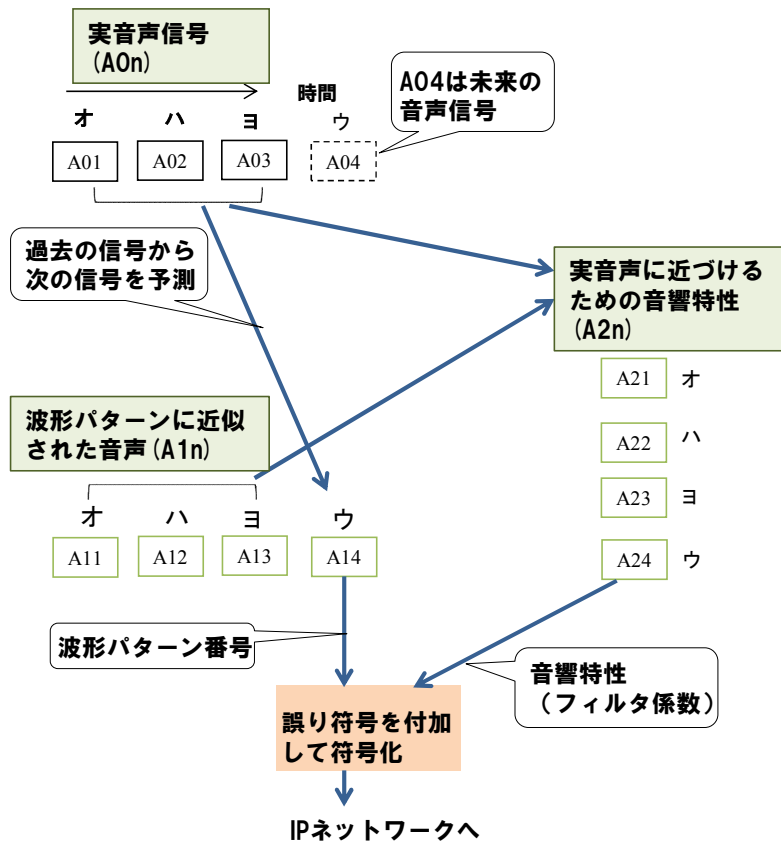


図 4: G.729a による符号化

今回は代表的な呼制御プロトコルである SIP について説明します。