

Optische Mustererkennung im Kanal-TV – Auf dem Weg von inhomogen zu homogen?

Marco Deubler

07.06.2023, 10:35 Uhr, aktualisiert 22.11.2023, 12:57 Uhr

FÜSSEN

Nachdem in Teil 1 der Fachartikelserie „KI-unterstützte Zustandsbewertung von Abwassersystemen“ die Funktionsweise der künstlichen Intelligenz beleuchtet sowie eine Einordnung in Workflow, Chancen und Grenzen vorgenommen wurde, erläutert Teil 2, was große Herausforderungen bei der Entwicklung einer KI-Bildererkennung für Kanalinspektionen sind.



Abbildung 1: Lageversatz mit Infiltration im Anschlussrohr, außerhalb der relevanten Systemgrenze | Foto: ISAS GmbH

Ein KI-System für Kanalzustandskodierung wird bekanntlich trainiert, indem ihm ein Experte die richtige Zuordnung zwischen repräsentativen Kanalvideos und entsprechender Zustandsbeschreibung vorgibt. Mit diesem „Wissen“ reproduziert das KI-Tool in seiner Anwendung dann die passende Zustandsfeststellung für neue Bilddaten mittels optischen Abgleichs. Ein derartiges System kann nur Zusammenhänge in den Anwendungsdaten finden, für die es in den Trainingsdaten geschult wurde. Daher erzielt es deutlich bessere Erkennungsquoten, wenn die betrachtete Umgebung viele wiederkehrende Muster, einheitliche Randbedingungen und ein Minimum an Sonderfällen aufweist, die nicht in den Trainingsdaten inbegriffen waren.



Abbildung 2: Einragender Stutzen im Endschacht, außerhalb der relevanten Systemgrenze | Foto: ISAS GmbH

Wie im ersten Artikel dieser Serie im Dezember 2022 geschrieben, entsteht schnell der subjektive Eindruck, dass die optischen Verhältnisse einer Kanalbefahrung homogen und durch eine KI gut bewältigbar seien. Letztere beherrscht heutzutage schließlich vermeintlich weitaus komplexere Systeme wie den Straßenverkehr. Skepsis über diese These kam bei uns jedoch im Zuge der Recherchen für diese Artikelserie auf. Schon in den ersten willkürlich gewählten Befahrungsvideos begegneten uns unerwartet viele „irreguläre“ Situationen.

In diesem Artikel werden derartige Unregelmäßigkeiten zunächst systematisch und anschaulich dargelegt. Darauf basierend wird der Schwierigkeitsgrad der automatisierten optischen Mustererkennung in Kanalbefahrungen objektiv beurteilt.



Abbildung 3: Versatz an der Anbindung zwischen Schachtgerinne und Rohranfang des folgenden Kanals, außerhalb der Systemgrenze | Foto: ISAS GmbH

Betrachtungshorizont: Aufnahmen von Haltungen beliebiger Kamerasysteme

Die nachfolgenden Ausführungen behandeln einerseits die Variabilität in den Eigenschaften des inspizierten Objekts selbst und andererseits Unwägbarkeiten, die aus der eingesetzten Kameratechnik resultieren. Der Einfluss des Letzteren ist insbesondere für Inspektionen per Dreh-Schwenkkopf-Kamera relevant, da die verwendeten Systeme ein breites Spektrum an Kameraeigenschaften wie Bildauflösung, Seitenverhältnissen oder Brennweiten abdecken. Zudem lässt diese Kameratechnik u.a. hinsichtlich Kameramontage und Schwenkverhalten Spielraum für die naturgemäß unberechenbare menschliche Komponente. Die beiden genannten Unsicherheitsfaktoren der Dreh-Schwenkkopf-Kameratechnik sind für Vollkugelbildscanner technisch weitestgehend ausgeschlossen, werden aufgrund der hohen Marktrelevanz der Dreh-Schwenkkopf-Systeme im Folgenden dennoch ausführlicher behandelt.

Die Aussagen dieses Beitrags werden durch Bilder aus Haltungsbefahrungen veranschaulicht, da sich die Entwicklung von KI-Systemen bisher auf diese Komponente, und eher nachrangig auf die Kodierung von Leitungs- bzw. Schachtvideos, konzentriert hat.



Abbildung 4: Gesamtansicht des exemplarischen Risses mit komplexer Struktur | Foto: ISAS GmbH

Zustände innerhalb und außerhalb der Systemgrenzen

Der Betrachtungsraum einer TV-Untersuchung beginnt am Rohranfang des inspizierten Hauptrohrs (meist Haltung) und endet an dessen Rohrende bzw. an einem Bewegungshindernis bei Inspektionsabbruch. Weist das Hauptrohr Anschlüsse auf, wird der Einmündebereich zwischen der letzten Rohrverbindung im Anschluss und dem Hauptrohr selbst aus sanierungspraktischen Gründen dem Hauptrohr zugeordnet. Die oberhalb der letzten Anschlussverbindung gelegenen Rohrabschnitte gehören dagegen zur jeweiligen Anschlussleitung.

TV-Inspektionen enthalten häufig Videosequenzen von Auffälligkeiten bzw. Schäden, die außerhalb dieses Betrachtungshorizonts liegen, z.B.:

- Anschlüsse und/oder Schäden innerhalb der zulaufenden Leitung, sichtbar beim Hineinschwenken in den Anschluss vom Hauptrohr aus (Abbildung 1)
- Anschlüsse und/oder Schäden innerhalb des Start-/Endschachtes der Inspektion, abgebildet beim initialen/finalen Rundum-Schwenken im Schacht (Abbildung 2 und Abbildung 3)
- Alle Auffälligkeiten bzw. Schäden, die beim Zurückziehen der noch angeschalteten Kamera im Videobild sichtbar sind – Gefahr der doppelten Dokumentation!
- Bei einer ausschließlich Bild-basierten Kodierung durch das KI-System würden solche Fälle in der Zustandsdokumentation des inspizierten Rohrs fälschlicherweise erscheinen. Das KI-Tool muss – ebenso wie der menschliche Betrachter – in der Lage sein, derartige Feststellungen für die inspizierte Haltung als irrelevant zu verwerfen.



Abbildung 5: Momentaufnahme radiales Abschwenken des breiteren Teils des exemplarischen Risses | Foto: ISAS GmbH

Zusammenfassung von auf mehreren Einzelbildern sichtbaren Zuständen

Von Dreh-Schwenkkopf-Kameras erzeugte Videos bestehen aus sehr vielen, nacheinander aufgenommenen Einzelbildern, die das menschliche Auge aufgrund der Kanal-TV-typischen hohen Bildfrequenz von 25 Bildern pro Sekunde („frames per second“) als bewegte Szene wahrnimmt. Die gängigen KI-Analysen beziehen sich jedoch auf die im einzelnen Bild enthaltenen Informationen.

Dies kann insofern zu einer Herausforderung für ein KI-System werden, da derselbe Zustand auf mehreren Einzelbildern oftmals nur teilweise und mit veränderlicher Beschaffenheit sichtbar ist, wie bspw. der komplexe Riss in Abbildung 4, der auf Abbildung 5 und Abbildung 6 radial abgeschwenkt wird. Im Videoframe aus Abbildung 5 ist ein verhältnismäßig breiter Abschnitt des Risses abgebildet – im Einzelbild in Abbildung 6 ist dagegen nur sein „Haarriss-artiger Ausläufer“ vorhanden.



Abbildung 6: Momentaufnahme radiales Abschwenken des schmalen Teils des exemplarischen Risses | Foto: ISAS GmbH

Auch hier muss eine KI-Anwendung mehr beherrschen als nur die direkte Übersetzung von Bildinformation in Zustandsfeststellung. Dann würde die Dokumentation im Beispiel nämlich fälschlicherweise zwei verschiedene Codes – einen für den breiteren und einen für den schmälern Riss – enthalten. Stattdessen muss das KI-System die Zusammengehörigkeit mehrerer Teilabbildungen zu ein und derselben Feststellung erkennen und mit einem einzigen Code beschreiben.

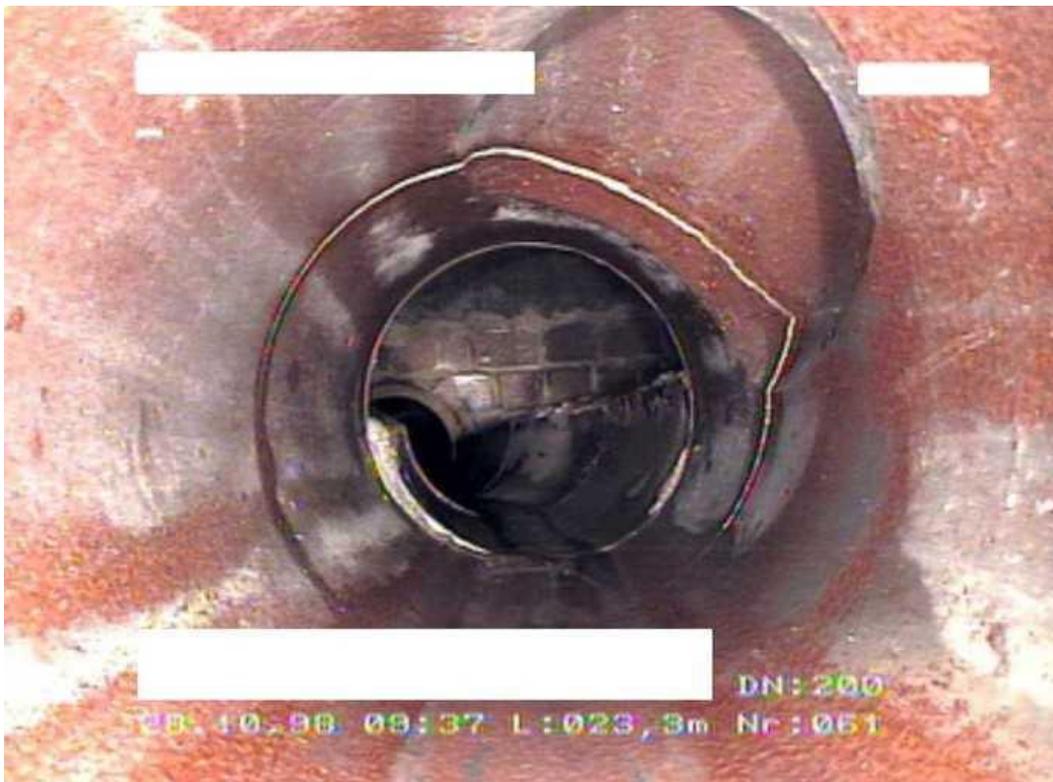


Abbildung 7: typischer Anschluss im Steinzeug-Rohr | Foto: ISAS GmbH

Optisch anspruchsvolle Situationen

Während der Recherchen für diesen Beitrag sind uns einige Fälle begegnet, für die selbst der menschliche Betrachter ein hohes Maß an Aufmerksamkeit und Konzentration für die richtige Zustandsbeschreibung benötigt. Die richtige Beschreibung derartiger Sonderfälle durch ein KI-System ist wesentlich schwieriger.

- Sehr große optische Ähnlichkeit zwischen einem Anschluss (Abbildung 7) und einer fehlenden Wandung (Abbildung 8) im Steinzeug-Rohr
- Identifizierung der weißen Linie am linken Kämpfer als Spinnwebe und nicht als Haarriss, indem bei laufendem Video sichtbar ist, wie das markierte Insekt beim Vorbeifliegen daran stößt und dadurch den weißen Faden zum Pendeln bringt (Abbildung 9)
- Optische Hindernisse auf der Linse, die einen Teil des Videobilds nur so weit stören, dass Objekte wie der Anschluss am linken Kämpfer durch das menschliche Auge weiterhin erfasst werden können (Abbildung 10). Derartige Anomalien können die Erkennungsgüte eines KI-Systems jedoch stark beeinträchtigen.



Abbildung 8: fehlende Wandung im Steinzeug-Rohr | Foto: ISAS GmbH

Individuelles Kamerahandling des Operators

Die Bediener von Kamerasystemen sind menschlich. Das bedeutet, dass die Entscheidungen, die ihnen im Rahmen der TV-Inspektion obliegen, nicht immer einheitlich

getroffen werden. Das gilt insbesondere für Befahrungen mit Dreh-Schwenkkopf-Kameras, die dem Operator naturgemäß mehr Freiheiten lassen, z.B. die unterschiedlich hohe Montage der Kamera auf dem Fahrwagen oder seine Schwenkfrequenz und -geschwindigkeit. Abbildung 11 und Abbildung 12 zeigen beispielhaft TV-Untersuchungen mit unterschiedlicher nicht-zentrischer Position des Kamerakopfes in der Rohrquerschnittsebene. Je nach Kamerahandling des

Bedieners kann ein KI-System einen bestimmten Zustand nicht identifizieren, weil er durch den Operator in einem anderen Winkel aufgenommen wurde als in den Referenzbildern, mit denen das KI-Tool zuvor trainiert worden war. Aus dieser individuellen Note resultieren also weitere Unsicherheiten für die KI-basierte Zustandsfeststellung.



Abbildung 9: Haarriss oder Spinnwebe am linken Kämpfer? Auf dem laufenden Video ist sichtbar, dass das mit dem roten Pfeil markierte Insekt die weiße Linie im Vorbeifliegen zum Schwingen bringt – es ist eine Spinnwebe! | Foto: ISAS GmbH

Bildauflösung und Brennweite

Für eine vollständige Zustandsbeschreibung ist das Ausmaß sämtlicher Feststellungen durch die KI-Anwendung zu beziffern. Im Jargon des Regelwerks sind dies die Quantifizierungen, also bspw. die Breite von Rissen oder der Versatz an Rohrverbindungen in Millimetern. Für die Messung solcher Größen aus den Bilddaten heraus sind Kameraeigenschaften wie die Bildauflösung in Pixel oder die Brennweite essenzielle Eingabeparameter. Diese Werte sind bei den allermeisten Dreh-Schwenkkopf-TV-Inspektionen jedoch sehr variabel und nicht bekannt, weswegen andere Wege zur Bestimmung der Zustandsquantifizierung durch ein KI-Tool gefunden werden müssen. Exemplarisch zeigen Abbildung 13 und Abbildung 14 beide jeweils einen Riss mit 1 Millimeter Breite – erstere mit höherer, zweitere mit geringerer Bildauflösung.

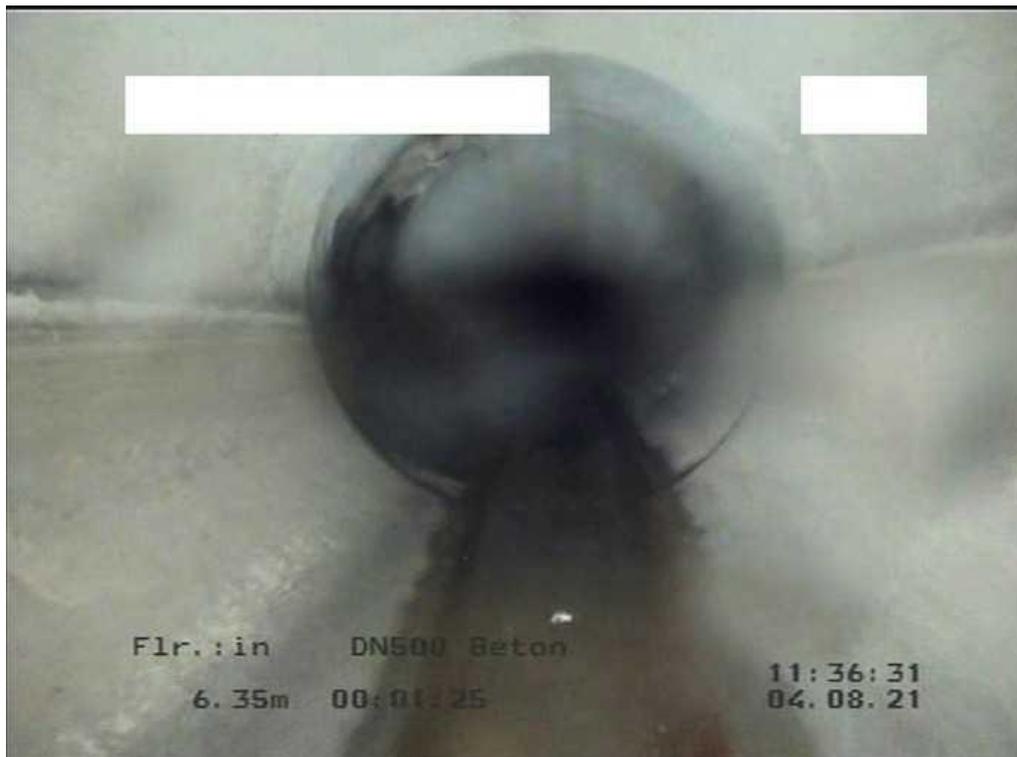


Abbildung 10: durch Wassertropfen auf der Linse beeinträchtigte Optik – das menschliche Auge kann Objekte wie bspw. den Anschluss am linken Kämpfer dennoch ausmachen | Foto: ISAS GmbH

Selbstversuch: Video ist nicht gleich Video

Abschließend möchten wir Sie als Leser dazu animieren, beliebig gewählte Videos aus Ihrem Kanalinspektionsfundus, sofern vorhanden, mit den Kernaussagen dieses Artikels im Hinterkopf zu betrachten. Möglicherweise entdecken Sie ähnliche optische Unregelmäßigkeiten, die in der Zustandsfeststellung eine Herausforderung für Mensch und/oder Maschine darstellen.



Abbildung 11: Kameraposition oberhalb des Querschnittmittelpunktes | Foto: ISAS GmbH

Fazit und Ausblick

Die erläuterten Inhomogenitäten in den Videodaten von Haltungsinspektionen belegen, dass die Übersetzung von Bildmaterial in eine vollumfängliche Zustandskodierung eine sehr anspruchsvolle Aufgabe darstellt. Insbesondere wird deutlich, dass zahlreiche fachtechnische Zusammenhänge über die reine Bildverarbeitung hinaus mit in den Dokumentationsprozess einbezogen werden müssen.



Abbildung 12: Kameraposition unterhalb des Querschnittsmittelpunktes | Foto: ISAS GmbH

Für KI-basierte Kodierung der Videodaten von Schächten und Anschlussleitungen ist aus unserer Sicht ein noch höheres Anforderungslevel zu erwarten, da die Objekte selbst variabelere, unberechenbarere Eigenschaften aufweisen. So sind Inspektionen insbesondere von Anschlussleitungen sehr heterogen, aufgrund ihrer dreidimensionalen Verläufe, kleinen Durchmesser, komplexen Schadensgeometrien und der daraus resultierenden, schwierigen Kameraführung.



Abbildung 13: Riss mit 1 mm Breite, höhere Bildauflösung | Foto: ISAS GmbH

Im Teil 3 der Fachartikelserie möchten wir beleuchten, inwieweit die Welt der KI-Entwickler eine Herausforderung in den geschilderten Aspekten sieht und wie damit umgegangen wird. In diesem Zuge werden wir diverse KI-Systeme detaillierter vorstellen.



Abbildung 14: Riss mit 1 mm Breite, geringere Bildauflösung | Foto: ISAS GmbH
Autor: Marco Deubler, Leiter KI-Implementierung, ISAS GmbH