

Zeitschrift: Horizons : le magazine suisse de la recherche scientifique
Herausgeber: Fonds National Suisse de la Recherche Scientifique
Band: 29 (2017)
Heft: 113

Artikel: Une machine à l'écoute
Autor: Gaitzsch, Sophie
DOI: <https://doi.org/10.5169/seals-821711>

Nutzungsbedingungen

Die ETH-Bibliothek ist die Anbieterin der digitalisierten Zeitschriften auf E-Periodica. Sie besitzt keine Urheberrechte an den Zeitschriften und ist nicht verantwortlich für deren Inhalte. Die Rechte liegen in der Regel bei den Herausgebern beziehungsweise den externen Rechteinhabern. Das Veröffentlichen von Bildern in Print- und Online-Publikationen sowie auf Social Media-Kanälen oder Webseiten ist nur mit vorheriger Genehmigung der Rechteinhaber erlaubt. [Mehr erfahren](#)

Conditions d'utilisation

L'ETH Library est le fournisseur des revues numérisées. Elle ne détient aucun droit d'auteur sur les revues et n'est pas responsable de leur contenu. En règle générale, les droits sont détenus par les éditeurs ou les détenteurs de droits externes. La reproduction d'images dans des publications imprimées ou en ligne ainsi que sur des canaux de médias sociaux ou des sites web n'est autorisée qu'avec l'accord préalable des détenteurs des droits. [En savoir plus](#)

Terms of use

The ETH Library is the provider of the digitised journals. It does not own any copyrights to the journals and is not responsible for their content. The rights usually lie with the publishers or the external rights holders. Publishing images in print and online publications, as well as on social media channels or websites, is only permitted with the prior consent of the rights holders. [Find out more](#)

Download PDF: 11.02.2026

ETH-Bibliothek Zürich, E-Periodica, <https://www.e-periodica.ch>

Une machine à l'écoute

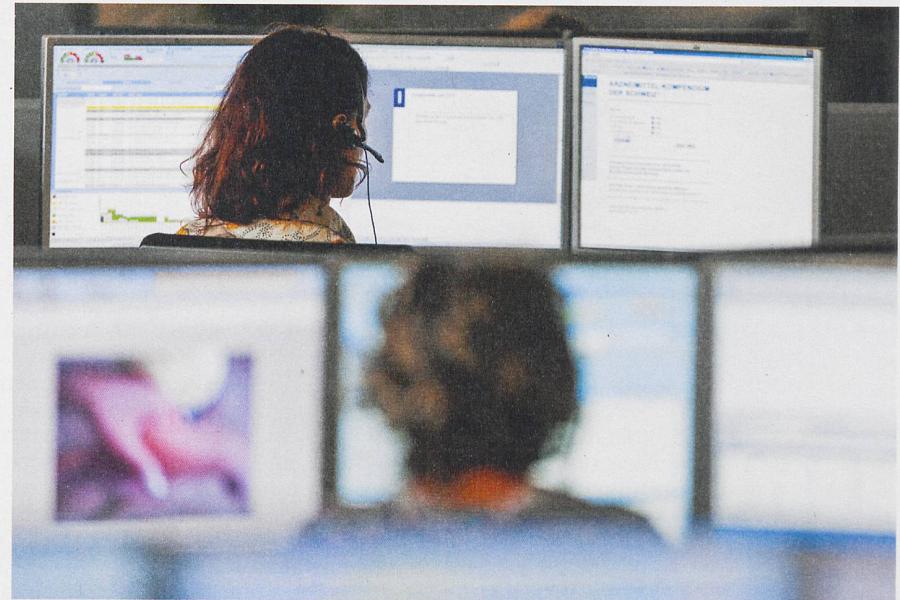
Des algorithmes apprennent à reconnaître les émotions dans notre voix. Principal défi: fonctionner en dehors du laboratoire. *Par Sophie Gaitzsch*

C'est souvent un serveur vocal qui nous répond lorsqu'on appelle sa banque ou son assurance. Pendant de longues minutes, il nous demande de choisir entre des options qui ne semblent pas toujours des plus pertinentes. Résultat: on soupire, on marmonne quelques jurons, on interpelle le système. Et là, comme par enchantement, un opérateur prend l'appel. Un hasard? Pas forcément: des programmes dopés à l'intelligence artificielle sont désormais capables de reconnaître les émotions, ce qui permet aux entreprises d'adapter leurs services en conséquence.

Sascha Frühholz, professeur de psychologie spécialisé dans les neurosciences à l'Université de Zurich, travaille sur la détection automatique des émotions dans la voix: «Les algorithmes sont de plus en plus efficaces, surtout pour discerner les six émotions de base que sont la colère, la peur, la joie, le dégoût, la tristesse et la surprise. Percevoir la honte ou la fierté est plus complexe pour les ordinateurs. Comme pour les humains d'ailleurs.»

Le principal défi pour ces systèmes est de se montrer suffisamment généralistes. «Ils sont entraînés sur des données spécifiques, explique Sascha Frühholz. Leur performance baisse lorsqu'ils travaillent dans un environnement acoustique différent ou dans une autre langue. Si l'algorithme apprend à reconnaître la colère dans des voix zurichoises, il obtiendra probablement de moins bons résultats avec des voix genevoises. Sa pertinence continuera de diminuer face à une langue tonale asiatique, dont le profil acoustique est plus éloigné encore.»

Pour surmonter cet écueil, Sascha Frühholz a mêlé des techniques d'apprentissage supervisé et non supervisé. «Concrètement, nous avons entraîné l'algorithme avec des données labellisées qui indiquent que telle voix est en colère ou joyeuse. Nous avons ensuite introduit des données non labellisées, ce qui a permis de le rendre plus indépendant.»



Contrairement à ces téléphonistes, les serveurs vocaux automatiques ne perçoivent pas l'impatience des clients – du moins pour l'instant. Photo: Keystone/Martin Ruetschi

Le taux de reconnaissance atteint 63%, sensiblement meilleur qu'avec l'utilisation indépendante de l'apprentissage supervisé ou non supervisé (taux entre 54% et 58%). Les êtres humains sont, eux, capables de détecter de manière correcte une émotion dans la voix d'un interlocuteur dans 85 à 90% des cas, note Sascha Frühholz, qui travaille également sur la perception auditive humaine. «Le niveau de reconnaissance dépend – autant pour les algorithmes que pour les humains – en grande partie du nombre d'émotions qui doivent être distinguées simultanément.»

Déceler la dépression

La reconnaissance automatique des émotions trouve des applications potentielles dans de nombreux domaines: services à la clientèle, marketing, surveillance, aide aux personnes âgées ou encore médecine. «Elle pourrait notamment aider à repérer les premiers signes d'une crise d'angoisse ou d'une dépression», indique David Sander, directeur du Pôle de recherche national «Sciences affectives» ainsi que du centre interfacultaire dédié à ce domaine à l'Université de Genève.

A l'EPFL, Jean-Philippe Thiran mène des recherches sur la reconnaissance visuelle des expressions du visage, en partenariat avec l'industrie automobile, un autre secteur qui s'intéresse de près à la détection des émotions. «L'objectif consiste à ras-

sembler des informations sur la personne au volant. Dans un contexte de voiture semi-autonome par exemple, il est primordial de savoir dans quel état émotionnel l'automobiliste se trouve au moment où le véhicule doit lui repasser la main: s'il est stressé, s'il est apte à prendre une décision.» La voiture pourrait diffuser une musique apaisante à un conducteur énervé ou intensifier la luminosité du tableau de bord en cas de coup de fatigue.

«La détection des expressions faciales dans des conditions peu favorables, lorsque le visage est mal éclairé, qu'il bouge ou qu'il n'est pas de face, constitue aujourd'hui les principaux enjeux de recherche, note Jean-Philippe Thiran. Tout comme la reconnaissance d'expressions plus subtiles ou individuelles.» Les algorithmes progressent rapidement, mais ne sont pas encore capables d'interpréter dans une situation réaliste toutes les émotions exprimées par notre voix et notre visage.

Sophie Gaitzsch est une journaliste suisse basée à Paris.