



Maschinen lernen sehen *Machines learn how to see*

Wie funktioniert Sehen? Wie unterscheidet man gute von schlechten Bildern? Wie filtert man aus Bildsignalen jene Informationen, die für das Erkennen wesentlich sind? Diese Fragen beschäftigen sowohl Hirnforscher/innen als auch Computerwissenschaftler/innen wie Thomas Pock, die in internationalen Kooperationen die Bildverarbeitung vorantreiben.

Thomas Pock hat seit 2014 eine AIT-Stiftungsprofessur für „Mobile Computer Vision“ am Institut für Maschinelles Sehen und Darstellen (ICG) inne. In seinen Forschungsarbeiten, bei denen er unter anderem mit Kollegen in New York und Paris kooperiert, beschäftigt er sich mit mathematischen Modellen, um „gute“ von „schlechten“ Bildern zu unterscheiden. Im Einzelnen geht es darum, aus den Bildsignalen, die Maschinen liefern, jene visuellen Informationen zu filtern, mit denen sich ein möglichst detailreiches bzw. aussagekräftiges Bild rekonstruieren lässt.

Forschungskooperation mit New York

In einer seiner aktuellen Arbeiten entwickelt Thomas Pock mathematische Methoden, um aus Magnetresonanztomographie-Signalen (MRT-Signalen) zweidimensionale Bilder zu rekonstruieren. Das Ziel: Mit so wenig Signaldaten wie möglich zu bestmöglichen Ergebnissen zu kommen. Denn je weniger Daten man benötigt, umso kürzer sind die Scanzeiten im MRT, umso mehr Patientinnen und Patienten können an einem Tag auf einer Maschine untersucht werden. Das senkt die Kosten.

In einer Forschungskooperation mit Florian Knoll und Daniel K. Sodickson

How does seeing work? What is a good image, and what is a bad image? How do we filter out the essential information from an image – the information we need to recognise what we see? These are key questions for brain researchers as well as computer scientists, such as Thomas Pock, who is always looking for international cooperation to achieve progress in image processing.

Since 2014, Thomas Pock has held an AIT-endowed professorship for Mobile Computer Vision at the Institute for Computer Graphics and Vision (ICG). His research work, developed in co-operation with colleagues from New York and Paris, focuses on mathematical models to distinguish between “good” images and “bad” images. Ultimately the objective is to filter machine-supplied image signals, extracting only the visual information that is absolutely essential for the reconstruction of a meaningful image with maximum detail.

Research cooperation with New York

One of Thomas Pock's current projects is to build mathematical models for the reconstruction of two-dimensional images from magnetic resonance imaging signals (MRI signals). His aim is to get the best possible result from the smallest possible amount of signal data. If you need less data, you automatically shorten the scanning time in the MRI, which in turn allows you to increase the number of patients who can be scanned with the machine in one single day and thus reduces costs.

In a research co-operation with Florian Knoll and Daniel K. Sodickson from the Department of Radiology at New York

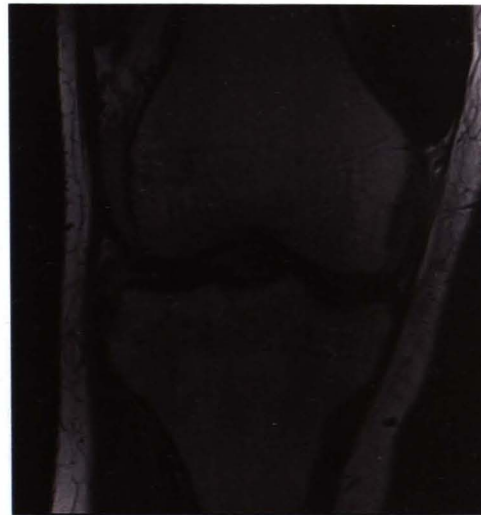
vom Department of Radiology an der New York University – School of Medicine hat Thomas Pock zusammen mit seiner PhD-Studentin Kerstin Hammernik einen Algorithmus entwickelt, der genau diese Aufgabe erfüllt: qualitativ hochwertige Bilder aus den unterabgetasteten MRT-Signaldaten zu rekonstruieren, die in lediglich einem Sechstel der Scanzeiten von bisherigen MR-Scans hergestellt werden können. Von den Forschungspartnern in den USA stammen die MR-Daten, die zur Entwicklung der Methode benötigt wurden, und das Know-how der Physik und Arbeitsweise der MR-Geräte. Thomas Pock und Kerstin Hammernik haben in Graz das mathematische Modell entwickelt, das aus den unterabgetasteten MRT-Signaldaten die Bilder rekonstruiert.

Ein Grundproblem der maschinellen Bilderkennung besteht in der schier unermesslichen Anzahl möglicher Bilder. Thomas Pock: „Wenn man alle theoretischen Bildvariationen allein bei einer Bildgröße von nur 65 mal 65 Pixel und einem Tonumfang von lediglich 256 Graustufen errechnet, dann ergeben sich daraus potenziell viel mehr verschiedene Bilder, nämlich $256^{65 \times 65} \approx 10^{10000}$, als es schätzungsweise Atome im Universum gibt, nämlich 10^{80} .“ Diese unvorstellbar große Zahl an möglichen Bildern erklärt, warum einfache Bildvergleichsdatenbanken zwangsläufig versagen müssen.

Maschinelles Lernen in Fusion mit Bildverarbeitung

„Als Menschen wissen wir, wie Bilder mit hoher Abbildungsqualität aussehen“, führt Thomas Pock aus. „Und das ist es auch, was wir Computern beibringen wollen: dass sie innerhalb von Sekundenbruchteilen Bilder erkennen und einordnen können.“ Während man früher quasi per Hand nach geeigneten mathematischen Modellen gesucht hat, um diese Aufgabe rechnerisch zu bewältigen, kombiniert Thomas Pock neue Methoden des maschinellen Lernens mit solchen der Bildverarbeitung. Konkret entwirft er Bildmodelle mit sehr vielen Freiheitsgraden, die – im ständigen Abgleich mit dem Idealbild – aus den Signaldaten zweidimensionale Bilder rekonstruieren können.

Thomas Pock: „In dem Lernproblem gibt es eine Loss-Funktion, die den Verlust feststellt und berechnet, wie sehr die momentan rekonstruierte Lösung von der Zielvorstellung abweicht – also davon, wie das Gewebe aussieht, das sich >



© TU Graz, data courtesy of Florian Kroil, NYU

Abbildung 1:
Schichtbild einer klinischen
MRT-Aufnahme eines Knies.
Figure 1:
Slice of the clinical MRI scan
of a knee.

University School of Medicine, Thomas Pock and his PhD student Kerstin Hammernik developed an algorithm that does precisely that: it builds high-quality images from the undersampled MRI signal data, taking just one sixth of the scan time of previous MRI scans. The research partners in the USA provide the MRI data that are needed to develop this method as well as the know-how in physics and the operating principles of the MRI devices, while Thomas Pock and Kerstin Hammernik devised the mathematical model that reconstructs the images from the undersampled MRI signal data in Graz.

One fundamental problem of machine-based image recognition is the sheer endless number of possible images. "If you calculate all theoretical image variations for an image size of no more than 65 times 65 pixels and 256 grey scales, you obtain potentially many more different images than there are atoms in the universe, namely $256^{65 \times 65} \approx 10^{10000}$ versus 10^{80} ." This unimaginably large number of possible images explains simple comparison-based databases cannot work," explains Thomas Pock.

Machine learning combined with image processing

"As human beings we know what high-quality images look like," says Thomas Pock. "Now we want computers to do the same – to be able to recognise and classify images within fractions of a second." In the past, we used to look for mathematical models capable of coping with this massive computing task virtually by hand. Now Thomas Pock combines new methods of machine learning with image processing methods. He designs image models with a large number of degrees of freedom that are able >

in den MRT-Daten spiegelt. Die Loss-Funktion stellt den Unterschied zum Soll fest und propagiert den Fehler zurück in das Modell. Das geschieht über die Berechnung des Gradienten der Verlustfunktion, der in die Richtung der stärksten Änderung zeigt. Auf diese Weise kann man die Modellparameter in eine Richtung lenken, sodass der Fehler kleiner wird. Das macht man so lange, bis keine Verbesserung mehr feststellbar ist.“

to reconstruct two-dimensional images from the signal data, constantly comparing the result with the ideal image.

Thomas Pock adds: “The learning problem involves a loss function that measures the loss and calculates how much the momentarily reconstructed solution differs from the target image, i.e. the appearance of the tissue as reflected in the MRI data. This error

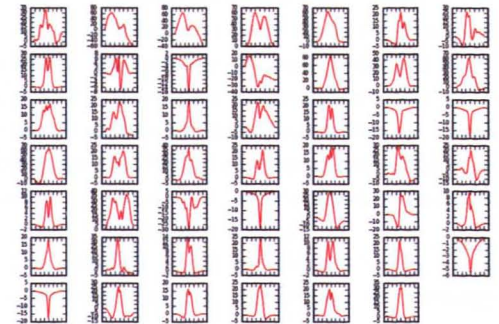
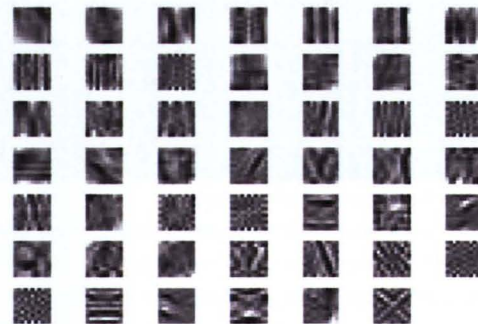


Abbildung 2:
Auszug aus den gelernten Modellparametern.

Links die Filterkerne und rechts die Bewertungsfunktionen.

Figure 2:

Extract from the learned model parameters. Filter cores on the left, evaluation functions on the right.

Abbildung 2 zeigt einen Auszug aus den gelernten Modellparametern, die im Prinzip aus einer großen Anzahl von verschiedenen Filterkernen und Bewertungsfunktionen bestehen.

is in turn back propagated into the model. This is done by calculating the gradient of the loss function which points in the direction of the strongest change. In this manner the model parameters can be varied to make the mistake smaller. Then you let the process continue until no detectable improvement can be achieved any more.”

Figure 2 shows an extract from the learned model parameters which basically consist of a large number of different filter cores and evaluation functions.

Der Designschritt in den von Pock und Hammernik entworfenen mathematischen Modellen mit ihren Tausenden von freien Parametern ist von neuronalen Netzwerken inspiriert und basiert auf den Erkenntnissen aus mehr als 50 Jahren Forschung. Ausgeführt wird die Berechnung auf einem Hochleistungsrechner der TU Graz an der TU Graz. Für die Forschungsarbeiten am TU Graz-Institut wurde in Kooperation mit dem Zentralen Informatikdienst (ZID) ein Großrechner angeschafft, der mit 16 der leistungsfähigsten Grafikkarten bestückt ist, von denen jede rund vier TeraFLOPS an Rechenleistung aufweist, das heißt, jede dieser Grafikkarten kann ca. vier Billionen (4×10^{12}) Rechenoperationen pro Sekunde ausführen.

The design step in Pock and Hammernik’s mathematical models, with its thousands of free parameters, is inspired by neuronal networks and based on the findings of more than 50 years of research. The calculation is carried out on a high performance computer of TU Graz. For the research work at the TU Graz institute it was decided in co-operation with the Central IT Service to purchase a supercomputer equipped with 16 of the most powerful graphic boards, each having a processing power of approximately four TeraFLOPS, meaning that each of these graphic boards is able to perform approximately four billion (4×10^{12}) computer operations per second.

US-Patent

„Mit unserer Methode haben wir es geschafft, in sehr kurzer Zeit MRT-Aufnahmen zu rekonstruieren, die eine ähnlich gute Qualität haben wie die derzeit erzeugten Bilder, aber wir brauchen dafür nur ein Sechstel der Aufnahmezeit“, freut sich Pock.

US patent

Pock’s method reconstructs MR images of comparable quality to that of currently generated images, but in a very short time – it only needs one sixth of

Abbildung 3 zeigt den klaren Vorteil der gelernten Methode bei sechsfach beschleunigter Scanzeit.

Ein US-Patent für diese neue Methode ist vorangemeldet, auch ein Produzent der MR-Scanner zeigt bereits starkes Interesse.

Aufsatz in „Acta Numerica“

Seine bisherigen theoretischen Erkenntnisse hat Thomas Pock vor Kurzem in einem ausführlichen Review-Artikel festgehalten, den er zusammen mit Antonin Chambolle, Professor am Centre de Mathématiques Appliquées an der Ecole Polytechnique in Paris, verfasst hat. Titel der Arbeit: „An Introduction to Continuous Optimization for Imaging“. Erscheinen wird die Arbeit in der derzeit renommiertesten Zeitschrift für Mathematik, dem in Cambridge herausgegebenen Journal „Acta Numerica“.

Text: Werner Schandor ■

the sampling time. Figure 3 shows the clear advantage of the learned method, in which the sampling time is reduced by a factor of six. A US patent for this new method is pending, and one particular manufacturer of MRI scanners has already shown strong interest.

Paper in Acta Numerica

Thomas Pock recently summarized his theoretical insights in a comprehensive review article written jointly with Antonin Chambolle, professor at the Centre de Mathématiques Appliquées at the Ecole Polytechnique in Paris. His paper with the title “An introduction to continuous optimization for imaging” will be published in the Cambridge journal “Acta Numerica”, currently the world’s top-cited journal in the field of mathematics.

Text: Werner Schandor ■

Abbildung 3:

Rekonstruktion eines MRT-Schichtbildes mit Beschleunigungsfaktor 6. Links das Ergebnis einer herkömmlichen Methode, die zu starken Artefakten führt, in der Mitte die Rekonstruktion mit der entwickelten Methode und rechts als Vergleich die Rekonstruktion aus den vollständigen Daten.

Figure 3:

The reconstruction of an MRI slice with an acceleration factor of 6. On the left, the result using a traditional method, which here leads to too many artifacts. In the centre, a reconstruction using the developed method; on the right, a reconstruction from the complete data as a comparison.

