



1 Zur Rolle des automatischen Mikrofonmischers in der Konferenztechnik und zum Ziel der Dissertation

Gegenstand dieser Arbeit ist der automatische Mikrofonmischer für Sprachanwendungen als zentrale Komponente der Konferenztechnik. Sein Einsatzbereich reicht von Sendestudios, Theatern, Kirchen und öffentlichen Einrichtungen bis hin zu großen Konferenzsälen. Dan Dugan, der Erfinder einer der ersten Automatismischer, hat sich wie kaum ein anderer mit der Thematik befasst. Seine Leidenschaft für das Thema zeigt sich auch in seiner Aussage zum Anwendungsfall automatischer Mikrofonmischer [1]:

"Well it's used for wherever you have a bunch of people talking. Even with two people it's useful. And when you get up to [...] six or seven or eight people it becomes essential and when you have [...] 20 or 30 people it becomes magic."

Der automatische Mikrofonmischer ist für viele Anwendungen nützlich und in manchen Fällen unentbehrlich. Durch seinen Einsatz wird die Rückkopplungsneigung einer elektroakustischen Anlage, unabhängig von der Anzahl benutzter Mikrofone, begrenzt. Zudem stellt er ein Mischsignal bereit, das hinsichtlich der Signalqualität und damit auch der Sprachverständlichkeit einem Mischsignal, das durch einfache Summation der Eingangssignale gewonnen wird, überlegen ist. Im Speziellen werden durch seinen Einsatz Störanteile, Nachhall und Kammfiltereffekte im Mischsignal reduziert.

Die ausschließliche Betrachtung des automatischen Mikrofonmischers als elektrotechnisches Gerät ist für das Verständnis der Thematik nicht zielführend. So wie sich einige Eigenschaften eines Mikrofons, beispielsweise die Richtcharakteristik, erst im Kontext einer Schallquelle zeigen, ergeben sich die Eigenschaften eines automatischen Mikrofonmischers erst im Kontext seiner Umgebung. Und so ergibt sich ein umfängliches Verständnis der Thematik erst dann, wenn der automatische Mikrofonmischer als Bestandteil eines elektroakustischen Systems bestehend aus Schallquellen, Raum, Mikrofonen,

Mischer, Lautsprechern und Hörern verstanden wird. Obwohl der automatische Mikrofonmischer ein rein elektrotechnisches Gerät ist, ist die Thematik des automatischen Mischens eng mit der Elektroakustik verbunden.

Das Ideal einer Konferenzanlage ist es, den Schall nur bei ausgewählten Sprechern „abzuholen“ und nur zu ausgewählten Hörern zu „bringen“. Dabei soll der automatische Mikrofonmischer eine scheinbar einfache Aufgabe lösen. Die Signale benutzter Mikrofone sollen den Mischer ohne Abschwächung passieren, Signale unbenutzter Mikrofone hingegen sollen abgeschwächt werden. Zusätzlich soll die Gesamtverstärkung des Mixers begrenzt werden. Dies sind die zwei essentiellen Prinzipien automatischer Mikrofonmischer, aus denen sich die technischen Realisierungen ableiten lassen. Dass zur Erfüllung dieser zwei Prinzipien anspruchsvolle technische Lösungen erforderlich sind, hat mehrere Gründe. Die Kopplung zwischen Schallquellen und *allen* Mikrofonen in einem Raum und der Einfluss von Störquellen zeigt sich auch in den Mikrofonsignalen, was die Erkennung benutzter und unbenutzter Kanäle erschwert. Zudem muss die Rückkopplung zwischen Lautsprechern und Mikrofonen berücksichtigt werden, wobei dem automatischen Mikrofonmischer hinsichtlich dieser Problematik eine besondere Rolle zukommt, da dieser direkten Einfluss auf die Schleifenverstärkung der elektroakustischen Anlage hat. Zuletzt ist der Hörer zu betrachten, dem das Mischsignal zugespielt wird. Das maßgebende Kriterium für eine zufriedenstellende Funktion der Beschallungsanlage ist die vom Hörer wahrgenommene Sprachverständlichkeit. Schon kleinste Fehlfunktionen eines automatischen Mikrofonmischer können dabei zu nicht tolerierbaren Störungen, beispielsweise dem Abschneiden von Anfangssilben, führen.

Wenngleich einzelne Komponenten des automatischen Mikrofonmischer bereits früher entwickelt wurden, hat sich der automatische Mikrofonmischer erst Anfang der 1970er-Jahre als eigenständiges Gerät in der Elektroakustik einen Namen gemacht. Wie kaum ein anderer in diesem Bereich hat sich Dan Dugan mit dem Thema auseinandergesetzt. Das von ihm entwickelte 1975 patentierte *Gainsharing*-Verfahren, das die verfügbare Verstärkung in Abhängigkeit der Signalpegel auf die einzelnen Mikrofonkanäle aufteilt, zählt, obwohl weitestgehend immer noch in seiner ursprünglichen Form betrieben,

auch heute noch zu den gefragtesten Verfahren und Geräte, die dieses Verfahren verwenden, werden aufgrund der Popularität seines Erfinders auch als *Dugan-Mischer* bezeichnet. Mit der Patentierung wurde das Verfahren lizenzpflichtig, was andere Firmen und Ingenieure veranlasste, neue Verfahren zu entwickeln und diese ebenfalls zu patentieren. In dieser Hinsicht beförderte die Patentierung der Verfahren die Neu- und Weiterentwicklungen auf diesem Gebiet. So patentierte Peters 1978 ein Verfahren mit einer ganz anderen Funktionsweise, das im *Voice-Matic Mixer* von *Industrial Research Products (IRP)* zum Einsatz kam. Dieser Mischer wird heute zwar nicht mehr produziert, hat sich in der praktischen Anwendung jedoch bewährt. Im Jahre 1982 patentierte Shure, ein etablierter Hersteller von Mikrofonen, ein Verfahren, das Spezialmikrofone mit zwei Kapseln nutzte, um die Aktivierung eines Mikrofonkanals nur aus einem vorgegebenen räumlichen Bereich zu erlauben. Dieses Verfahren konnte sich trotz seiner Tauglichkeit zumindest langfristig nicht durchsetzen, offenbar da die Verwendung von Spezialmikrofonen die Flexibilität bei der Mikrofonwahl einschränkte und höhere Kosten mit sich brachte. Und so patentierte Shure 1986 ein weiteres Verfahren, das sogenannte *Maxbus*-Verfahren, das anstrebt nur ein Mikrofon pro Schallquelle zu aktivieren. Dieses Verfahren hat sich in der Praxis besonders bewährt und das Prinzip wird nahezu unverändert in den aktuellen Mischern von Shure eingesetzt. Bis heute basieren viele Automatismischer auf dem *Gainsharing*- oder dem *Maxbus*-Verfahren.

Die Mehrheit der Verfahren wurde im Zeitalter der Analogtechnik entwickelt. Da die heute weit entwickelte Digitaltechnik gegenüber der Analogtechnik enorme Vorteile hinsichtlich Kosten, Speicherbarkeit und Zuverlässigkeit aufweist, werden die Verfahren heute mit digitaler Technik realisiert. Am Funktionsprinzip der Verfahren hat sich dadurch wenig geändert – die heute digital betriebenen Verfahren stellen Digitalisierungen der analogen Verfahren dar.

Bis heute hat sich keines der Verfahren als generell überlegen erwiesen und so wundert es nicht, dass bis heute rege Diskussionen darüber stattfinden, welches der Verfahren für eine Beschallungsanwendung das beste sei. Zu dieser Frage findet man im *Handbook for Sound Engineers* die folgende

Aussage [2]:

"There is no definitive answer to this question. It is impossible to tell which automatic microphone mixer design will operate best in a given situation by studying technical specifications, believing the marketing literature, poring over circuit schematics, deciphering lines of computer code, or rereading this chapter. Human speech is very complex and human hearing is very discerning. Like so many areas in professional audio, the critical ear is the final judge."

Aus wissenschaftlicher Perspektive ist diese Antwort ernüchternd, doch zeigt sie auch die Schwierigkeiten auf, die eine objektive Bewertung der Verfahren begleiten. So ist ohne die Berücksichtigung der Umgebungsbedingungen und durch eine rein technische Betrachtung des automatischen Mikrofonmischers keine eindeutige Aussage über die Eignung eines Verfahrens möglich. Zudem wird auf die Komplexität des menschlichen Sprechens und die hohe Empfindlichkeit des menschlichen Hörens verwiesen. Soll eine Bewertung erfolgen, müssen die spezifischen Eigenschaften von Sprachsignalen berücksichtigt werden. Außerdem muss bei der Bewertung der Hörer in die Betrachtung einbezogen werden. Der Antwort auf die Frage, welches der Verfahren das beste ist, etwas näherzukommen, ist ein Ziel der vorliegenden Arbeit.

Nachdem nun die Rolle des automatischen Mikrofonmischers in der Beschallungstechnik kurz beschrieben wurde und die Entwicklungsgeschichte automatischer Mikrofonmischer skizziert wurde, werden im Folgenden Aufbau und Ziele der vorliegenden Arbeit herausgestellt.

In Kapitel 2 wird die Problematik des Mischens von Mikrofonsignalen aus Perspektive der Elektroakustik diskutiert. Es wird gezeigt, dass der ordnungsgemäße Betrieb einer elektroakustischen Anlage bei Verwendung mehrerer Mikrofone ohne den Einsatz eines automatischen Mikrofonmischers in der Regel nicht möglich ist.

Ausgehend von diesen Betrachtungen werden in Kapitel 3 die zwei wesentlichen Prinzipien automatischer Mikrofonmischer abgeleitet. Es werden

die grundlegenden Anforderungen an automatische Mikrofonmischer festgehalten und typische Einstellparameter und Funktionen beschrieben.

Kapitel 4 gibt eine detaillierte Übersicht der bestehenden Verfahren. Neben Verfahren, die heute noch gebräuchlich sind, werden auch solche Verfahren vorgestellt, die sich nicht durchsetzen konnten, da dies zum Verständnis der Thematik beiträgt. Die Kenntnis über die technische Funktionsweise eines Verfahrens ist auch bei der praktischen Anwendung von Vorteil. Insbesondere bei der Einrichtung einer elektroakustischen Anlage, kann das Wissen über die Wirkungsweise eines automatischen Mikrofonmischers von großem Nutzen sein. Nur so können die Ursachen für ein Fehlverhalten der Anlage schnell identifiziert, potenzielle Lösungen gefunden und die Auswirkungen von Einstellparametern abgeschätzt werden. Da viele der bestehenden Verfahren patentiert wurden, wird auch kurz auf die Bedeutung und Auswirkungen der Patentierung eingegangen.

Die Beurteilung der Verfahren beruht bis heute überwiegend auf individuellen Erfahrungen. Es hat sich bisher kein Gütekriterium etabliert, das vergleichende Aussagen zwischen den Verfahren zulässt. In Kapitel 5 wird ein Gütekriterium entwickelt, das die objektive Bewertung der Verfahren ermöglicht. Die zur Berechnung des Gütekriteriums erforderlichen Signale werden durch eine Simulationsumgebung gewonnen. Mithilfe des Gütekriteriums und der Simulationsumgebung werden die bestehenden Verfahren bewertet und verglichen. Außerdem wird die Fähigkeit der Verfahren, den Kammfiltereffekt zu verhindern, genauer untersucht.

In Kapitel 6 werden Vorteile digitaler Signalverarbeitung und digitaler Systeme vorgestellt, die auch in automatischen Mikrofonmischem nutzbringend eingesetzt werden können.

Das Vorkommen mehrerer gleichzeitig aktiver Sprecher ist ein nicht zu vernachlässigendes Merkmal von Konversationen [3]. Die Bewertung der Verfahren in Kapitel 5 wird jedoch zeigen, dass die Aktivierung eines Kanals bei den bestehenden Verfahren verzögert erfolgt, wenn bereits ein Sprecher aktiv ist. In Kapitel 7 wird ein Multiband-Automatikmischer vorgestellt, der den bestehenden Verfahren hinsichtlich dieser Problematik überlegen ist.

Dennoch erzielen die heute zur Verfügung stehenden Verfahren in den

meisten Anwendungen bereits gute Ergebnisse. Von einer Verbesserung dieser Verfahren kann daher kein allzu großer Zugewinn erwartet werden. Die bestehenden Verfahren sind jedoch nicht in der Lage, zwischen Sprachsignalen und transienten Störgeräuschen zu unterscheiden. Ruft ein Sprecher in der Nähe eines Mikrofons ein transientes Störgeräusch hervor, etwa durch Blättern in einem Papierstapel, wird dieses auch im Mischsignal hörbar sein. Akustisches Übersprechen zwischen einer Schallquelle und den Mikrofonen erschwert den Einsatz von Sprachdetektoren. Außerdem darf die Aktivierung eines Kanals durch den Sprachdetektor nicht verzögert werden. In Kapitel 8 werden zwei Algorithmen zur Sprachdetektion entwickelt und ein Konzept vorgestellt, das die Nutzung von Sprachdetektoren in einem automatischen Mikrofonmischer, trotz der bestehenden Schwierigkeiten, möglich macht.

Obwohl ihr Name es vermuten lässt, sorgen automatische Mikrofonmischer nicht für eine Angleichung der Pegel mehrerer Mikrofonsignale. Ein weiterer Beitrag dieser Arbeit liegt deshalb in der Erweiterung des automatischen Mikrofonmischer um eine automatische Verstärkungsregelung. In Kapitel 9 wird die übliche Sichtweise, die automatische Verstärkungsregelung als vom automatischen Mikrofonmischer unabhängige Komponente anzusehen, infrage gestellt und ein Konzept vorgestellt, das die Verwendung einer automatischen Verstärkungsregelung in einem automatischen Mikrofonmischer, trotz des akustischen Übersprechens zwischen einem Sprecher und allen Mikrofonen, möglich macht.

Die vorliegende Arbeit richtet sich sowohl an den interessierten Anwender von Automatismischern, für den in den Kapiteln 2, 3, 4 und 5 die grundlegenden Konzepte der Thematik dargelegt werden, als auch an den Entwickler, für den die Kapitel 6, 7, 8 und 9 Inspirationsquellen für Neu- und Weiterentwicklungen sein können.

2 Problemstellung: Das Mischen von Mikrofonsignalen

Das Mischen, der von mehreren Mikrofonen aufgenommenen Sprachsignale, zum Zweck der Aufnahme, Übertragung oder der verstärkten Wiedergabe mittels Lautsprechern, ist für viele Anwendungen erforderlich. Im einfachsten Fall besteht das Mischen im Summieren aller Mikrofonsignale, wie in Abbildung 1 dargestellt.

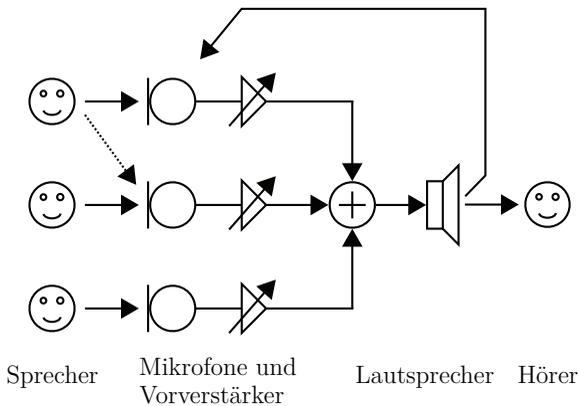


Abbildung 1: Schematische Darstellung einer einfachen elektroakustischen Beschallungsanlage bei der das Mischsignal durch Summation der Mikrofonsignale gewonnen wird

Eine so simple Anordnung ist jedoch in vielerlei Hinsicht unzureichend. Da Störgeräusche und Raumhall von allen Mikrofonen aufgenommen werden, der durch einen Sprecher hervorgerufene Direktschall jedoch größtenteils nur von einem Mikrofon übertragen wird, reduziert sich das Signal-Rausch-Verhältnis und das Direktsignal-Raumsignal-Verhältnis des Mischsignals mit der Anzahl der verwendeten Mikrofone. Zudem führt akustisches Übersprechen, das in Abbildung 1 durch den gestrichelten Pfeil angedeutet ist, zu einer Überlagerung zeitversetzter Signale und somit zu Kammfiltereffekten, die die Qualität des Mischsignals ebenfalls herabsetzen. Zuletzt erhöht sich die Schleifenverstärkung des elektroakustischen Systems, aufgrund der Rück-

kopplung des Lautsprecherschalls zu den Mikrofonen, mit steigender Anzahl der Mikrofonkanäle. Schon bevor die Stabilitätsgrenze überschritten wird, kann eine hohe Schleifenverstärkung in einzelnen Frequenzbereichen zum Nachschwingen der Beschallungsanlage in diesen Frequenzbereichen führen, wodurch die Sprachverständlichkeit abermals reduziert wird [4]. Im Folgenden werden die beschriebenen Effekte genauer betrachtet und deren Einfluss auf die Qualität des Mischsignals diskutiert. Die daraus resultierenden Erkenntnisse werden Ausgangspunkt zur Formulierung der Prinzipien automatischer Mikrofonmischer sein.

2.1 Störanteile und Nachhall im Mischsignal

Bei der Betrachtung des Mischsignals stellt die Stimme eines Sprechers das Nutzsinal dar. Störschall kann beispielsweise durch Klimatisierungsanlagen, Lüfter oder Hintergrundgespräche hervorgerufen werden. Hinzu kommt das Eigenrauschen der Mikrofone und Mikrofonvorverstärker. Auch der durch akustisches Übersprechen resultierende Schall eines entfernten Sprechers ist als Störschall anzusehen. Es wird nun der Fall betrachtet, dass alle Mikrofone einen Störanteil mit dem Effektivwert N_{eff} übertragen und nur ein Mikrofon zusätzlich einen Nutzanteil mit dem Effektivwert S_{eff} überträgt. Befinden sich mehrere Mikrofone in ausreichendem Abstand zueinander, ergibt sich der Effektivwert des Störanteils im Mischsignal aus der pythagoreischen Summe der einzelnen Effektivwerte [5]. Würde nur ein Mikrofon verwendet werden, ergäbe sich das Signal-Rausch-Verhältnis des Mikrofonsignals aus dem Verhältnis des Effektivwerts S_{eff} des Nutzanteils und dem Effektivwert N_{eff} des Störanteils. Werden hingegen N Mikrofonsignale summiert, von denen alle einen Störanteil enthalten und nur eines den Nutzanteil, ergibt sich das Signal-Rausch-Verhältnis aus dem Verhältnis S_{eff} zur pythagoreischen Summe der N Störanteile mit den Effektivwerten N_{eff} . Setzt man diese beiden Werte ins Verhältnis, ergibt sich die Reduktion des Signal-Rausch-Verhältnisses des Mischsignals gegenüber dem Signal-Rausch-Verhältnis des benutzten Kanals

in Abhängigkeit der Anzahl verwendeter Kanäle [5]:

$$\Delta SNR = 20 \log_{10} \left(\frac{\frac{S_{\text{eff}}}{\sqrt{N N_{\text{eff}}^2}}}{\frac{S_{\text{eff}}}{N_{\text{eff}}}} \right) = 20 \log_{10} \sqrt{\frac{1}{N}} = -10 \log_{10} N \quad . \quad (1)$$

Gleichung (1) zeigt, dass sich das Signal-Rausch-Verhältnis des Mischsignals gegenüber dem Signal-Rausch-Verhältnis des benutzten Kanals um 3 dB pro Verdopplung der Kanalanzahl N reduziert. Das bestmögliche Signal-Rausch-Verhältnis ergibt sich, wenn nur das dem Sprecher zugeordnete Mikrofon aktiv ist. Jedes weitere sorgt für eine Verschlechterung.

Die Tragweite dieses Zusammenhangs soll an einem praktischen Beispiel demonstriert werden: In einem Konferenzraum mit acht Teilnehmern herrscht ein Störschallpegel von 42 dB und ein Sprecher ruft an einem Mikrofon einen Schallpegel von 60 dB hervor. Aus diesen Werten ergibt sich ein Signal-Rausch-Verhältnis von 18 dB. Das Signal-Rausch-Verhältnis soll für eine ausreichende Sprachverständlichkeit mindestens 15 dB betragen, ein in der Praxis bewährter Richtwert [6]. Der Raum ist für eine Konferenz daher durchaus geeignet. Wird nun jedem der Sprecher ein Mikrofon zur Verfügung gestellt und das Mischsignal durch Summation der Mikrofonsignale erzeugt, reduziert sich das Signal-Rausch-Verhältnis nach Gleichung (1) um $10 \log_{10} 8 \approx 9$ dB auf 9 dB und liegt damit unter dem angestrebten Richtwert.

Neben Störgeräuschen findet sich auch der von jedem Mikrofon aufgenommene Raumschall im Mischsignal wieder. Auch hierbei ergibt sich unter vereinfachenden Annahmen eine Reduzierung des Direktsignal-Raumsignal-Verhältnisses um 3 dB pro Verdopplung der Kanalanzahl N [5]. Insbesondere bei langen Nachhallzeiten führt ein geringes Direktsignal-Raumsignal-Verhältnis zu einer Reduzierung der Sprachverständlichkeit [7].

Es ist außerdem zu berücksichtigen, dass sich das Signal-Rausch-Verhältnis beziehungsweise das Direktsignal-Raumsignal-Verhältnis am Hörerort bei der Wiedergabe des Mischsignals über Lautsprecher durch den Störschall, beziehungsweise Nachhall am Hörerort, weiter reduziert.

Nach dem *Cocktailparty-Effekt* kann ein Hörer aus dem von mehreren Schallquellen hervorgerufenen Schall, den Schall einer gewünschten Schall-

quelle extrahieren. Für die Teilnehmer einer Konferenz kann durch diesen Effekt, der nur bei beidseitigem Hören funktioniert, eine Verbesserung des Signal-Rausch-Verhältnisses zwischen 5 und 15 dB erzielt werden [8]. Eine scheinbar gute Akustik im Konferenzraum kann für einen Hörer an einem entfernten Ort, dem das Mischsignal über einen Lautsprecher als Monosignal zugespielt wird und der diesen Effekt deshalb nicht nutzen kann, unzureichend sein.

Zusammenfassend führt die einfache Summation der Mikrofonsignale in vielen praktischen Fällen zu einem unzureichenden Signal-Rausch-Verhältnis und Direktsignal-Raumsignal-Verhältnis und damit zur Unbrauchbarkeit des Mischsignals.

2.2 Kammfiltereffekt

Ein weiteres Problem, das sich beim Summieren von Mikrofonsignalen ergibt, ist das Auftreten des Kammfiltereffekts. In Abbildung 1 ist durch den gestrichelten Pfeil angedeutet, dass der von einem Sprecher hervorgerufene Schall nicht nur das nahe, sondern auch ein entferntes Mikrofon erreicht. Werden die Mikrofonsignale summiert, führt dieses akustische Übersprechen zu einer Überlagerung zeitversetzter Signale. Abbildung 2 zeigt den Betragsfrequenzgang des Kammfilters, der bei einer Wegdifferenz von 1 m und einer daraus resultierenden Pegeldifferenz von 6 dB entsteht. Die frequenzabhängigen Überhöhungen und Abschwächungen führen zu einem dünnen und hohlen Klang. Außerdem werden Sprachanteile, die im Frequenzbereich einer Senke liegen abgeschwächt, wodurch die Sprachverständlichkeit merklich reduziert werden kann. Brunner et al. zeigten, dass Kammfiltereffekte selbst bei einer Abschwächung des zeitversetzten Signals um 18 dB, insbesondere bei Laufzeitunterschieden zwischen 0,5 und 3 ms, noch wahrnehmbar sind [9]. Die entsprechenden Wegdifferenzen liegen zwischen etwa 0,17 und 1 m und bewegen sich damit in einem Bereich, der in einigen praktischen Anwendungsfällen, z. B. bei der Positionierung mehrerer Mikrofone an einem Tisch, anzutreffen ist. Bei der Einrichtung der Beschallungsanlage sollte die 3:1-Regel angewandt werden, die besagt, dass wenn eine Schallquelle einen