

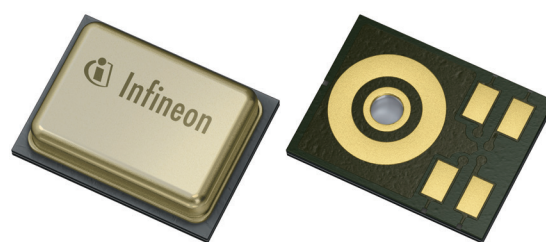
AN547 - 高性能、高SNR MEMSマイクロフォンが必要な理由

目次

1. 要約	1
2. SN比	2
3. 最大入力音圧レベル (AOP)	2
4. 録音用マイクロフォンにおける性能の重要性	4
4.1 録音におけるSN比の重要性	4
4.2 録音におけるAOPの重要性	5
5. 音声認識におけるマイクロフォン性能の重要性	5
6. ノイズキャンセルアルゴリズムにおけるマイクロフォン性能の重要性	6
7. まとめ	6
7.1 略語リスト	7
免責事項	8

1. 要約

自動音声認識システムの普及と、情報や経験を共有するためのビデオコンテンツの使用は劇的に増加しています。優れたユーザーエクスペリエンスを確保するには、音をとらえるために使用されるマイクロフォンの性能と品質が高くなければなりません。マイクにおける重要なエレメントとして、ノイズ、歪み、周波数応答、部品マッチングなどがあります。本アプリケーションノートでは、SN比(SNR)と最大入力音圧レベル(AOP)に焦点を当て、音声認識およびオーディオ/ビデオキャプチャシステムで高いマイク性能を起用する利点について説明します。



2 SN比 (SNR)

2. SN比 (SNR)

マイクロフォンの出力におけるノイズは、意図された入力源ではない信号として定義され、一般に出力信号に望ましくない要素とみなされます。ノイズレベルが高いほど、オーディオ信号の品質が低下します。ノイズは、マイクロフォンの外部にある場合もあれば、マイクロフォン自体から発生する場合があります。マイクロフォンの自己ノイズは、人には通常サーという音に聞こえ、知覚される音質に影響を与えます。アルゴリズムの場合、ノイズは信号の忠実度を劣化させ、それによってシステム性能も低下します。マイクロフォンのノイズは、さまざまな方法で説明されます。

- **自己ノイズ**(ノイズ= N_{Vrms} , dBV, dBFS)は、外部音によって励起されないときにマイクロフォン自体が生成する実効ノイズ電圧です。
- **信号対ノイズ比**(SNR(dB))は、目的の入力信号に対するマイクの自己ノイズを表します。SNRは通常、94dBSPL(1Pa)の正弦波で所望の音を表すために標準化された音響入力信号を使用して測定されます。
- **等価入力ノイズ EIN**(dBSPL)は、マイク自体が発生する雑音で、この雑音成分を入力音圧に換算したものが等価雑音レベルです。

3. 最大入力音圧レベル(AOP)

すべてのオーディオトランスデューサは、通過する信号にコンテンツを追加する点で非線形システムです。歪みの場合には、追加されたコンテンツは、元の信号に存在する周波数の高調波です。歪みは、通常、全高調波ひずみ、THD(自己ノイズが含まれる場合はTHD+N)として測定されます。それは、マイクロフォンが正弦波で励起されたときの信号高調波(通常2番目から5番目)のエネルギーと基本周波数のエネルギーの比です。試験信号は、典型的には、比較的高い音圧レベル(SPL)、通常は94dBSPLかそれ以上の1kHzの信号となります。THDは百分率(%)で示されます。

最大入力音圧レベル (AOP) は、一般に、THDが10%を超える音圧レベルとして定義されます。AOPの単位はdBSPLです。



ほとんどの場合、マイクロフォンに入ってくる音の原形と内容を保存することが有益で重要となります。元の信号に歪みなどのコンテンツが追加されると、キャプチャされたサウンドを聞いている人に不愉快に聞こえる可能性があります。より多くの追加エネルギーがある(すなわち高いTHD)と、知覚されるオーディオ品質は悪化します。歪みはまた、着信信号の内容の非常に細かい分析を実行する音声認識システムのようなアルゴリズムを混乱させる可能性があります。

AN547 - 高性能、高SNR MEMSマイクロフォンが必要な理由

3 最大入力音圧レベル (AOP)

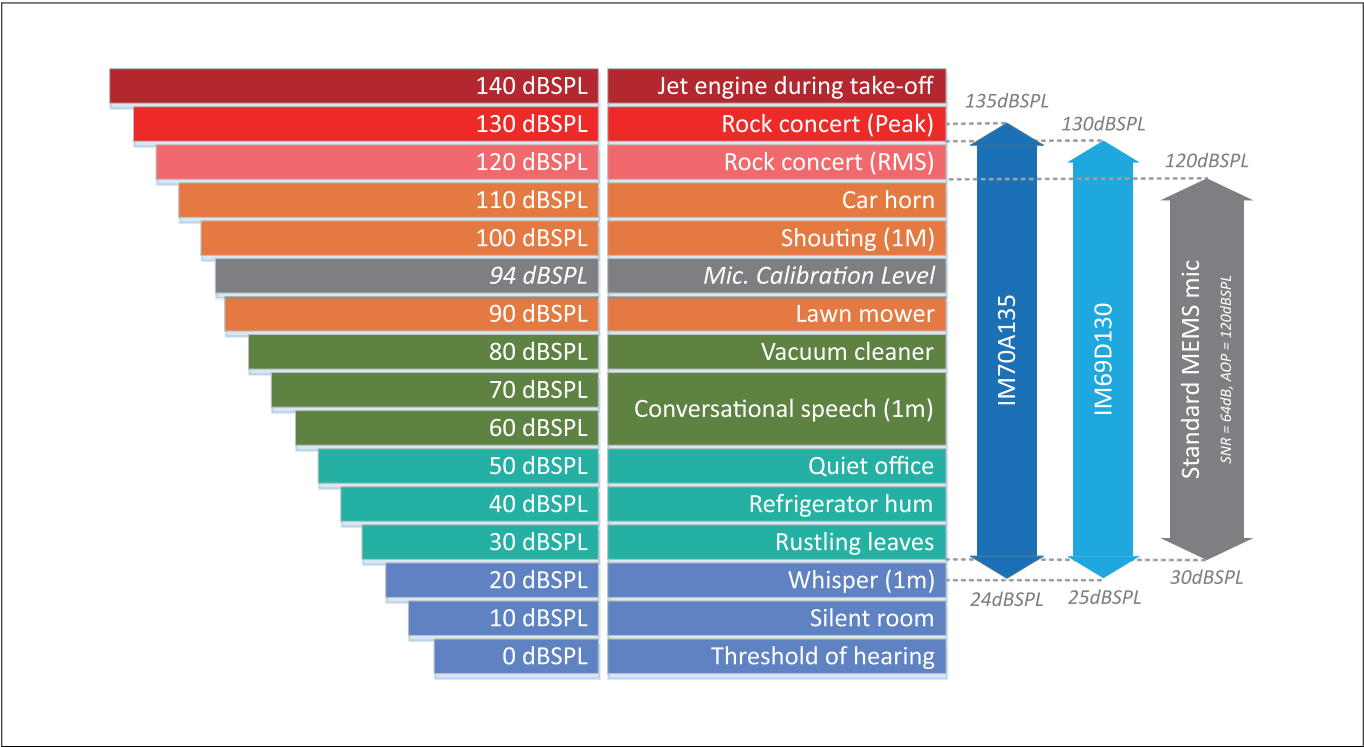


図1 IM69D130とIM70A135のダイナミックレンジと標準MEMSマイクロフォン

4 録音用マイクロフォン性能の重要性

4. 録音用マイクロフォン性能の重要性

4.1 録音におけるSN比の重要性

オーディオ/ビデオ録画の目的は、対象物からの着信音を捕捉し、それをマイクロフォンシステムの出力に再現することです。録音が人間の耳に向けられている場合、電気出力信号が音響信号を可能な限り近づけて「自然な」音声録音を提供することが望ましいです。マイクとそのSNRは、サウンドキャプチャ信号チェーンの重要な部分であり、オーディオ録音の質に影響します。典型的な使用例を以下の表に示します。

使用例	詳細と課題
ホームビデオ	通常、家はマイクノイズが容易に優勢になる静かな環境です。 様々なキャプチャリング、再生条件、機器があります。
子供	撮影対象物はよく動き、ソフトな(静かな)声をしています。
ソーシャルメディア	視聴者の関与を最大限得るために高画質が必要とされます。
プロフェッショナルビデオ	求人応募、面接、人材紹介、プレゼンテーションなど。 高画質は、応募者やビジネスの差別化を図るために重要です。
音楽	自然な録音を確保するためには高音質が重要です。 多様なキャプチャリングや再生条件が課題となっています。
パフォーマンス	学校の演奏は難しいことがあります。 小さな声、対象までの距離、周囲の騒音
自然	録音された音は、低い音圧レベルまたは非常に低い音圧レベルです。
監視	とらえる音は、小さく、離れたところからの音になります。

フリーフィールドでは、距離が倍になるごとに音圧が半減(6dB低下)します。とらえる音源が離れているほど、マイクロフォンに届く音響信号は小さくなります。マイクロフォンの自己ノイズが実質的に一定であるので、入力信号レベルの低下は、マイクロフォンの出力信号のSNRを低下させることになります。典型的には、弱い信号を増幅してデバイスの信号経路に適切なレベルにする必要があります。信号を増幅すると、出力に存在するノイズも増幅されます。増幅が増えるほど、ノイズはとらえた信号の品質を著しく低下させるレベルになるリスクが高くなります。

高マイクロフォンSNRは、信号が増幅されてもノイズフロアを聞こえるようにします。キャプチャリング距離が長くなればなるほど、マイクの自己ノイズは問題を避けるために低くしなければなりません。これは、距離が長く音源自体が小さい場合には特に問題となります。距離が倍増することにより音圧が6dB減衰するので、SNRが6dB高いマイクロフォンを使用すると、信号品質を低下させることなくキャプチャリング距離を2倍にすることができます。POLQA(Perceptual Objective Listening Quality Assessment)は、デジタル音声分析を使用するITU-T標準モデルであり、記録された音声信号の品質および明瞭度を非常に客観的に決定します。高いSNRを持つマイクロフォンは、POLQAテストで明らかに優れたパフォーマンスを発揮し、優れたスピーチの明瞭度をもたらします。より高いSNRマイクロフォンで録音した場合、同じレベルの信号がより分かりやすくなります。

再生条件およびビデオ画質は、知覚されるノイズレベルに影響を及ぼします。

- 再生環境の周囲ノイズレベル
- 再生音量
- 聴取機器の品質(例えば、ノイズおよび周波数応答)
- 高いビデオ品質では、全体的なオーディオ/ビデオ品質の低下を防ぐため、高音質が要求されます。

5 音声認識におけるマイクロフォン性能の重要性

4.2 録音におけるAOPの重要性

AOPはSNRと同様に、重要なオーディオ/ビデオ品質ファクターです。歪みにより、ビデオは簡単に使い物にならなくなってしまいます。多くのスマートフォン動画がオンラインでポップ/ロックコンサートで撮影され、オンライン上にありますが、オーディオの歪みために視聴に耐えないものになっています。意図する音(または外乱)の入ってくる音圧レベルが高いか非常に高い場合、AOPは音質を向上します。高AOPは、平均音圧レベルがそれほど高くなくても、入ってくる音響信号に現れる非常に高い信号ピークをマイクロフォンシステムが処理するのに役立ちます。典型的な使用例を以下の表に示します。

使用例	詳細と課題
ポップ/ロックミュージックコンサート	コンサートは通常は大音量で行われます。 高音質は、良好で自然な音響性能への鍵となります。
スポーツイベント	スポーツ(モータースポーツなど)または観衆(アイスホッケー競技場など)のいずれかが非常に大きい状況です。
交通	低周波ノイズが多くあります。
風	風は屋外で撮影されたオーディオ/ビデオ録画の音質を下げる一般的な原因です。 高いAOPは特定の種類の風条件に有効です。

数年前まで、家電製品のマイクロフォンAOPの標準レベルは110~120dBSPLでした。最近では、AOPの要求が高まっています。顧客を満足させる音質と音声認識性能を確保するために、デバイス設計者は130dBSPLマークに近いかそれ以上のAOPを持つより優れたマイクロフォンを選択する必要があります。

低い音圧レベルでは、AOPに指定された10%よりも低いTHDレベルの方が理にかなっています。高いAOPを持つことに加えて、THDは、2%以下の低いままで、意図する用途(例えば、最大120dBSPL)のために十分高い音圧レベルまで維持することも重要です。

5. 音声認識におけるマイクロフォン性能の重要性

とらえた音がアルゴリズムを対象とするシステムの場合、音質の目標は、信号が人間の耳のための場合と異なることがあります。アルゴリズムに最適化されていれば、信号は必ずしも自然な音である必要はありません。使用例にかかわらず、信号が外乱、アーチファクト、歪みおよびノイズをきれいに保つことが常に重要です。

自動音声認識(ASR)は、音声信号を文章に自動的に転記するタスクです。転記精度は、人間レベルに近づいており、約95%です。

しかしながら、これまでのところ、このレベルを達成することは、周囲条件が整った実験室においてのみ可能でした。

実際の環境や離れた場所での音声認識には、バックグラウンドノイズ、残響、エコーキャンセレーション、マイク測位などの重要な音響課題があります。優れた音声認識エンジンのためにはそれだけでは不十分です。システム内のすべての要素は、品質のボトルネックを防ぐために高い標準で性能を発揮しなくてはなりません。マイクロフォンの仕事は、最良の入力信号をもつ音声認識システムを提供することです。高い入力信号品質は、ASRシステムが入ってくる音を分析することと、音声内容を認識できる特徴を見つけることに役立ちます。重要なパラメータは、ノイズ、歪み、周波数応答、位相です。

高AOPは、騒々しい環境での音声認識システムを可能にします。

時には、音声信号そのものが小さくなく、他の外乱が存在する場合もあります例えば、音声制御された家庭用娯楽システムおよびデジタルアシスタントのマイクロフォンに近いスピーカがあり、大きな音量の音楽または音声情報を出力することができます。

高AOPは歪みを低く保ち、ノイズとエコーのキャンセレーション

6 ノイズキャンセルアルゴリズムにおけるマイクロフォン性能の重要性

を改善します。

音声源までの距離が長いほど、ASRアルゴリズムに供給される信号の信号対雑音比が低くなります。従って、意図されたキャプチャリング距離が長くなると、マイクロフォンSNRが高くなるべきです。



6. ノイズキャンセルアルゴリズムにおけるマイクロフォン性能の重要性

音声認識システムの重要な機能は、転記されるべき音声以外の音声と雑音を無視することができることです。オーディオ/ビデオキャプチャと人間と人間との間の通信品質もまた、信号から不要な音を排除することで改善されます。目標はSNRを上げること、ここではSNRは所望の音(信号)と不要な周囲音(ノイズ)との比になります。

ノイズキャンセレーションと指向性は、アルゴリズムと組み合わせて複数のマイクロフォンを使用することで達成されます。ビームフォーミングのような指向性マイクロフォンシステムは、目的の音源を強調表示します。望ましくないサウンドも2つのマイク間のレベル差などのパラメータに基づいてキャンセルすることができます。ブラインドソース分離は、より高度なノイズリダクションシステムです。これは、方向、距離と場所の点で個別にノイズキャンセルを可能にします。これらのノイズキャンセル方法はすべて、信号の精度と高い品質によって可能となるものです。マイクロフォンは、高いSNR、低歪み、平坦な周波数応答(そして位相応答改善)、および低群遅延を備える必要があります。

ノイズキャンセルアルゴリズムの機能を最適化するために、システムで使用されるマイクロフォンは同じ特性を持つ必要があります。マイクロフォンのマッチングに対するマイクロフォンの役割は非常に重要です。感度、位相挙動、および遅延度がマイクロフォンごとに差が少ないほど良いとされます。

7. まとめ

2005年から2015年にかけて、マスマーケット家電製品の最先端マイクロフォンのSNRIは、60dB以下から約65dBまで改善されました。新しい高性能音声認識システムとその他のキャプチャリング使用例によって設定された要件により、65dBでさえもはや十分ではありません。現在のハイエンドマイクロフォンは70dB SNRに近づいています。高いマイク性能は、高い音声認識とオーディオキャプチャの品質にとって重要な要素です。自動音声認識アルゴリズムやカメラなどの技術の性能は急速に向上しており、デバイス購入者のユーザーエクスペリエンスの期待は高まっています。マイクが改善のボトルネックにならないようにすることが重要です。幸いなことに、高性能マイクが開発されています。ノイズ性能は、ここ数年で飛躍的に改善しました。SNRは70dBレベルを超えて上昇しており、品質劣化歪みはAOPが130dB SPLマークに達した過去のものになりつつあります。このレベルのマイクロフォン性能は、デバイスのユーザーエクスペリエンスを、最も要求の厳しい顧客でさえも満足いただけるものにします。

7 まとめ

7.1 略語リスト

SNR:信号対雑音比

EIN:等価入力ノイズ

THD:全高調波歪

AOP:最大入力レベル

ASR:自動音声認識

SPL:音圧レベル

Pa:パスカル、圧力の単位

CE:家電製品

Trademarks

All referenced product or service names and trademarks are the property of their respective owners.

Edition 2017-10-17

Published by

Infineon Technologies AG
81726 Munich, Germany

© 2017 Infineon Technologies AG
All Rights Reserved.

Do you have a question about any aspect of this document?

Email: erratum@infineon.com

Document reference
IFX-ovy1506346816951

IMPORTANT NOTICE

The information contained in this application note is given as a hint for the implementation of the product only and shall in no event be regarded as a description or warranty of a certain functionality, condition or quality of the product. Before implementation of the product, the recipient of this application note must verify any function and other technical information given herein in the real application. Infineon Technologies hereby disclaims any and all warranties and liabilities of any kind (including without limitation warranties of non-infringement of intellectual property rights of any third party) with respect to any and all information given in this application note.

The data contained in this document is exclusively intended for technically trained staff. It is the responsibility of customer's technical departments to evaluate the suitability of the product for the intended application and the completeness of the product information given in this document with respect to such application.

WARNINGS

Due to technical requirements products may contain dangerous substances. For information on the types in question please contact your nearest Infineon Technologies office.

Except as otherwise explicitly approved by Infineon Technologies in a written document signed by authorized representatives of Infineon Technologies, Infineon Technologies' products may not be used in any applications where a failure of the product or any consequences of the use thereof can reasonably be expected to result in personal injury