「AIエッジ」を実現する画像センシング技術の活用事例

沖電気工業株式会社
イノベーション推進センター
センシング技術研究開発部
チームマネージャー

山本（やまもと）　一真（かずま）

沖電気工業株式会社
イノベーション推進センター
センシング技術研究開発部

ファン　チョンフィ

1. はじめに

近年、2030年までに達成する持続可能な開発目標SDGsが世の中に浸透し、17のゴール・169のターゲットの実現に向けて様々な取組みが行われている。OKIでは、2030年に「社会の大丈夫をつくっていく。」を実現することをキーメッセージとして掲げ、その実現に向けて、SDGsの17のゴールへ貢献する2つの重要課題（マテリアリティ）を定義している。1つ目は、社会価値をつくりだすためにSDGsの実現に向けて解決すべき7つの社会課題として、「老朽化問題」、「自然災害」、「交通問題」、「環境問題」、「労働力不足」、「労働生産性」、「感染症拡大」を定義している。2つ目は、サステナブルな企業活動の実現へ向けた取組みとして、ステークホルダーの期待に応える企業活動の実現とモノづくりを支える基盤の強化を挙げている。OKIの強みである顧客基盤・インストールベース・技術力を活かして、ユーザーに直接的または間接的にソリューションを提供していくことで、これらの重要課題へ対応している。

重要課題を解決するための強みとする注力技術を「AIエッジ」と定め、この「AIエッジ」を強化する先端的な技術領域として**図1**に示す5つの技術領域を定義している。その中の1つであるセンシング領域においては、現場を確実に見る技術で、製造や交通、建設などの間違うことや止まることが許されないクリティカルな現場の課題を解決することを目指し、光・画像・音等を対象としたOKI独自のセンシング技術の研究開発を進めている[1]。

■図1.　AIエッジを強化する注力技術領域

本稿では、OKIの画像センシング技術について、「AIエッジ」を実現するために必要な特長について述べ、その活用事例を紹介する。

2.「AIエッジ」を実現する画像センシング技術

クリティカルな現場の課題解決に向けた画像センシングの対象は、製造現場における人やモノ、道路における様々な交通事象など多種多様である。これらの現場における画像センシングに求められることとして、24時間365日リアルタイムに動作すること、細部に至るまで詳細に把握すること、屋外での雨天・吹雪といった天候変化や時間経過による照明環境変化がある状況でも確実に動作することなどがある。

また、現場に近いエッジでの処理に対するニーズも強い。近年のネットワーク技術の進展により、様々なシーンで大容量の映像伝送が可能となっている。一方で、既存のネットワーク環境を継続利用することを余儀なくされる現場も多く、細部に至る認識をするための高解像度の映像を伝送し、高性能なサーバーを使った映像認識の実施が難しいことがある。また、即応性を求めるような現場では、通信による遅延が許容されないこともある。多くの現場のニーズに応える画像認識には、高性能なサーバーを前提とせずエッジで処理可能なことが求められる。

OKIでは、これらの現場の要望に応えるために、以下のような特長を持つ画像センシング技術を開発している。まず、画像認識におけるAI処理を軽量化することにより、比較的省電力で処理性能の低いエッジ端末においてリアルタイムに動作することを実現している。また、現場の映像を使った開発に注力するとともに、現場では稀にしか起こらない事象に対してリアルな模擬映像を取得・生成することにより、現場が必要とする詳細な認識を実現している。さらに、天候変化や照明変化に頑健な認識を実現している。

3. 画像センシング技術の活用事例

現場の課題を解決する特長を持つ画像センシング技術の活用事例を紹介する。主に、製造現場での作業者行動認識について詳細を述べ、その他の現場での活用事例についても紹介する。

3.1 製造現場での作業者行動認識技術

近年、製造現場では、ロボットによる自動化が進んでいる。しかし、自動化が難しいため、人手作業が残っている工程は多い。製造現場では、熟練作業員の減少によるモノづくり品質の低下が懸念され、作業者のノウハウ・技術のデジタル化による継承が急務とされている。このような課題の解決に向けた、エッジ端末での作業者の詳細な行動認識の事例を紹介する。

製造現場での想定アプリケーション

製造ラインにおいて、各工程の作業者の行動をモニタリングすることや作業内容を数値化することは、品質確保や生産性向上のために非常に重要である[2]。このような適用先に向け、画像センシングによる詳細な行動認識技術を活用することで「現場を見える化」するアプリケーションの例を図2に示す。アプリケーション例①の「ポカヨケ」では、対象工程に必要な一連動作の順番やそれぞれの継続時間などを認識し、ミスがある場合に作業者に即時フィードバックすることにより、ヒューマンエラーを防止できる。ヒューマンエラーの例として、異なる部品を取り付けることや手順間違い、作業漏れなどがある。アプリケーション例②の「作業分析」では、詳細な行動認識技術を活用することで、製造ラインの改善に役立てることができる。改善の例として、蓄積された認識結果である個々の工程の動作に要した時間や熟練度を数値化し統計分析することで、作業のばらつきやトラブル発生時の要因分析が可能になる。また、このような詳細な行動認識技術を利用したアプリケーションを現場で運用することにより、「副産物」として膨大な作業映像を収集することができる。行動認識の結果と作業映像を利用して、未経験作業者の育成などにも活用することができる。

■図2. 画像センシングを用いた行動認識技術の概要と想定アプリケーションの例

詳細な行動認識技術

画像センシングを用いた人物の行動認識は「『粗い』行動認識」(Coarse-grained Action Recognition) と「『詳細な』行動認識」(Fine-grained Action Recognition) に分類できる。「『粗い』行動認識」では、人物の体全体が映像中に映り、走る・跳ぶ・座るといった一般的な行動を認識する。このような行動の識別は、人物や背景などの「見え」情報及び全体の「動き」情報を用いることで実現できる。一方、「『詳細な』行動認識」は、主に腕などの体の一部に注目し「モノを取る」や「モノを置く」などのある作業に関する一連の細かな動作を識別するという比較的難しいタスクである。そのため、識別対象の行動同士を見分ける特徴選択や時系列情報を参照できる識別モデルの利用が重要となる。

製造現場向けの作業者の詳細な行動認識を実現する要素技術を図3に示す。用途に応じて「骨格情報などの有効な特徴」を抽出し「時系列ディープラーニング技術」を用いて識別処理を行う。また、アプリケーションによって行動識別結果の時系列情報から、作業手順の正しさや作業の継続時間、同様の手順を行った回数などを用いて作業自体の良否を判定する。以降に詳細な行動認識を実現するための要素技術を紹介する。

■図3. 詳細な行動認識を支える要素技術

(1) 骨格情報などの有効な情報の抽出技術

製造ラインでの部品組立や加工など様々な工程で起こる動作（「パーツを取り付ける」や「ネジを締める」など）の認識を想定すると、背景・服装に依存しない作業者の姿勢や腕などの関節位置とその変化情報が重要である。したがって、行動識別に利用する情報として、主に映像の各フレームにおける作業者の骨格情報を用いる。骨格情報の抽出方法は、OpenPose[3]に代表されるディープラーニングを利用した手法が多数提案されており、これらの手法を利用することができる。

また、骨格情報を基本として、対象とする用途に合わせて他の情報も抽出する。例えば、手順間違い判定などのアプリケーションでは、同じような動作に対して向きを見分ける必要がある際に、骨格情報に加えてオプティカルフローなどの向きを示す情報を利用することができる。さらに、異なる部品取付け判定などの用途では、部品を検出した上でその位置・領域情報を利用することができる。

(2) 時系列ディープラーニング技術

作業者が行っている動作を識別するために、抽出した骨格情報などの特徴の時系列的な変化をモデリングする必要がある。一般に、時系列の骨格情報を用いた行動識別のディープラーニング技術はRNN（Recurrent Neural Network）、CNN（Convolutional Neural Network）、TCN（Temporal Convolutional Network）、GCN（Graph Convolutional Network）がある[4]。我々は、処理速度が比較的速く、骨格情報以外の他の特徴を容易に組合せることができるTCN[5]を用いる。TCNは、画像認識分野で広く用いられる一般的な畳み込み処理（Convolution）を時間軸方向に適用した時系列畳み込み処理（Temporal Convolution）を用いることで、ある時刻の前後の時系列情報を基に行動を認識する。時系列畳み込み処理を複数回行うことで、長時間の時系列情報を効率よく扱い、精度の高い行動認識が実現できるようになる。また、ディープラーニング技術の発展とともに画像認識分野の畳み込み処理に対して行われてきた改良を応用することができ、より複雑な行動を認識することも可能となる。このようにTCNは、高い汎用性を持ち、様々な改良を行うことができるため、様々な工程の現場に共通的に適用可能だと考える。

自社工場での活用事例

詳細な行動認識技術を適用して、OKI本庄工場のPHS組立工程で異物を除去するエアー吹き作業に対して、作業の良否を判定する行為判定システム[6]を検討している。エアー吹きでの作業漏れ・ミスがあった場合、後工程の検査で不良を検出することが困難であるため、作業自体の良否を判定することで、エアー吹き作業を確実に行ったことを確認することが求められている。

想定する行為判定システムを**図4**に示す。このシステムは、作業者が作業を行う場面をカメラで撮影し、エッジ端末などにより画像認識技術を用いて作業内容や作業手順が正しかったかどうかを判定する。その判定結果を作業者へリアルタイムに通知することで、作業者はその場で作業漏れや作業ミスに気付き、確実にエアー吹き作業を行うことができる。また、判定結果を管理サーバーに保存することで作業行為の証跡データとして活用できる。

現在、**図4**に示した行為判定システムを構築し、実際のPHS組立工程で動作させることで検証を開始している。今後は、作業漏れ・作業ミスの防止状況の分析や、作業者への効果的な通知方法の検証などを進める予定である。

■図4. 行為判定システム

3.2 道路インフラでの車両・人物計測技術

道路インフラへの応用として、映像中から車両や人物を検出し追跡する活用事例を紹介する。道路インフラの実態を把握するために、5年ごとに全国規模の道路交通量調査が実施されている。交通量調査は従来、主に人の目視による計測に頼ってきた。しかし、人手不足等に対応するため、監視カメラ映像を利用した画像センシングを用いた計測へと切り替わってきている。また、将来的には、カメラ映像を利用することで人による監視ではできない常時監視を実現し、交通状況をリアルタイムに把握することで、災害や渋滞に対する早期の対応をすること等が検討されている。

このようなニーズに応えるため、OKIは、特定の場所を通過する車両台数の計測や速度超過や逆走による事故防止を、エッジ端末によるリアルタイム映像解析システムとし

て実現するAISION車両センシングを提供している。また、自動二輪車の計測や立往生車両の検知、渋滞の検知といった機能追加によるバージョンアップも行うことで、多様化するニーズに対応している[7]。また、図5に示すような屋外での時間に応じた「見え」の変化に頑健な認識や、霧／雪／雨などを再現させた画像合成を行うことによる天候変化に頑健な認識を実現している。

■図5. 同一地点での時間ごとの映像（左：夕方／右：夜）

さらに、今後は対象を車両から人物に拡大していくことを検討している。車両に加えて人物も同時に検知することにより、歩行者の乱横断検知や混雑による車道へのはみだし検知等の実現が可能となり、より安全な道路インフラの実現に貢献することができると考える。

3.3 異常検知分野での画像生成技術

異常検知分野への応用として、判別対象の異常画像を正常画像から生成する画像生成の活用事例を紹介する。例として、画像認識技術を用いて、製造分野における不良品検知を実現することを考える。通常の製造時にできる製品・部品において、良品に比べて不良品は非常に少ない。そのため、認識処理開発時に参照できる不良品の画像も非常に少なく、検知できる不良品の網羅性が下がってしまう可能性がある。また、十分な数の自然な不良品画像を人手により作成することは困難である。

このような課題へ対応するため、pix2pix[8]に代表されるImage-to-Image Translationと呼ばれる画像変換技術を利用することにより、正常画像から様々な異常画像を生成することで、開発時に参照できる異常画像の網羅性を高める取組みを検討している。MVTec Anomaly Detection Dataset[9]のHazelnut画像の正常画像を異常画像に変換した例を図6に示す。この例から、入力した正常画像に対して、ヒビや穴のある画像を生成できていることが分かる。今後、

■図6. Hazelnutの異常画像生成の例

社内での認識処理開発へ適用し、有効性の検証を進めていく予定である。

4. おわりに

本稿では、「AIエッジ」を実現する画像センシング技術に必要となる特長をまとめ、製造現場・道路インフラ・異常検知へ向けた技術の活用事例を紹介した。今後も様々な現場課題を解決できる特長を持った画像センシング技術を開発していく。

参考文献

[1] 増田誠：「「AIエッジ」を強化するOKIのセンシング技術」、OKIテクニカルレビュー第239号、2022年5月.
[2] “Improving Manual Assembly Lines at Scale”, Drishti Technologies Inc, 2021.
[3] Cao et al.：“OpenPose：Realtime Multi-person 2D Pose Estimation using Part Affinity Fields”, TPAMI, 2019.
[4] Sun et al.：“Human Action Recognition from Various Data Modalities：A Review”, ArXiv, 2021.
[5] Lea et al.：“Temporal Convolutional Networks for Action Segmentation and Detection”, CVPR, 2017.
[6] 川面 怜哉、浅野 将仁、蘭 浩二、山本 一真、小林 司：「作業内容や作業手順の正しさを判定する行為判定システム」、OKIテクニカルレビュー第238号、2021年11月.
[7] 高橋 佑輔・渡辺 孝光：「AISION®車両センシング～安全・安心な車流管理の実現へ向けて～」、OKIテクニカルレビュー第238号、2021年11月.
[8] Isola et al.：“Image-to-image translation with conditional adversarial networks”, CVPR, 2017.
[9] Bergmann et al.：“The MVTec Anomaly Detection Dataset：A Comprehensive Real-World Dataset for Unsupervised Anomaly Detection”, IJCV, 2021.