

音声自動字幕システム ToScLive

講師の音声をリアルタイムに字幕化し、学生に配信



ToScLiveの概要

Outline of ToScLive automatic captioning system

新型コロナウイルスの影響により、大学などの講義のオンライン化が急速に進行している。それに伴い、通信障害による音声の聞き逃しや、聞き手の集中力欠如による理解度低下などへの対応が課題となっている。

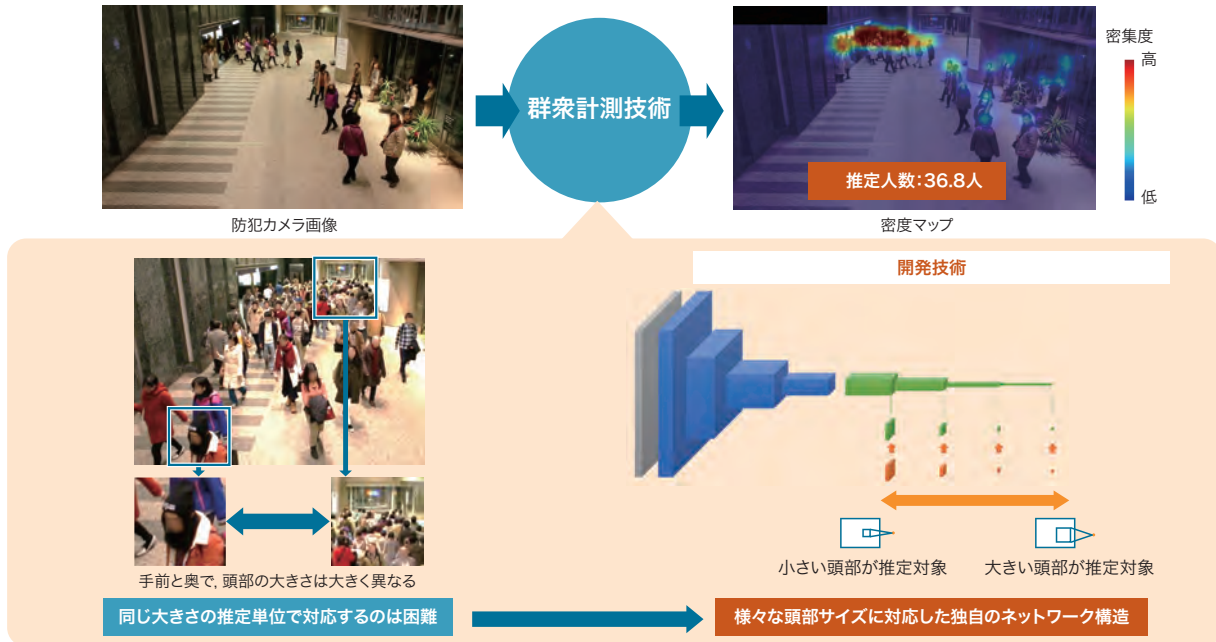
そこで、テキスト字幕による聞き逃し箇所の確認や講義の振り返りを実現するため、オンライン講義向けの音声自動字幕システムToScLiveを開発した。講師の声をリアルタイムに字幕化することで、オンライン講義の理解度向上に貢献できる。

ToScLiveの音声認識には、当社で開発してきた最新の深層学習手法を適用した。その結果、字幕を見たときに容易に内容を読み取ることができる字幕の生成が可能である。

講義特有の専門用語を含む発話を認識するには、専門用語辞書が必要であるが、これを人手で作成するのは、非常に手間が掛かるという問題があった。ToScLiveは、講義資料から専門用語を自動で抽出する機能と、複数のユーザー間で辞書を共同編集できる辞書プラットフォームを持っており、ユーザーが専門用語辞書を手軽に作成し、共有できるのが特長である。これにより、専門用語を含む高精度な音声字幕を実現した。

現在、複数の大学とともにToScLiveの実証実験を進めており、今後はToScLiveの活用による、コロナ禍に役立つ、新しい音声コミュニケーションの実現に貢献していく。

一般的なPCで高速・高精度に群衆の人数をカウントする 群衆計測技術



一般的なPCで高速処理が可能な高精度群衆計測技術の概要

Overview of highly accurate high-speed crowd counting technology utilizing general-purpose PC processors

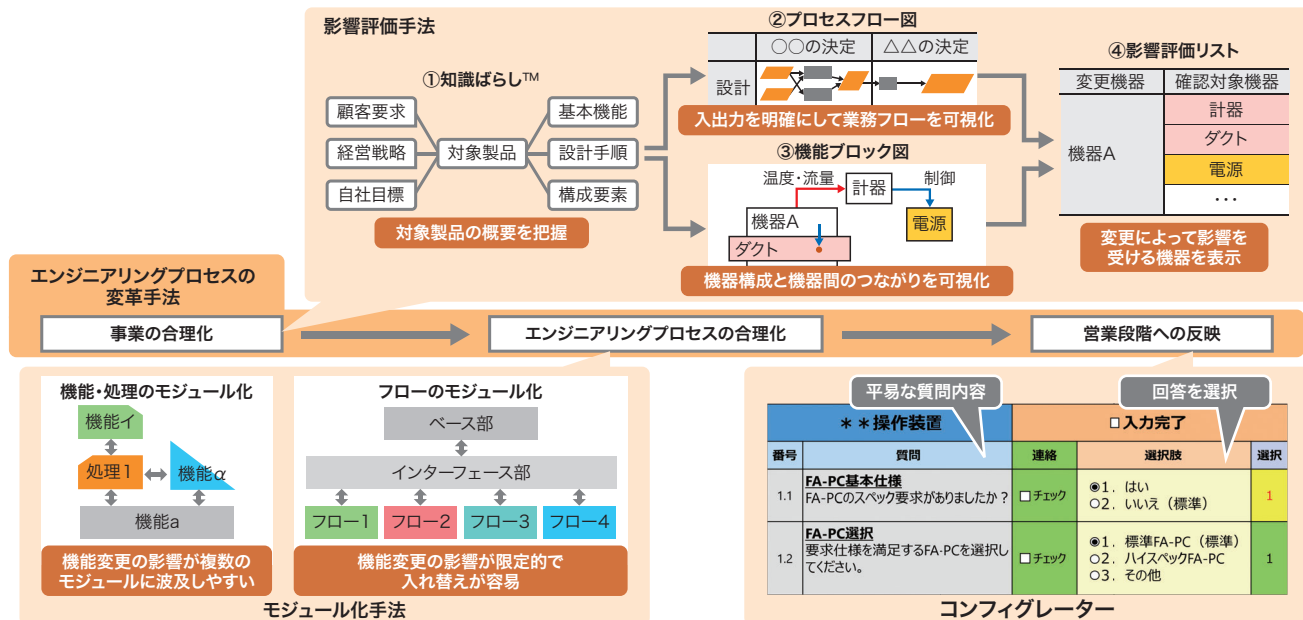
世界の防犯カメラの設置台数は、2022年までに10億台に達するといわれ、その後も加速度的に増加すると見られている。防犯カメラの映像を基に、駅や空港などの公共施設や商業施設で、トラブルの原因となりやすい人の密集状態を迅速に検出し、人数や混雑状況を高精度に計測するニーズが高まっている。また、昨今の新型コロナウイルスの感染拡大を受けて、密集状態を迅速かつ高精度に検知する技術が注目されている。

近年、深層学習技術を応用することで、密集度が極めて高い群衆の人数推定が可能になっている。しかし、一般に深層学習を用いた解析は、データの処理量が膨大でGPU (Graphical Processing Unit) などの高価な専用演算装置が必要となるため、コストの増加が避けられず、普及が進まなかった。また、カメラからの距離によって画像に写る人物の大きさが変わると人数を正確に推定できず、精度が低下していた。

そこで、一般的なPC (パソコン) で高速・高精度に人数をカウントできる技術を開発した。まず、低計算量動作に適した独自の深層学習手法により、GPUなどの専用の演算装置を用いず、PCに内蔵されているCPU上での高速処理を可能とした。更に、深層学習のネットワークを、画像に写る人物が大小どのような頭部サイズであっても解析できる構造にすることで、既存手法の課題を解決し、画像1枚当たりの人数の推定誤差を16.0%から14.7%に向上させた。

この技術により、低コストで多数のカメラの画像を処理できるため、特に多くの防犯カメラが設置されている施設において、効率的な混雑監視が可能となる。

基礎収益力を向上させるエンジニアリングプロセスの変革手法



FA-PC:産業用PC

エンジニアリングプロセス変革手法の確立

Establishment of method to reform engineering processes

従来、製造業の基礎収益力を向上させるために、工程や部品の標準化や共通化などのように、固定要素を増やしてエンジニアリングコストを下げる手法が使われてきた。今回、受注生産など多品種少量生産への対応を考慮し、顧客の様々な要求に応えながら変動要素に効率良く対応し、基礎収益力を向上させるエンジニアリングプロセスの変革手法を確立した。変革手法は“事業の合理化”、“エンジニアリングプロセスの合理化”、“営業段階への反映”の3段階で構成される。

事業の合理化では、製品戦略や、製品仕様、設計変更などによる事業への影響を、部品コストや設計工数などで定量的に把握する評価手法を開発した。知識ばらし™により体系化した事業情報を、プロセスフロー図と機能ブロック図に展開し、事業全体の整合性と合理性を評価する。この評価手法により、仕様変化がシステム全体に及ぼす影響を明確にでき、受注設計型製品での手戻り抑制や、研究開発要素の重要性評価に有効である。

エンジニアリングプロセスの合理化では、機能や処理をモジュール化して組み合わせる従来の手法に加えて、機能や処理のフローに着目したモジュール化手法を考案した。この手法は、ハードウェア・ソフトウェアに限らず適用でき、機能が多く組み合わせが複雑な製品においても、機能の追加や入れ替えによる影響範囲が小さく、カスタマイズが容易になった。この手法を適用したことで、組み込みソフトウェア製品の設計工数が80%低減した。

営業段階への反映では、概略仕様を入力しただけでも、禁則を考慮しながら未確定仕様を自動で設定する機能を開発した。また、見積もり手順をモジュールにすることで、平易で少ない質問に対する回答を基に、見積書を出力するスプレッドシート型コンフィグレーターを開発した。これにより、インフラシステム製品の見積もり工数を30%削減した。

■ 少量多品種の半導体製造に適用可能な不良解析技術



異なる半導体製品に共通して発生する不良の可視化システム
Failure visualization system to automatically classify failures common to different products in high-mix low-volume semiconductor manufacturing processes

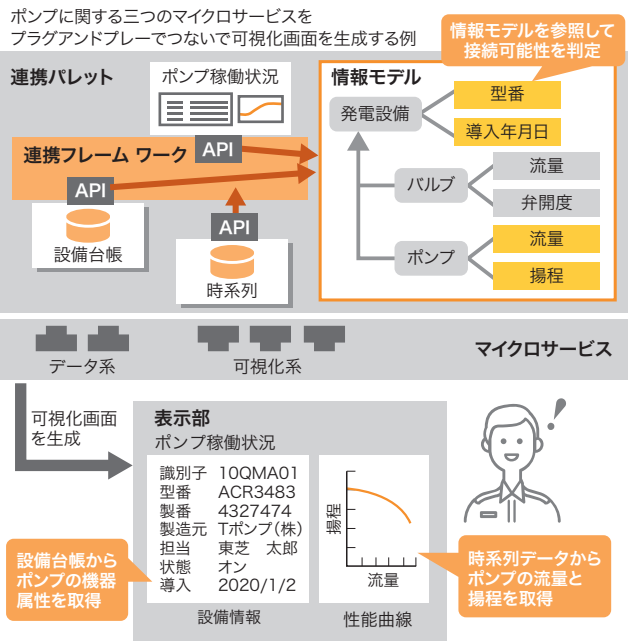
半導体工場では、製造過程で発生した不良に関する品質データをAIで自動解析し、作業時間の短縮に役立っている。顧客ニーズに合わせて必要な量だけ製造する少量多品種の工場では、製品によっては十分な品質データが得られず、解析精度低下の問題があった。また、多種類の製品の解析結果を何度も確認する必要があり、時間短縮が課題であった。

そこで、異種製品の品質データを一つに統合し、個々の製品の品質データが十分に得られない場合でも、高精度に不良解析できる技術を開発した。公開データWM-811K^(注)を用いた実験では、44製品の品質データの分類精度が75.3%から83.3%に向上した。この技術を応用し、異種製品に共通する不良を一目で確認できる製品またぎ不良可視化システムを開発し、東芝グループ向けに実用化した。1日につき1人当たり4.2時間掛かっていた不良解析業務を、約1/8の30分に短縮し、作業効率の向上に貢献している。

(注) データ分析コンペティションのプラットフォーム Kaggle が公開している半導体ウエハマップのデータセット。

研究開発センター

■ 情報モデルを活用したマイクロサービス連携技術



API: Application Programming Interface

情報モデルを活用したマイクロサービス連携技術の適用例
Example of microservice integration technology based on information model

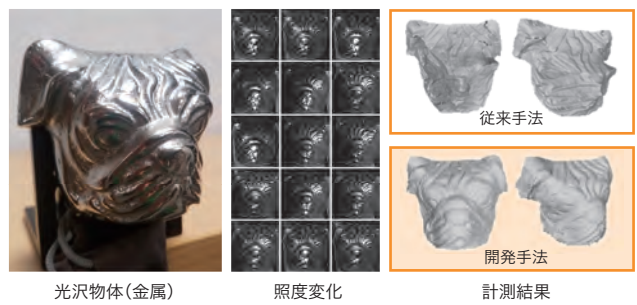
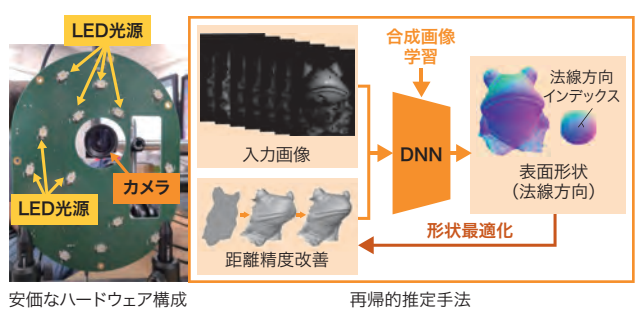
プラントでは、データを用いて設備稼働状態を把握し、運用保守に役立てる動きが進んでいる。設備の名称や、データ項目、単位などが設備ごとに異なると、それらに対する分析機能や可視化画面を個別に開発する必要が生じ、迅速なサービス開発の妨げとなる。そこで、情報モデルを活用して、データ取得などの共通機能を標準モジュール化し、これらを組み合わせて、迅速に運用保守向けサービスを構築できるマイクロサービス連携技術を開発した。

情報モデルは、設備の名称や、データ項目、単位などを扱うデータ辞書であり、当社が国際標準化を主導するIEC 61360 (国際電気標準会議規格 61360) 及びIEC 62656を基礎としている。情報モデルで名称などを共通化し、マイクロサービス間の接続可能性を自動判定できる。例えば、ポンプの型番や流量などのデータ項目を参照することで、設備台帳データ、時系列データ、可視化画面の三つのマイクロサービスをつないで、直ちにポンプの稼働状況を可視化できる。

今後は、既存システムのデータと接続する機能を追加し、運用保守向けサービス基盤に拡張していく。

研究開発センター

■ 光沢物の高精度な計測が可能な近接3D計測手法



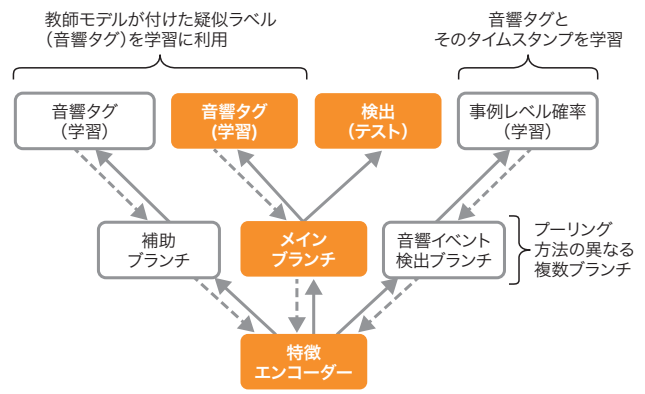
光沢物の近接3D計測手法と計測結果の例
Close-up three-dimensional (3D) reconstruction method and results of measurement of reflective metallic object

照度差ステレオ法は、照明変化による画像の変動情報に基づいて、被写体の3次元(3D)形状を計測する手法である。安価なLED(発光ダイオード)光源とカメラを用いて、数cmの近距離から1mm未満の解像度で3D形状を計測できるが、金属などの光沢物については、複雑な光学的物理現象を正確にモデル化することが困難で、従来の技術では計測できなかった。

今回、CG(コンピュータグラフィックス)によって光の反射をシミュレーションして作った大量の画像から、深層ニューラルネットワーク(DNN)を用いて主要な光学的物理現象をモデル化し、これを用いて近距離の光沢物の形と距離の推定を交互に繰り返し行う独自の3D形状推定手法を開発した。シミュレーション画像を用いた定量評価の結果、従来の手法に比べて、誤差を約46%削減できた。また、実画像を用いた計測でも破綻なく光沢物(金属など)を計測できることを確認した。

東芝欧州社

■ 音によるイベント検知技術



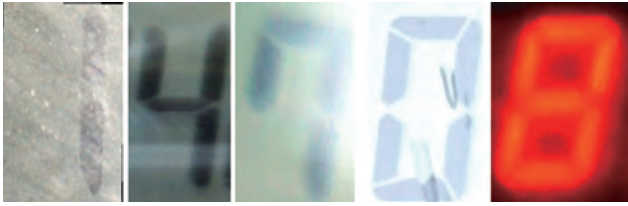
指導型学習とマルチブランチ学習による音響イベント検知技術の性能向上
Improvement in performance of sound event detection technology by means of teacher-student model-based guided learning and multi-branch learning

新型コロナウイルスの感染拡大で、機器やボタンに直接触れず、音声コマンドで操作できる機器の導入が増えている。これは、音声による機器操作という利便性の提供だけでなく、音で多様な生活行動のログを取得するデータサービス事業や、状況に応じたタイムリーなサービス提供にもつながる。

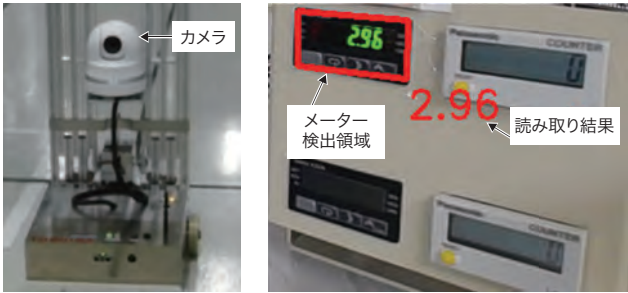
当社は、中国科学院と共同で、音でユーザーの行動や周辺状況を把握する音響イベント検知技術の研究開発を進めている。注意機構を用いた指導型学習(教師-学生モデル)とマルチブランチ学習を導入した独自手法を開発し、家庭内での多様なイベントの検知精度を向上させた。IEEE(Institute of Electrical and Electronics Engineers)のAASP(Audio and Acoustic Signal Processing Technical Committee)が公認する音響イベント検知の国際コンペティションDCASE2020チャレンジ(IEEE AASP Challenge on Detection and Classification of Acoustic Scenes and Events)に合同チームで参加し、音源分離・音響イベント検知部門でF値49.5%を達成し、1位を獲得した(2位は42.3%)。今後も、オープンイノベーションで更なる精度向上を図り、早期の実用性能達成を目指す。

東芝中国社

■ ディープラーニングによる高精度なメーター読み取り技術



読み取りにくい数字画像の例
Examples of difficult-to-read numerical images



メーター読み取り技術を搭載したAGV
Automatic guided vehicle (AGV) equipped with meter reading technology

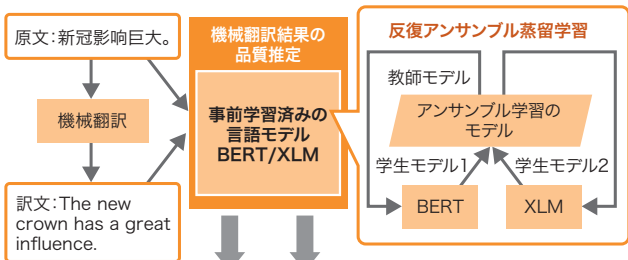
デジタルメーターの検出と数値の読み取り
Detection of digital meters and reading of numerical values

インフラ・エネルギー系の施設には、ネットワークに接続されていないメーターが多数存在し、作業者が巡回点検、帳票作成を行っている。省力化のために、カメラの画像からメーターを自動的に検出し、表示値を読み取って数値データ化するメーター読み取り技術が、求められている。従来は、数値の読み取りに2値化や細線化などの画像処理技術を用いていたため、低コントラストや、かすれ、環境光の反射などの影響を受けやすく、読み取り精度の向上が難しかった。

今回、数値読み取りアルゴリズムに6層ディープラーニング (CNN: Convolutional Neural Networks) を適用し、読み取り精度が従来の56.8%から99.7%へ向上した。また、画像取得、メーター位置探索、及び数値読み取りの三つの処理を並列化するソフトウェアの最適化を行い、更に、シングルボードコンピュータ上で推論アクセラレーターを用いることで、撮像から読み取り結果出力までの処理時間を130msから33msに短縮した。これにより、作業員や、AGV (Automatic Guided Vehicle)、ドローンなどの移動中に撮影した画像からの、高精度なメーター読み取りを実現した。

生産技術センター

■ 機械翻訳結果の品質推定技術



単語レベルの品質: The new (BAD) crown (BAD) has a great influence.
文レベルの品質: 翻訳エラー率 (数値)

事前学習済みモデルと反復アンサンブル蒸留学習による機械翻訳の品質推定技術

Automatic machine translation quality estimation technology using pre-trained models and iterative ensemble distillation

機械翻訳結果の品質を推定し、品質に応じて色を変えるなどの表示をすることは、ユーザーが翻訳結果を参照する際に有用であるだけでなく、翻訳サービスのドメイン適応において、翻訳品質の悪い結果を効率的に改善 (再学習) して翻訳精度を向上させる際にも、コスト削減の観点で非常に重要である。

当社は、北京交通大学と共同で機械翻訳品質推定の研究開発に取り組んでいる。最先端の言語モデルであるBERT (Bidirectional Encoder Representations from Transformers)、及びXLM (Cross-lingual Language Model)による事前学習済み言語モデルと反復アンサンブル蒸留学習の導入により、中英翻訳における単語レベル、文レベルの翻訳品質推定精度を、それぞれ49.76%、62.48% (2019年のThe 15th China Conference on Machine Translationでの最高値は47.39%、58.31%)と大きく改善した。大学との連携を継続して更なる精度改善を図るとともに、翻訳品質が悪い場合に自動で修正する機能の開発にも着手し、更なるコスト削減を目指す。

東芝中国社