

特定話者データについて

2版 2008年06月18日

1. 概要

SRM-10P または SRM-20P では、特定話者データを登録することにより任意の単語で音声認識を行うことが出来ます。
本文では、特定話者データを登録する際のいくつかの注意事項を解説します。

2. 特定話者データの単語選定について

①1 グループあたりの単語数を減らす

特定話者データでは、1 グループあたり最大で 15 単語まで登録可能です。
しかし、単語の登録数が多くなると誤認識をする割合が高くなります。
1 グループあたりの登録単語数を出来るだけ少なくしてください。(10 単語以下が目安となります。)

②同一グループ内に類似する発音を含む単語を入れない

同じグループ内に、類似する発音の単語が含まれている場合に、誤認識をする割合が高くなります。(例 あか、あさ)
出来るだけ別グループにするか、単語自体を変更するか、ご検討ください。

③長い単語、短い単語は使わない

特定話者データでは、3 秒以内で有れば任意の単語が登録可能です。
しかし、極端に短い単語 (1 文字) や、極端に長い単語 (10 文字以上) の場合、誤認識をする割合が高くなります。
出来るだけ登録する単語の文字数は 2~10 文字程度としてください。

3. 特定話者データの登録について

①登録時の環境について

特定話者データの登録は、マイク入力から直接音声を入力して登録を行います。
そのため、反響や暗騒音 (※) を考慮して、使用する環境下で登録を行う方が、より認識精度を高くすることが出来ます。
但し、騒がしい (他の人の声、機械音等) 環境下で登録を行った場合、雑音を一緒に登録してしまい認識精度が低下する場合があります。特定話者データの登録を行う際には、静かな環境下で登録を行うようにしてください。

※暗騒音とは、環境騒音とも言い、その環境における騒音のことです。例えば、エアコン等から、常に発せられている騒音等が、それにあたります。

②単語の途中に無音がある単語について

特定話者データの登録は、音声を入力した後、0.5 秒の無音が続くと登録を完了します。
そのため、「まわれ右」のように単語の途中で無音が入る単語の場合、途中の無音が 0.5 秒以上続くと「まわれ」で登録が完了してしまいます。
1 つの単語として認識させるためには、単語の途中の無音を 0.5 秒以内としてください。

【注意】

上記の注意事項は、誤認識や認識精度を低下させないための対策ですので、必ずしも認識精度が改善されるわけではありません。
実際の認識精度に関しては、お客様の仕様に合わせて十分な評価を行ってください。

ご注意

- ・本文書の著作権は（株）アルファプロジェクトが保有します。
- ・本文書の内容を無断で転載することは一切禁止します。
- ・本文書に記載された回路図およびサンプルプログラム等の著作権は（株）アルファプロジェクトが保有しますが、お客様のアプリケーションで使用される場合には、ご自由にご利用いただけます。
- ・本文書の内容は、将来予告なしに変更されることがあります。
- ・本文書に記載されている内容およびサンプルプログラムについての質問等のサポートは一切受け付けておりませんのでご了承ください。
- ・本文書の内容については、万全を期して作成いたしました。万が一不審な点、誤りなどお気づきの点がありましたら弊社までご連絡下さい。
- ・本文書の内容およびサンプルプログラムに基づき、アプリケーションを運用した結果、万が一損害が発生しても、弊社では一切責任を負いませんのでご了承下さい。

- ・その他の会社名、製品名は、各社の登録商標または商標です。



株式会社アルファプロジェクト
〒431-3114
静岡県浜松市東区積志町 834
<http://www.apnet.co.jp>
E-MAIL : query@apnet.co.jp