
テキスト音声合成ライブラリ

「AquesTalk pico」

標準インターフェース仕様書(フレーム合成版)

株式会社アクエスト

www.a-quest.com

1. 概要

本文書は、弊社のテキスト音声合成ライブラリ **AquesTalk pico** の関数インターフェースを示したものです。
本文書の内容は弊社標準の基本仕様であり、カスタマイズによってはインターフェース仕様が異なることがあります。実装の際には、ライブラリ提供の際に付属されるマニュアルを参照してください。

AquesTalk pico は、よみ記号にアクセント等の情報を付与したローマ字表記の音声記号列から音声波形データを生成する日本語の音声合成ライブラリです。

なお、**AquesTalk pico** には、DA コンバーターや PWM タイマーなどのペリフェラルを操作する DA (Digital to Analog) 出力関数は含まれておらず、DA 出力関数はシステムに応じて別途実装する必要があります。

2. 基本仕様

入力	ローマ字表記音声記号列 (別紙「ローマ字音声記号列仕様」参照)
出力	10KHz または 8KHz サンプルング(声種により異なる) 16bitPCM (符号付)
声種	女声 4 種、男声 1 種、ロボット声 2 種の中から 1 種を予め選択
必要メモリ	ROM: 23Kbyte、RAM: 500byte (動作環境により変動あり)
その他	話速可変 約 50%~300%、DA 出力関数は含まず。実装が別途必要

3. 音声合成関数 API

以下に、本音声合成のライブラリに含まれ、システム側から呼び出す関数を示します。

CAqTkPicoF_Init

CAqTkPicoF.h

説明	初期化 処理に必要なワークバッファを指定。
構文	uint8_t CAqTkPicoF_Init(uint32_t *buf, uint16_t lenFrame)
引数	
buf	ワークバッファのアドレスを指定。 バッファのサイズは AQ_SIZE_WORKBUF (AqTkPicoF.h 内に規定) 音声合成が終了するまで、解放や内容の書き換えは不可。
lenFrame	フレーム長(1 度の SyntheFrame()呼び出しで生成するサンプル数)を 30-320 の間で指定。 デフォルト: 160
戻り値	0:正常終了 0 以外:エラー

CAqTkPicoF_SetKoe

CAqTkPicoF.h

説明	音声記号列をセット
構文	uint8_t CAqTkPicoF_SetKoe(const uint8_t *koe, uint16_t speed, uint16_t lenPause)

引数	
koe	音声記号列(1 文程度を指定)を指定(文字列の先頭アドレス。NULL 終端)
speed	発話速度の指定。 値を大きく設定するほど、速くなる。発話速度 [%] 50-300 の間で指定。デフォルト: 100
lenPause	最後のポーズ(無音区間)の長さを指定。 0xffffU を指定すると、内部の標準ポーズ長。 最後のポーズが不要な場合は、256 の指定を推薦。
戻り値	0:正常終了 0 以外:エラー

CAqTkPicoF_SyntheFrame

CAqTkPicoF.h

説明	音声波形を1フレーム生成 CAqTkPicoF_SetKoe()でセットした音声記号列から、音声を1フレーム生成して戻る。 (1フレームの長さは CAqTkPicoF_Init で先に指定)
構文	uint8_t CAqTkPicoF_SyntheFrame(short *wav, uint16_t *pLen)
引数	
wav	音声波形出力バッファ。生成した1フレームの音声波形データがここに書き込まれて戻る。 バッファのサイズは lenFrame で指定した値以上とし、外部で確保する。
pLen	生成したデータのサンプル数を返す。通常、最終フレーム以外は lenFrame で指定した値。 最終フレームは lenFrame 以下になる。
戻り値	0:正常終了 1:End of Data それ以外:エラー 最終フレームも0を返す。その次に呼び出したときに1を返す。

4. 関数呼び出しシーケンス

次コードを例に関数の呼び出しシーケンスを示します(コードはシーケンスを示すためのもので、実際に有効なソースコードではありません)。

1. CAqTkPicoF_Init()にて初期化とワークバッファを指定。
2. 続いて、CAqTkPicoF_SetKoe()にて合成するメッセージを指定。
3. 続いて、CAqTkPicoF_SyntheFrame()を繰り返し呼び出し、生成した1フレーム分の波形データをD/Aに出力する。
4. これを、1(End of Data)が返るまで繰り返す。

続いてメッセージを合成する場合は、2.のステップから繰り返します。ただし、ワークバッファを解放した場合は 1.のステップから行います。

なお、音声合成を他の処理と並列に行う場合は、4.のループ処理を展開するようにします。

```
#define FRAME_LEN 160
uint32_t buf[AQ_SIZE_WORKBUF];
const uint8_t koe[]="konnichiwa.";

uint8_t iret;
CAqTkPicoF_Init(buf, FRAME_LEN);
iret = CAqTkPicoF_SetKoe(koe, 100, 256);
if(iret!=0) <<<error proc>>>
```

```

for(;;){
    short wavFrame[FRAME_LEN];
    uint16_t size;
    iret = CAqTkPicoF_SyntheFrame(wavFrame, &size);
    if(iret==1) break; // End of Data
    if(iret!=0) <<<error proc>>>
    <<< Copy "wavFrame" data to D/A output buffer. Length is "size" or FRAME_LEN >>>
}

```

5. エラーコード表

関数の戻り値とそのエラー内容を以下に示します。

0	正常終了
100	その他のエラー
105	音声記号列に未定義の読み記号が指定された
106	音声記号列のタグの指定が正しくない
107	タグの長さが制限を越えている(または[>]が見つからない)
121	CAqTkPicoF_Init()で初期化されていない
122	関数の引数の指定が正しくない
123	ワークバッファが外部から書き換えられた
200	音声記号列が長すぎる

以上