

FELDERHOFF Uwe, MACKENSEN Philip, THEILE Günther [IRT]

**STABILITÄT DER LOKALISATION BEI VERFÄLSCHTER REPRODUKTION
VERSCHIEDENER MERKMALE DER BINAURALEN SIGNALE**

1. Einleitung

Vorausgegangene Untersuchungen belegen, daß die durch Kopfdrehung hervorgerufenen dynamischen Merkmale der binauralen Signale notwendig sind, um insbesondere in der Medianebene eine stabile Lokalisation zu gewährleisten. Daher ist für hochwertige Verfahren der virtuellen Akustik die Anwendung von Head-Tracker Systemen unverzichtbar. Der damit verbundene hohe Aufwand bei der Berechnung der binauralen Signale ließe sich möglicherweise beachtlich reduzieren, wenn die Toleranz des Gehörs gegenüber verschiedener Vereinfachungen bekannt wäre. Vor diesem Hintergrund wurde mit Hilfe des bereits vorgestellten drehbaren Kunstkopfes zunächst die maximale Latenzzeit für die Nachführung des Kunstkopfes untersucht, um Daten für die erforderliche Rechengeschwindigkeit zu ermitteln. Ferner ist die Frage zu beantworten, in welchem Maße die Anwendung genauer (z.B. individueller) Außenohrübertragungsfunktionen lohnend ist. Die hier beschriebenen Richtungshörversuche mit einer Abstraktion des Kunstkopfes belegen zunächst die relativ große Bedeutung der dynamischen Merkmale.

2. Untersuchung zur maximalen Latenzzeit von Auralisationssystemen

Eine wahrnehmbare Latenzzeit eines Auralisationssystems mit Head-Tracking bewirkt bei der Wiedergabe von stationären Schallquellen und schnellen Kopfauslenkungen, eine kurzzeitige, ungewollte Auslenkung des Hörereignisortes (*siehe Abbildung 1*). Dieser störende Effekt tritt bei Latenzzeitüberschreitung von ΔT_{\max} auf.

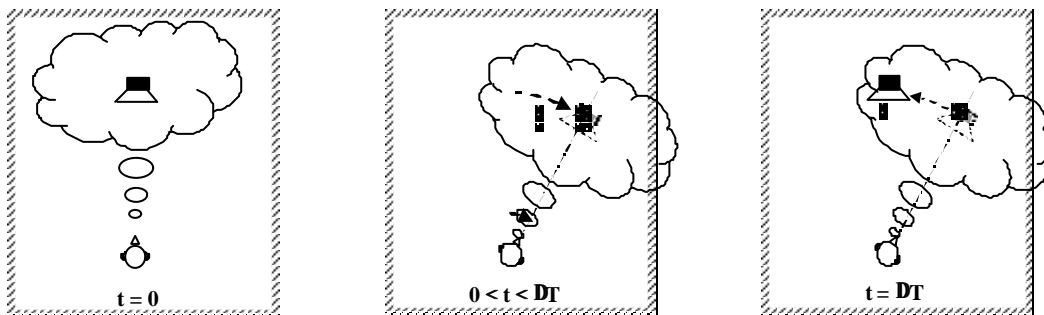


Abb. 1: Auswirkung von wahrnehmbaren Latenzzeiten ($\Delta T > \Delta T_{\max}$)

2.1 Wo treten bei Auralisationssystemen Latenzzeiten auf?

Zur Berechnung der binauralen Signale werden bei modellbasierenden Auralisationssystemen Schallereignisse, je nach Schalleinfallrichtung beim Hörer, mit den entsprechenden HRTF (aus einer Datenbank) gefaltet. Dabei werden auch Laufzeiten und Schalleinfallrichtungen der frühen Reflexionen, sowie der Absorptionsgrad der Reflexionsflächen berücksichtigt. Der Rechenaufwand solcher Systeme steigt mit Anzahl der gerechneten Schallquellen, berücksichtigten frühen Reflexionen und mit Länge der benutzten HRTF beliebig an. Je nach Rechenleistung der benutzten Hardware kann so der Rechenaufwand eine Signalverzögerung hervorrufen, welche die maximale Latenzzeit überschreitet. Hinzu kommt eine zusätzliche Verzögerung infolge des Tracking-Systems.

2.2 Versuchsanordnung

Die Wahrnehmbarkeitsschwelle der Latenzzeit wurde mit Hilfe eines Kunstkopfes durchgeführt (Neumann KU100), der mittels Tracking-System den spontanen Kopfdrehbewegungen der Versuchsperson folgt. Der Aufbau geht aus Abbildung. 2 hervor.

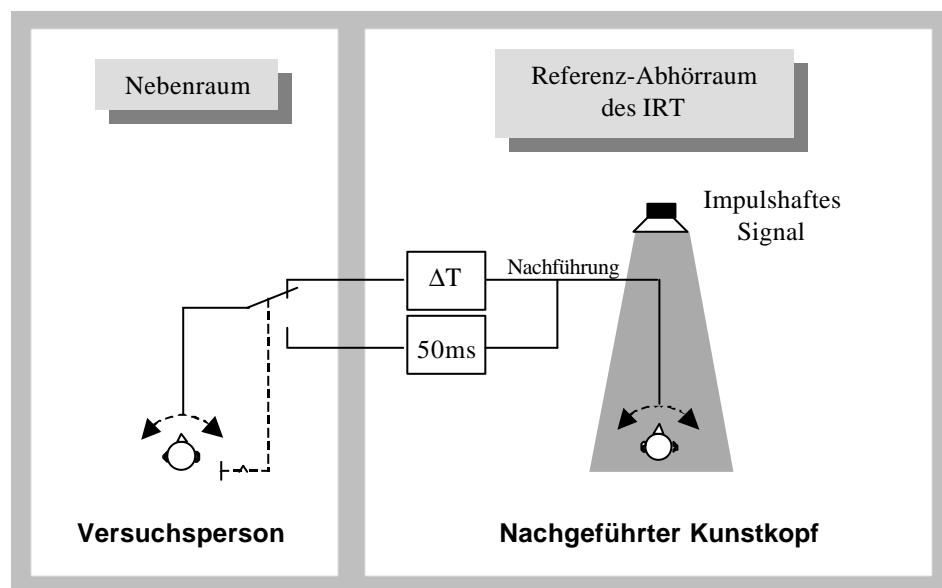


Abb. 2: Hörversuchsanordnung zur Bestimmung der maximalen Latenzzeit \mathbf{DL}_{max} .

Der Kunstkopf wird mittels Motor-Dreheinheit (TR3 von *THOMA Filmtechnik*) und Tracking-System (Polhemus FasTrak von *The VR Source*) den Rotationsbewegungen des Kopfes der Versuchsperson entsprechend nachgeführt. Die Umsetzung der Tracking-Signale zur Ansteuerung des Motors erfolgte mittels PC, der über die seriellen Anschlüsse mit dem Tracking-System und der Motor-Dreheinheit verbunden war. Die Updaterate des Head-Trackers betrug 120 Updates pro Sekunde (bzw. alle 8,333 ms). Die Baudraten der seriellen Schnittstellen waren so eingestellt, daß Datenübertragung (FasTrak \Rightarrow PC \Rightarrow Motorsteuerung) und Datenkonvertierung zwischen zwei Updates erfolgte. Dadurch wurden zusätzliche Verzögerungen vermieden. Die gesamte Latenzzeit des Systems (Auslenkung Kopf \Rightarrow Auslenkung Kunstkopf) beträgt 50 ms.

2.3 Versuchsdurchführung

Die Aufgabe der Versuchspersonen bestand darin, durch Umschaltung zwischen verzögerter Nachführung (ΔT) - und „nicht verzögerter“ Nachführung (50 ms) des Kunstkopfes, Unterschiede in Bezug auf Lokalisation der Schallquelle festzustellen (siehe Abb. 2). Dabei wurden die Probanden angewiesen, kleine schnelle Kopfauslenkungen durchzuführen, um so die Auswirkungen der Latenzzeit leichter wahrzunehmen. Erst wenn die Versuchsperson mündlich ihre Entscheidung mitgeteilt hatte, ob, und wenn ja, bei welcher Schalterstellung das verzögerte Signal erkannt wurde, wechselte der Versuchsleiter den Wert der Latenzzeit ΔT (siehe Tab.1) auf einen weiteren zu testenden Wert. Bei jeweils fünf eingestellten Latenzzeiten im kritischen Bereich, wurde die Entscheidung durch Wiederholung überprüft.

<u>ΔT / ms</u>												
50	58	67	75	83	92	100	108	117	125	133	142	150

Tab. 1: Latenzzeiten

In Vorversuchen wurde mit verschiedenen Testsignalen (Musik, Sprache, Soloinstrumente) und unterschiedlichen Lautsprecheraufstellungen (Mono, Stereo, Surround) die kritischste Versuchsbedingung ermittelt. Als Stimulus zeigten sich die impulshaften Eigenschaften der Castagnetten (EBU SQAM; Track 27), wiedergegeben über einen Lautsprecher in 2,5 m Entfernung vor dem Kunstkopf, als besonders geeignet. Die Versuchspersonen konnten während einer Übungsphase die auftretenden Lokalisationseffekte kennenlernen.

2.4 Ergebnisse

An der Untersuchung haben 17 Testpersonen teilgenommen, die bei jeder eingestellten Latenzzeit entweder für *Unterschied wahrnehmbar* \Rightarrow „1“ oder *Unterschied nicht wahrnehmbar* \Rightarrow „0“ votiert hatten. Die Mittelwerte aller Entscheidungen, sowie der 95%-Vertrauensbereich, sind in Abhängigkeit der eingestellten Latenzzeit in *Abbildung 3* dargestellt.

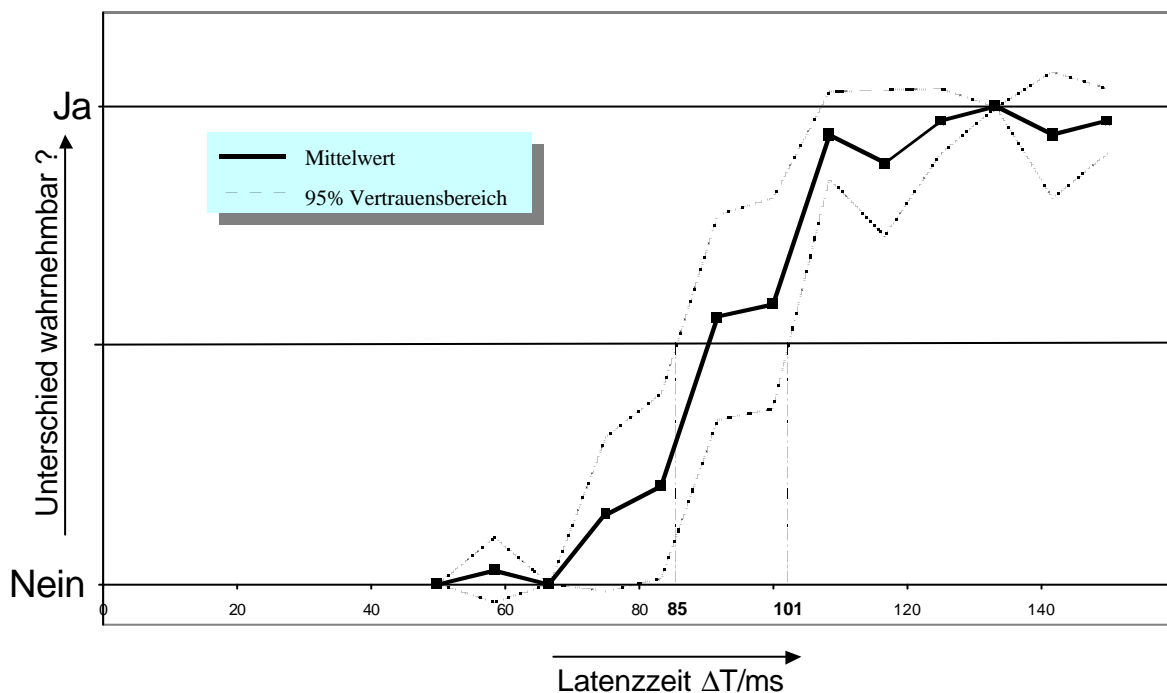


Abb. 3: Wahrnehmbarkeit der Latenzzeit; 17 Versuchspersonen

Als Wahrnehmbarkeitsschwelle wurde ein Mittelwert von 0,5 angenommen. Das bedeutet, wenn der 95%-Vertrauensbereich des Mittelwertes diese Grenze übersteigt, gilt die entsprechende Latenzzeit als wahrnehmbar.

Das Ergebnis dieser Untersuchung zeigt, daß Latenzzeiten kleiner als 85 ms von den Versuchspersonen nicht mehr wahrgenommen wurden. Bei Latenzzeiten zwischen 85 ms und 101 ms liegt der Übergang zwischen „nicht wahrnehmbar“ und „wahrnehmbar“. Ab Latenzzeiten > 101 ms sind die Lokalisierungseffekte eindeutig wahrnehmbar.

Für das Entwerfen von Auralisationssystemen mit Head-Tracking sollte also eine maximale Latenzzeit von $\Delta T_{\max.} = 85$ ms nicht überschritten werden.

3. Richtungshörversuche mit dem Kugelflächenmikrofon (KFM 6 von Schoeps)

Ein weiterer Aspekt bei der Entwicklung von Auralisationssystemen ist die Frage, in welchem Maße eine Vereinfachung der HRTF's und dadurch eine Reduktion des Rechenaufwandes möglich ist. Die folgende Untersuchung befaßt sich deshalb mit den Auswirkungen einer starken Vereinfachung der HRTF's auf die Lokalisation. Es wurde in der oben beschriebenen Kunstkopf-Tracking-Einrichtung der Kunstkopf KU100 durch ein Kugelflächenmikrofon [2] (KFM 6) ersetzt.

3.1 Versuchsaufbau

Die Hörversuche fanden im Referenz-Abhörraum des IRT und einem ruhigen Nebenraum statt (siehe Abbildung 4). Die Lautsprecheraufstellung entsprach der ITU-Rec. BS.775 (3/4 Stereoformat). Auf der optimalen Abhörposition befand sich das Kugelflächenmikrofon, welches den Kopf-Drehbewegungen der Probanden im Nebenraum nachgeführt wurde. Das Kugelflächenmikrofon befand sich auf gleicher Höhe mit den Lautsprechern (1,25 m). Die Kugelflächenmikrofonsignale wurden den Versuchsteilnehmern über einen diffusfeld-entzerrten Kopfhörer (STAX SR-Lambda) dargeboten.

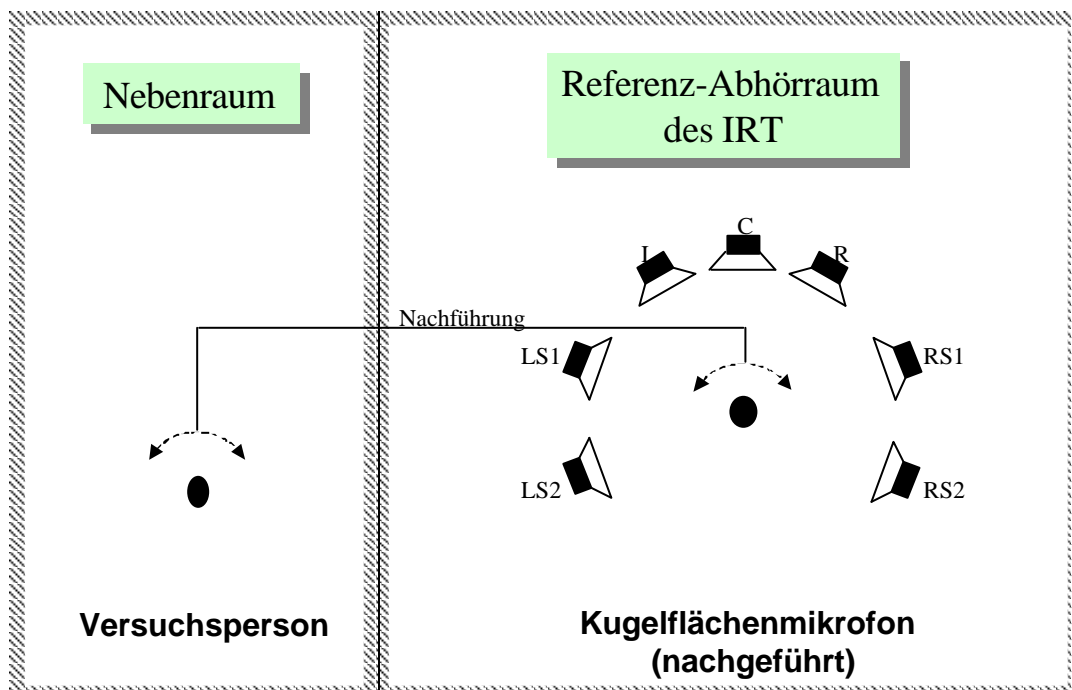


Abb. 4: Hörversuchsanordnung zur Untersuchung der Lokalisation von Hörereignissen unter Verwendung vereinfachter HRTF

3.2 Versuchsdurchführung

Am Hörversuch nahmen 14 Versuchspersonen teil. Jede absolvierte zwei Durchgänge. Im ersten Versuchsdurchgang war die Nachführung des Kugelflächenmikrofons ausgeschaltet, wodurch die Probanden die Kopf-Drehbewegung zur Lokalisation der Schallquelle nicht einsetzen konnten. Im zweiten Durchgang konnten dann die Versuchspersonen spontane Drehbewegungen des Kopfes zur Unterstützung der Lokalisation ausüben.

Als Stimulus diente männliche Sprache (EBU SQAM; Track 50). Das Testsignal wurde als direkte Schallquelle (Lautsprecher) oder als Phantomschallquelle, die durch ein Pegelverhältnis (2:1, 1:1 oder 1:2) zweier benachbarter Lautsprecher zustande kam, den Probanden dargeboten. Insgesamt waren 30 Schallquellen zu lokalisieren, wobei davon drei zweimal vorkamen, um die Zuverlässigkeit der Versuchsperson überprüfen zu können.

Vor den eigentlichen Versuchsdurchgängen absolvierte jeder Proband eine kurze Trainingsphase, um sich auf die unterschiedlichen Abhörbedingungen einzustellen.

Während des eigentlichen Versuchs wurde das Testsignal dann für jede zu überprüfende Richtung (zufällige Reihenfolge) zweimal hintereinander dargeboten, so daß den Probanden ausreichend Zeit blieb, den Hörereignisort durch Azimut- und Elevationswinkel, auf den ihnen vorliegenden Fragebögen zu markieren.

3.3 Ergebnisse

3.3.1 Richtungshören

In *Abbildung 5* und *Abbildung 6* sind die Ergebnisse des Richtungshörversuches mit nicht nachgeführtem und nachgeführtem Kugelflächenmikrofon dargestellt. Auf der Ordinatenachse sind der Azimut des Hörereignisortes und auf der Abszissenachse der Azimut des Schallereignisortes in Grad, sowie die dazugehörigen Lautsprecherpositionen aufgetragen. Die senkrechten, durchgezogenen Linien markieren die realen Lautsprecherpositionen und zwei besondere Phantomschallquellen (RS2-LS2, L-R).

Bei exakter Lokalisation der Schallquellen, würden sich die Richtungsurteile auf einer Geraden von „unten links“ nach „oben rechts“ befinden.

Starke Abweichungen von der Ideallinie sind in *Abbildung 5* zu erkennen. Wie zu erwarten, treten für das starre Kugelflächenmikrofon in hohem Maße vorn-hinten-Vertauschungen auf, besonders im Bereich der Medianebene (0° bzw. 360° und 180°).

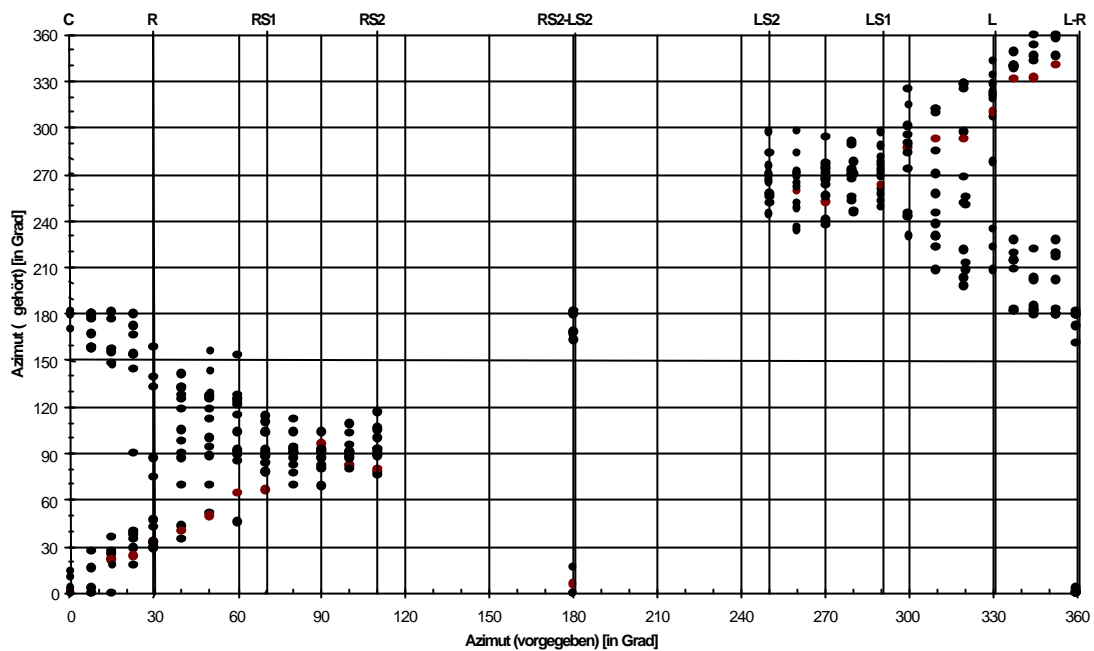


Abb. 5: Wahrgenommene Richtungen mit nicht nachgeführtem Kugelflächenmikrofon

Dagegen sind die „vorn-hinten-Vertauschungen bei unterstützenden Kopf-Drehbewegungen weitgehend verschwunden (Abbildung 6). Geblieben ist nur eine geringe Streuung der Richtungsurteile.

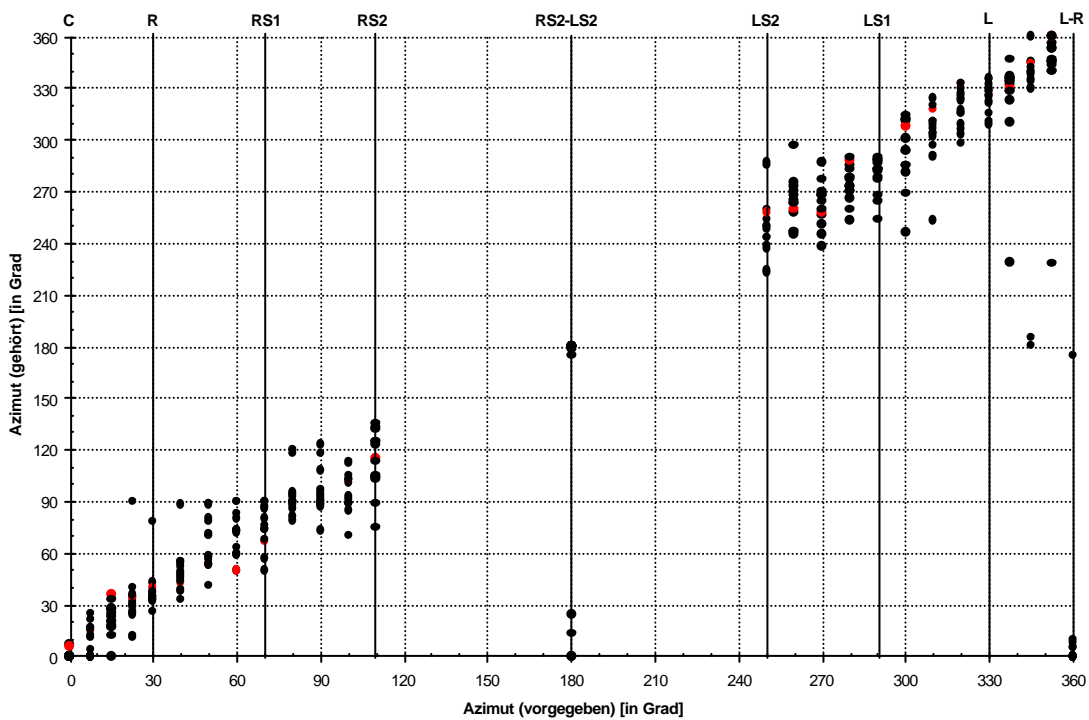


Abb. 6: Wahrgenommene Richtungen mit nachgeführtem Kugelflächenmikrofon

3.3.2 Elevation

In *Abbildung 7* und *Abbildung 8* sind die Ergebnisse der wahrgenommenen Elevation des Hörereignisortes in Abhängigkeit des Azimut der Schallereignisorte, bei nicht nachgeführtem und nachgeführtem Kugelflächenmikrofon dargestellt. Die Mittelwerte der Elevation bei jeder Schalleinfallrichtung sind durch kleine Quadrate gekennzeichnet. Eine angegebene Elevation von $+90^\circ$ bedeutet, das die Schallquelle über dem Kopf wahrgenommen wurde. Eine Elevation von 0° bedeutete, daß das Testsignal in Ohrhöhe wahrgenommen wurde.

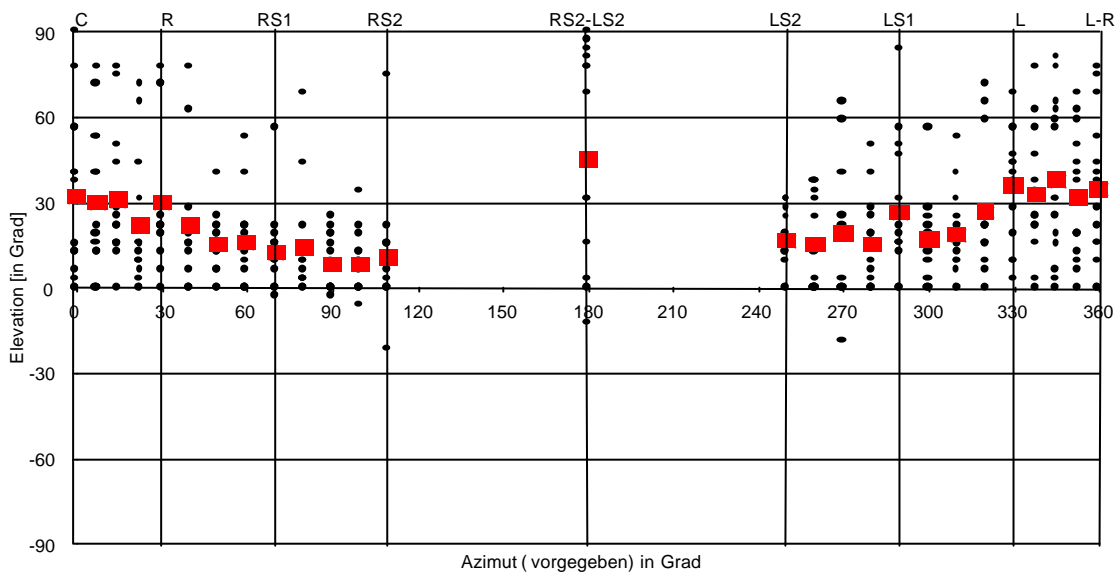


Abb. 7: Wahrgenommene Elevation mit nicht nachgeführtem Kugelflächenmikrofon

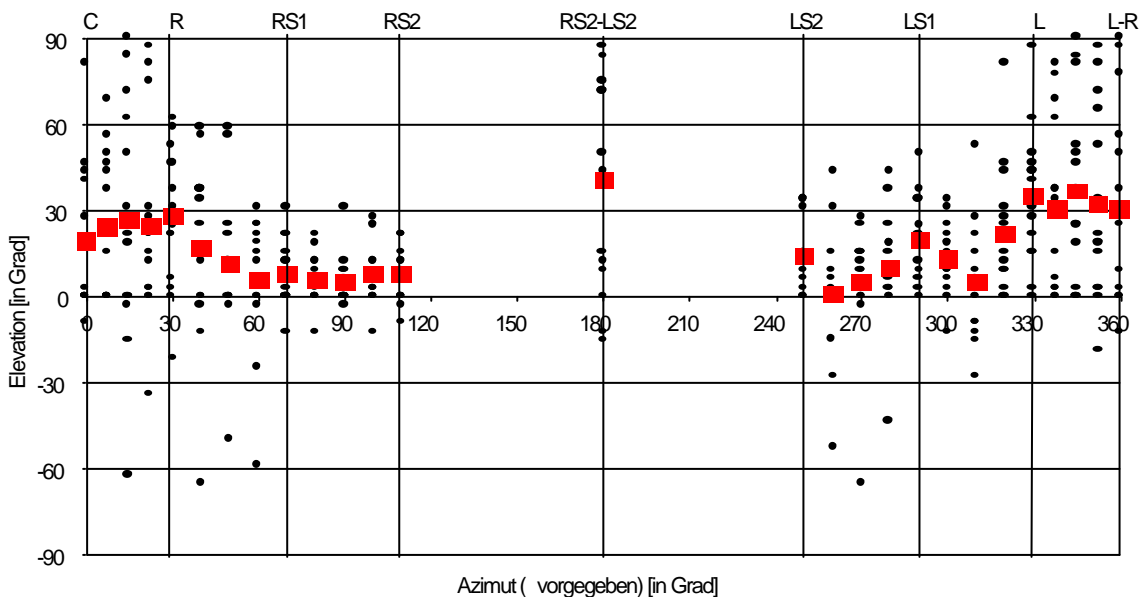


Abb. 8: Wahrgenommene Elevation mit nachgeführtem Kugelflächenmikrofon

Man erkennt, daß gravierende Elevationseffekte insbesondere im Bereich der Medianebene auftreten, und zwar unabhängig davon, ob das Kugelflächenmikrofon nachgeführt wird oder nicht.

4. Zusammenfassung

Hochwertige Auralisationsverfahren erfordern die Anwendung von Head-Tracker Systemen und damit einen hohen Aufwand bei der Berechnung der binauralen Signale. Ein wichtiger Faktor ist die erforderliche Rechengeschwindigkeit. Um sie zu ermitteln, wurde im Teil 1 der Untersuchungen die maximal zulässige Latenzzeit von Auralisationssystemen ermittelt. Es zeigte sich, daß Latenzzeiten von **DT < 85 ms** nicht wahrgenommen werden, auch nicht unter kritischen Versuchsbedingungen (impulshaftes Signal von vorne).

Teil 2 der Untersuchungen betraf die Frage, in welchem Maße die Anwendung genauer (z.B. individueller) HRTF's lohnend ist, wenn im Fall der Head-Tracker Systeme die dynamischen Merkmale die Lokalisation unterstützen. Richtungshörversuche mit einem Kugelflächenmikrofon erbrachten die folgenden Ergebnisse:

Die Auswertung der dynamischen Merkmale verhindert das Auftreten von vorn-hinten-Vertauschungen auch dann, wenn für die Medianebene keine spektralen Richtungsmerkmale im binauralen Signal enthalten sind. Jedoch läßt sich das Phänomen der Elevation im Bereich der Medianebene dadurch nicht vermeiden. Die exakte Lokalisation im Bereich der Medianebene setzt voraus, daß neben den dynamischen Merkmalen auch die relevanten spektralen Merkmale der binauralen Signale vorhanden sind. Weiterführende Untersuchungen werden zeigen, mit welcher Genauigkeit diese „richtungsbestimmenden Bänder“ [3] reproduziert werden müssen.

Literatur

- [1] MACKENSEN, Philip / REICHENAUER, Klaus / THEILE, Günther (1998): Einfluß der spontanen Kopfdrehungen auf die Lokalisation beim binauralen Hören, in diesem TMT-Tagungsband
- [2] Theile, Günther (1986): Das Kugelflächenmikrofon. 14. Tonmeistertagung München, 1986, VDT Tagungsbericht, S. 277-293
- [3] Blauert, Jens (1969/70): Sound localization in the median plane. Acustica 22, S. 205 - 213