

Infineons neue Xensiv-MEMS-Mikrofone

Hochwertige Audiosignalerfassung

Endgeräte mit Sprachsteuerung und die Verwendung von Video- und Audioaufnahmen zum Übermitteln von Informationen und Erlebnissen werden immer beliebter.

Einschränkender Faktor bisher war oftmals die Qualität der verwendeten MEMS-Mikrofone. Mit dem kürzlich vorgestellten IM69D130 überwindet Infineon die bisherigen Beschränkungen in der Audiosignalkette.

VON MARCEL KNECHT,
MARKETING MANAGER
MEMS MICROPHONES,
INFINEON TECHNOLOGIES



Der Markt für MEMS-Mikrofone ist in den letzten Jahren rasant gewachsen. MEMS-Mikrofone werden hauptsächlich in Mobiltelefonen, Laptops, Kopfhörern, digitalen Sprachassistenten, Smart-Home-Geräten, Kameras, Robotern sowie in der Automobilbranche eingesetzt.

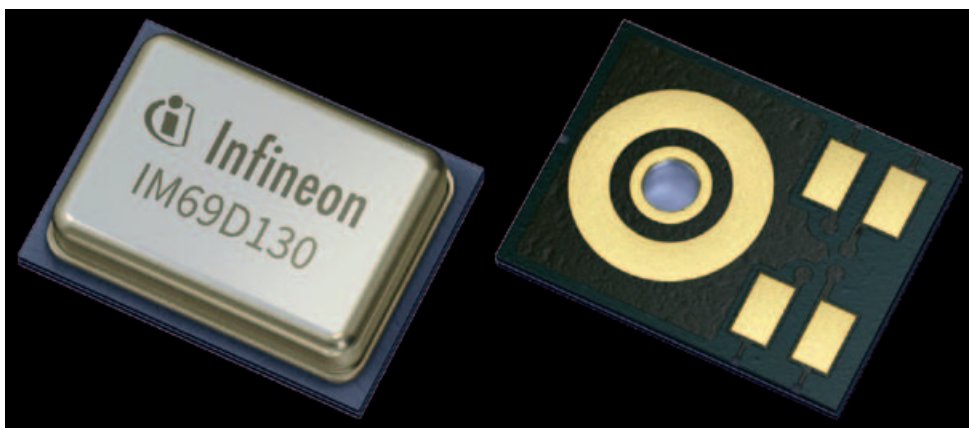
MEMS-Mikrofone bieten gegenüber herkömmlichen Elektret-Kondensatormikrofonen (ECM) viele Vorteile. Zum einen sind sie in derselben Leistungsklasse kleiner, zum anderen sind Amplitude und Phase von Audiosignalen aus mehreren Mikrofonen aufeinander abgestimmt. Außerdem ist die MEMS-Technologie robust gegen hohe Temperaturen und eignet sich für das Reflow-Löten, was die automatisierte Bestückung von Platinen erlaubt.

Ein Mikrofon setzt Schalldruckwellen in elektrische Signale um. Allerdings sind nicht alle Mikrofone gleichermaßen leistungsfähig, und es gibt verschiedene Parameter, die entscheidend dafür sind, ob sich ein Mikrofon für einen bestimmten Anwendungsfall eignet.

Kriterien für leistungsfähige Mikrofone

Unter dem elektrischen Rauschen im Ausgangssignal des Mikrofons versteht man alle Signale, die nicht aus dem gewünschten Eingangssignal stammen. Verschiedene Parameter bzw. Spezifikationen beschreiben das Rauschen bei Mikrofonen. Zum einen das Eigenrauschen. Darunter versteht man das Rauschen, das vom Mikrofon selbst erzeugt wird, wenn kein Tonsignal vorhanden ist. Dies wird in V_{rms} , dBV oder dBFS gemessen. Unter Äquivalent-Eingangsrauschen versteht man den imaginären akustischen Rauschpegel, der dem elektrischen Rauschpegel am Mikrofonausgang entspricht. Es wird ausgedrückt in dB SPL Schalldruck (dB Sound Pressure Level). Ein wichtiges Kriterium ist der Signal-Rausch-Abstand (SNR). Der SNR-Wert, angegeben in dB, ist ein Maß für das Eigenrauschen des Mikrofons relativ zum vorgesehenen Eingangssignal.

Weitere wichtige Kenngrößen für die Qualität von Mikrofonen sind die Verzerrung (Total Harmonic Distortion, THD) und der Acoustic Overload Point (AOP). In der Realität sind Mikrofone wie alle Signalwandler nichtlinear, d.h. sie erzeugen eine gewisse Verzerrung. Im Fall von Verzerrungen liegen die zusätzlichen Signale als Harmonische vor (üblicherweise 2. bis 5. Harmonische). Unter THD versteht man das Verhältnis der Energie, die diese Harmonischen enthalten, zur Energie der Grundfrequenz. Sie wird als Prozentwert ausgedrückt. Grundsätzlich definiert der AOP den Punkt, an dem die THD 10 Prozent überschreitet. In anspruchsvolleren Anwendungsbereichen wird der AOP allerdings gelegentlich auch als der Punkt spezifiziert, an dem die THD 1 Prozent überschreitet.



Das Xensiv-MEMS-Mikrofon IM69D130 läutet eine neue Leistungsklasse für digitale Premium-MEMS-Mikrofone ein, die die bisherigen Beschränkungen in der Audio-Signalkette überwindet.

Optimierte Spracherkennung

Für Systeme, die mit Algorithmen arbeiten, werden Tonsignale anders erfasst, als sie vom menschlichen Ohr wahrgenommen werden. Daher sind auch die Zielvorgaben hinsichtlich der Tonqualität unterschiedlich. Das Signal muss nicht unbedingt natürlich klingen, solange es für die verwendeten Algorithmen optimiert ist. Unabhängig vom Anwendungsfall ist es wichtig, dass das Signal frei von Störungen, Verzerrungen und Rauschen bleibt.

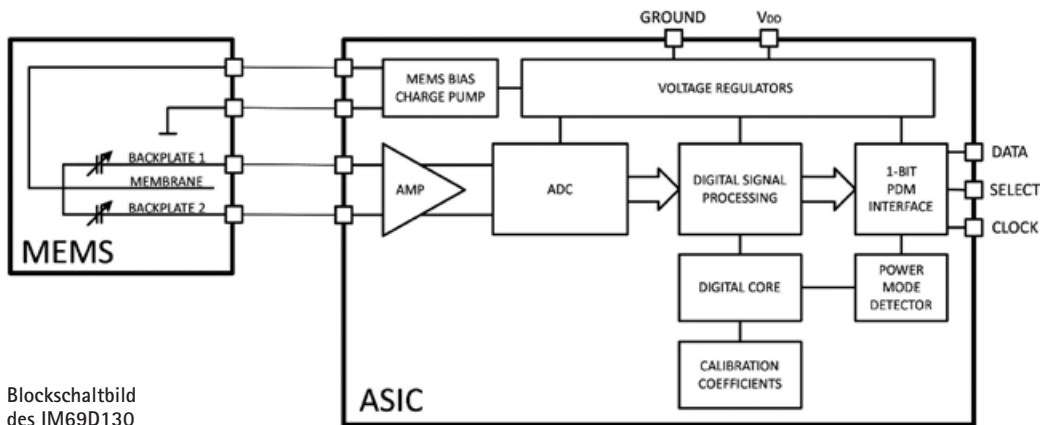
Unter automatischer Spracherkennung versteht man den Prozess, durch den Sprachsignale automatisch in geschriebenen Text übertragen werden. Die Genauigkeit liegt derzeit bei etwa 95 Prozent und kommt bereits sehr nahe an das menschliche Niveau heran. Dieser Wert wurde bisher aber nur in Labors mit sehr günstigen Umgebungsbedingungen erreicht.

In der Praxis hat die Sprachsteuerung, insbesondere wenn sich der Sprecher nicht in unmittelbarer Nähe befindet, meist mit erheblichen Schwierigkeiten in Bezug auf die Akustik zu kämpfen, zum Beispiel mit Hintergrundgeräuschen, Nachhall, Echounterdrückung und der Mikrofonposition. Daher reicht es nicht, einfach nur eine gute Spracherkennungs-Software zu haben. Jeder Baustein des Systems sollte bestmögliche Leistungen liefern, damit keine Qualitätseinbußen auftreten. Ein qualitativ hochwertiges Eingangssignal hilft, den sprachlichen Inhalt eines Geräusches zu analysieren. Zu den entscheidenden Parametern gehören dabei Rauschen, Verzerrung, Frequenzgang und Phase.

In Umgebungen mit starken Störgeräuschen kann die Spracherkennung erheblich verbessert werden, wenn das verwendete Mikrofon eine hohe Linearität, das heißt möglichst wenige Verzerrungen aufweist. Ein hoher AOP trägt dazu bei, Verzerrungen gering zu halten und die Unterdrückung von Störgeräuschen und Echos zu verbessern. Manchmal ist das Sprachsignal selbst nicht laut genug und es sind andere Geräusche vorhanden, die Störungen verursachen. Das ist beispielsweise der Fall, wenn digitale Sprachassistenten laute Musik oder gesprochene Informationen wiedergeben.

Möglichst großer Signal-Rausch-Abstand

Je größer der Abstand zur Quelle des Sprachsignals ist, desto höher sollte das Signal-Rausch-Verhältnis des Mikrofons sein.



Blockschaltbild des IM69D130

Die Erfassung von Audio- und Videosignalen kann ebenso wie die Qualität eines Gesprächs von Mensch zu Mensch verbessert werden, wenn unerwünschte Geräusche aus dem Signal ausgeblendet werden. Ziel ist es, den Signal-Rausch-Abstand zu vergrößern, also in diesem Fall das Verhältnis zwischen dem erwünschten Tonsignal und unerwünschten Umgebungsgeräuschen. Rauschunterdrückung und Richtcharakteristik lassen sich durch Verwendung mehrerer Mikrofone in Verbindung mit geeigneten Algorithmen erreichen.

Richtmikrofonsysteme, die beispielsweise mit Beamforming-Algorithmen arbeiten, können die Empfindlichkeit der Mikrofone in der gewünschten Richtung erhöhen und die erwünschten Schallquellen hervorheben. Eine ausgefeilte Methode zur Unterdrückung von Störgeräuschen sind „Blind Source Separation“-Algorithmen. Diese ermöglichen die Unterdrückung von Störgeräuschen unabhängig von der Richtung, von der Entfernung und vom Ort der Entstehung. Alle diese Verfahren zur Unterdrückung von Störgeräuschen profitieren von der Genauigkeit und Qualität des Signals, das sie empfangen. Das Mikrofon sollte daher einen möglichst großen Signal-Rausch-Abstand, eine geringe Verzerrung, einen linearen Frequenzgang (der auch den Phasengang verbessert) und eine geringe Gruppenlaufzeit aufweisen.

Funktionsmerkmale und Kennwerte

Das IM69D130-Xensiv-MEMS-Mikrofon wurde speziell für Applikationen entwickelt, bei denen ein geringes Eigenrauschen, ein großer Dynamikbereich, niedrige Verzerrungen und ein hoher AOP erforderlich sind. Die Mikrofone haben ein Signal-Rausch-Verhältnis (SNR) von 69 dB.

Das Mikrofon verfügt über die Infineon-Dual-Back-Plate-Technologie, die auf einem miniaturisierten symmetrischen Mikrofondesign ba-

siert, wie es in Studio-Kondensatormikrofonen verwendet wird. Diese Konstruktion bietet eine hohe Linearität des Ausgangssignals in einem Dynamikbereich von 105 dB. Das Grundrauschen des Mikrofons liegt bei 25 dB (69 dB Signal-Rausch-Verhältnis); die Verzerrung steigt auch bei einem Schalldruckpegel von 128 dB SPL nicht über 1 Prozent. Das bedeutet, dass die verzerrungsfreie Erfassung von Sprachbefehlen auch bei laufender Musik aus einem Lautsprecher möglich ist. Der lineare Frequenzgang (28 Hz Niederfrequenz-Roll-off) und die engen Fertigungstoleranzen führen zu einer genauen Phasenübereinstimmung der Mikrofone. Das Bauteil ist in einem Gehäuse von 4 mm x 3 mm x 1,2 mm untergebracht.

Mit seiner Empfindlichkeit (± 1 dB) und Phasenübereinstimmung ($\pm 2^\circ$ bei 1 kHz) unterstützt das IM69D130 ein extrem genaues Audio-Beamforming, das innovative Audio- und Sprachalgorithmen mit hoher Leistungsfähigkeit ermöglicht. Da es über eine digitale Schnittstelle verfügt, benötigt es keine analogen Bauteile. Damit reduziert sich auch der Aufwand, die Platine vor Hochfrequenzstörungen zu schützen; bei Konstruktionen mit mehreren Mikrofonen werden weniger Datenleitungen benötigt.

Der digitale Mikrofon-ASIC enthält einen extrem rauscharmen Vorverstärker und einen schnellen Sigma-Delta-AD-Wandler (nur 6 μ s Latenzzeit bei 1 kHz). Verschiedene Power-Modi ermöglichen eine ideale Anpassung an spezifische Stromversorgungsanforderungen. Jedes IM69D130-Mikrofon ist kalibriert, was in sehr kleinen Toleranzen für die Empfindlichkeit (± 1 dB) resultiert.

In Verbindung mit innovativen Algorithmen zur Verarbeitung von Audiosignalen liefert das IM69D130 von Infineon ein hochwertiges, digitales Audio-Rohdatensignal, das sich auch für anspruchsvollste Spracherkennungsszenarien wie die Fernfeldererkennung oder die Erkennung von leisen Stimmen eignet. (nw)