
テキスト音声合成ライブラリ

「AquesTalk10」

標準インターフェース仕様書(フレーム合成版)

AQUEST

www.a-quest.com

1. 概要

本文書は、テキスト音声合成ライブラリ **AquesTalk10** の低レベル関数インターフェースを示したものです。
本文書の内容は標準仕様であり、カスタマイズにより仕様が異なることがあります。実装の際には、SDK に付属のマニュアルを参照してください。

AquesTalk10 は、よみ記号にアクセント等の情報を付与した、かな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。

外部ライブラリやヒープメモリを使わないため、OS 無しでも動作します。

AquesTalk10 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、音声波形のデータを固定のフレームに分割してメモリ上に生成します。この音声波形データをアナログ出力する部分は、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	かな表記音声記号列 (別紙「音声記号列仕様」参照) UTF-8
出力	16KHz サンプリング 16bitPCM (符号付)
声種	ベース音素片 3 種から選択 (複数可) パラメータでユーザカスタマイズ可
処理量 (CPU)	20DMIPS (各種 32/64bit CPU)
RAM	11KB (ワーク:9KB, 出力バッファ:640B, Stack:700B)
ROM	134~321KB (プログラム+音素片データ1つ~3つ)
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

* 数値は Cortex-M3 での実測値。構成・プラットフォームにより多少変動します。

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。

CAqTk_Init	AqTkF.h
説明	初期化 処理に必要なワークバッファを指定。
構文	HAQTK CAqTk_Init(void *buf, int bufSize)
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[byte]は AQ_SIZE_WORKBUF (AqTkF.h 内に規定) アライメントは sizeof(int)以上 音声合成が終了するまで、解放や内容の書き換えは不可。
bufSize	バッファのサイズ (AQ_SIZE_WORKBUF) を指定
戻り値	0 以外: AquesTalk インスタンスのハンドル 0: ライブラリが異なる
CAqTk_SetVoice	AqTkF.h
説明	声質パラメータの設定
構文	int CAqTk_SetVoice(HAQTK hAqTk, int bas, int vol, int spd, int lmd, int pit, int acc)
引数	
hAqTk	AquesTalk インスタンスのハンドル
bas	声色のベースとなる基本素片 F1E, F2E, M1E のいずれかを指定 (0/1/2)
vol	音量 0-300 default:100

spd	話速 値が大きいほど発話速度が速い 50-300 default:100
lmd	音程1 主に声質の高低を表現するが、より複雑な声質の変化がある。0-200 default:100
pit	高さ 声の高さがパラメータに比例する。20-200 default:基本素片に依存
acc	アクセントの強さ。値が大きいほどアクセントによるピッチの高低が大きくなる。 0-200 default:基本素片に依存
戻り値	0:正常終了 0以外:エラー

CAqTk_SetLastPauseLen

AqTkF.h

説明	最後のポーズの長さを指定
構文	int CAqTk_SetLastPauseLen(HAQTK hAqTk, int lenPause);
引数	
hAqTk	AquesTalk インスタンスのハンドル
lenPause	最後のポーズ長[sample] -1 を指定すると、内部の標準ポーズ長。(デフォルト) 0 で最後のポーズ長が0になる 最後のポーズが不要な場合は、512 の指定を推薦。
戻り値	0:正常終了 0以外:エラー

CAqTk_SetKoe

AqTkF.h

説明	音声記号列をセット
構文	int CAqTk_SetKoe(HAQTK hAqTk, char *koe)
引数	
hAqTk	AquesTalk インスタンスのハンドル
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)。 文字コードは UTF-8 関数内部で内容を書き換えることがあるため、const 文字列は指定できない。
戻り値	0:正常終了 0以外:エラー

CAqTk_SyntheFrame

AqTkF.h

説明	音声波形を1フレーム生成 CAqTk_SetKoe()でセットした音声記号列から、音声波形を1フレーム生成して戻る。 (1フレームの長さは SAMPLE_PER_FRAME で先に指定)
構文	int CAqTk_SyntheFrame(HAQTK hAqTk, short *wav, short *pSize)
引数	
hAqTk	AquesTalk インスタンスのハンドル
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 SAMPLE_PER_FRAME。バッファは呼び出し側で確保する。
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 SAMPLE_PER_FRAME となる。 最終フレームは SAMPLE_PER_FRAME 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。

CAqTk_SetDevKey AqTkF.h

説明	開発ライセンスキー設定 音声波形を生成する前に一度呼び出す。
構文	int CAqTk_SetDevKey(const char *key)
引数	
key	ライセンスキーの文字列(半角英数)を指定
戻り値	ライセンスキーが正しければ 0、正しくなければ 1 が返る 不正なキーでも 0 を返す場合がある。このとき制限は解除されない。

CAqTk_SetUsrKey AqTkF.h

説明	使用ライセンスキー設定 音声波形を生成する前に一度呼び出す。
構文	int CAqTk_SetUsrKey(const char *key)
引数	
key	ライセンスキーの文字列(半角英数)を指定
戻り値	ライセンスキーが正しければ 0、正しくなければ 1 が返る 不正なキーでも 0 を返す場合がある。このとき制限は解除されない。

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 CAqTk_Init()にて初期化とワークバッファを指定。
- 2 続いて、CAqTk_SetKoe()にて合成するメッセージを指定。
- 3 続いて、CAqTk_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。

```
#include <AqTkF.h>
char workbuf[AQ_SIZE_WORKBUF];
char koe[]="こんにちわ。";

uint8_t iret;
HAQTK aqtk = CAqTk_Init((void*)workbuf, AQ_SIZE_WORKBUF);
iret = CAqTk_SetKoe(aqtk, koe);
if(iret!=0) <<<error proc>>>

for(;;){
    short wavFrame[SAMPLE_PER_FRAME];
    short size;
    iret = CAqTk_SyntheFrame(aqtk, 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

関数呼び出しがエラーのときの戻り値とそのエラー内容を以下に示します。

100	その他のエラー
102	未初期化またはワークバッファが壊れている
103	音声記号列指定エラー (語頭の長音、促音の連続など)
104	音声記号列に有効な読みがない
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 [>] がみつからない)
108	タグ内の値の指定が正しくない
109	関数呼び出し時の引数の指定が正しくない
120	音声記号列が長すぎる
121	音声記号列の BP 句が長すぎる
122	音声記号列の BP 句が長すぎる

以上