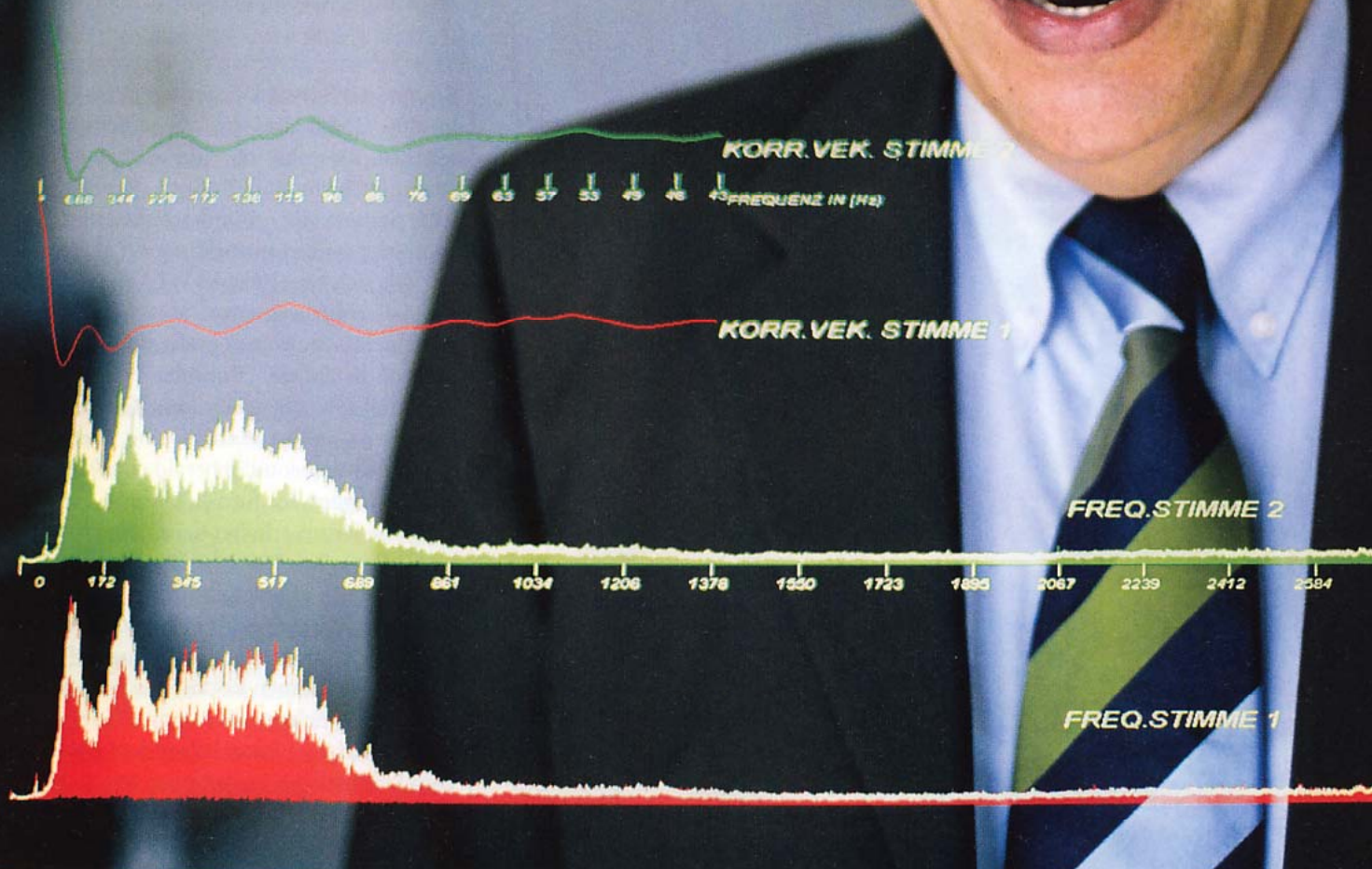


Quereinsteiger: Ein Bekannter weckte bei Sameh Rahman das Interesse an Verfahren zur Identifizierung von Stimmen



Kriminelle Schwingungen

Ein Physiker ist mit einem neuen Verfahren in die Domäne der kriminalistischen Stimm-Identifizierung eingedrungen. Er trifft auf Unglauben – und ist womöglich doch auf der richtigen Spur

VON VERONIKA SZENTPÉTERY; MANFRED KLIMEK (FOTOS)

Stalker, die ihre Opfer telefonisch belästigen, haben gegen die TV-Kriminalisten beim CSI keine Chance: Die Forensik-Experten befreien die Stimmufnahmen mit ein paar Mausklicks von Hintergrundgeräuschen, vergleichen sie mit einer Sprechprobe von Verdächtigen und haben den Täter dann rasch identifiziert. Doch was im Fernsehen problemlos funktioniert, bereitet echten Experten Kopfzerbrechen: Der Identifikation von Menschen anhand von Stimmproben sind im richtigen Leben Grenzen gesetzt.

Diese Grenzen hat Sameh Rahman, Physikprofessor an der Universität Hannover, nach eigenen Angaben ein ganzes Stück weit verrückt: Basierend auf Erkenntnissen der Strukturforschung hat er ein physikalisches Modell für die Analyse der menschlichen Stimme entwickelt, mit dem sich Störgeräusche wie Motorenlärm in Stimmufnahmen elegant herausfiltern lassen. Obendrein soll Rahmans Software in der Lage sein zu erkennen, ob zwei Proben vom selben Sprecher stammen – nützlich nicht nur für Kriminalisten, sondern auch für Telefon-Hotlines von Banken, die wissen müssen, ob sie tatsächlich einen Kunden am Hörer haben. Altgediente Phonetiker allerdings wollen dem Fachfremden nicht glauben, dass er ein Problem gelöst hat, mit dem sie seit Jahren kämpfen.

Der aus Ägypten stammende Rahman ist eigentlich Experte für die Untersuchung von Strukturen bei kristallinen Feststoffen. Doch als ihn vor Jahren ein Bekannter um die Übersetzung einer verrauschten arabischen Tonaufnahme bat, packte ihn der Ehrgeiz: Ihm fiel

auf, dass Stimmen und Kristalle Gemeinsamkeiten haben. Die dreidimensionalen Strukturen von kristallinen Stoffen wie Kochsalz haben oft Unregelmäßigkeiten, wenn etwa die Abfolge der Atome von Fremdatomen unterbrochen wird. Ebenso wechseln bei menschlichen Stimmen periodische und unperiodische Frequenzanteile, was man laut Rahman als eindimensionale Struktur ansehen kann.

Zu den periodischen Anteilen gehören die Frequenzen der einzelnen Laute. Jede Kombination mehrerer Laute aber verursacht eine Art Rauschen – die unperiodischen Frequenzanteile. Beide Anteile ergeben bei jedem Menschen eine charakteristische Mischung, die Rahman mit Hilfe eines mathematischen Verfahrens erfasst. Auf ähnliche Weise isoliert er das Frequenzmischprofil von Störgeräuschen und filtert sie aus Aufnahmen heraus.

Beim zweiten Einsatzgebiet seines Modells, der Stimmidentifizierung, digitalisiert Rahman Referenzprobe und Tatprobe. Dann schlüsselt er mit einem mathematischen Verfahren auf, welche Frequenzen in den Proben vorkommen. Dieses Frequenzspektrum wird in 24 so genannte kritische Frequenzbänder zerlegt – in Bereiche also, in denen die häufigsten und damit am deutlichsten zu hörenden Frequenzen stecken; eine ähnliche Unterteilung nutzt auch unser Ohr beim Hören. Dann werden die Frequenzbänder beider Proben miteinander verglichen. Da aber ein reiner Frequenzspektrum-Vergleich keine eindeutigen Ergebnisse liefern würde, hat Rahman einen zusätzlichen mathematischen Kniff ent-

**WAS IM FERN-
SEHEN SCHNELL
GEHT, IST IM
ECHTEN LEBEN
KOMPLIZIERT**

www.min.uni-hannover.de/mitarbeiter/srahrman.htm

Sameh Rahman nutzt Erkenntnisse aus der Psychoakustik und der Strukturforschung zum Filtern und Vergleichen von Stimmufnahmen

RAHMANS VERFAHREN KLINGT WIE DER TRAUM EINES JEDEN KRIMINALISTEN

wickelt, um der Stimme weitere Merkmale zu entlocken. Die Software liefert anschließend einen Übereinstimmungsfaktor: Ab 80 Prozent geht Rahman von einer Stimmidentität aus, zwischen 50 und 80 Prozent sind weitere Tests nötig, unterhalb von 50 Prozent schließt er eine Übereinstimmung aus.

Mit seinem Verfahren ist Rahman als Gutachter bereits gefragt. So lieferte er entscheidendes Beweismaterial gegen zwei Kokain-Dealer: Seine Software zeigte nach Ansicht des Gerichts, dass es tatsächlich diese beiden waren, deren Stimmen auf abgehörten Telefonaten über Drogen-Deals zu hören waren.

„Wir erreichen beim Vergleich der Stimmen nahezu die Genauigkeit des menschlichen Gehörs“, sagt Rahman selbstbewusst. Und anders als eine Auswertung durch einen Phonetiker unterliege seine Methode keinem subjektiven Einfluss. Sie sei unabhängig von Sprache, Grammatik und Sprechrhythmus und brauche auch keine Proben mit dem gleichen

der Vergleich von Frequenzspektren „Stand von vor 40 Jahren“. Ähnlich kritisch äußert sich Angelika Braun, Phonetik-Professorin an der Universität Marburg und zuvor Leiterin von Kösters Fachbereich im BKA: „Mir stellen sich die Nackenhaare hoch, wenn jemand sich weigert, mit seinem Verfahren an Qualitätsprüfungen durch die Wissenschaftsgemeinde teilzunehmen.“ Für sie zählen bei der Stimmidentifizierung Merkmale wie Grundfrequenz, Obertöne, Sprechtempo, Stimmlage und Dialekt – und da führe bei aller elektronischen Unterstützung noch kein Weg an einem geschulten Gehör vorbei.

Allerdings ist Rahman nicht der einzige, der den menschlichen Faktor überflüssig machen will. Phonetik-Forscher aus aller Welt arbeiten an Projekten zur vollautomatischen Stimmidentifizierung – basierend auf einer vom Frequenzspektrum abgeleiteten Größe. Ebenso die Voice.Trust AG in München: Sie verkauft bereits ein System, das Stimmen eindeutig identifiziert – allerdings sind dafür noch gespeicherte Textproben erforderlich, die ein zu authentifizierender Anrufer wortgetreu zu wiederholen hat. Voice.Trust arbeitet aber nach eigener Aussage an Software, die Stimmen auch durch den Abgleich mit Proben beliebigen Inhalts identifizieren kann.

Details will man auch hier nicht verraten, doch zumindest Rahman muss ohnehin bald mit genaueren Informationen herausrücken: Sein Verfahren soll im Rahmen eines Projektes der Deutschen Forschungsgemeinschaft (DFG) auch extrem teure industrielle Diamantsägen schützen können – es soll warnen, wenn es die typische Veränderung im Säegeräusch erkennt, das einige Zeit vor dem Riss der Säge auftritt. Und die DFG-Richtlinien sehen vor, dass Methoden und Ergebnisse in regelmäßigen Berichten publiziert werden. 



Mehrplatzabhörsystem: Bis zu vier Ermittler können sich Stimmaufnahmen anhören

Text; die Qualität von mitgeschnittenen Telefonaten reiche für aussagekräftige Ergebnisse.

Tatsächlich klingt das wie der Traum aller Ermittler – doch die Phonetik-Experten unter ihnen stören sich daran, dass Rahman wichtige Elemente seiner Methode unter Verschluss hält und sich auch nicht an internationalen Vergleichstests für derartige Systeme beteiligt.

„Es gibt keinen akustischen Fingerabdruck“, sagt Olaf Köster vom Fachbereich Sprechererkennung und Tonträgeranalyse des Bundeskriminalamts (BKA). Eine Stimme sei gar nicht anders als subjektiv zu beurteilen,