

新ビジネスを創出するクラウドサービスとエッジデバイス向けミドルウェアのハイブリッド化

Creation of New Businesses by Hybridization of Cloud Services and Edge Devices

松本 信幸 MATSUMOTO Nobuyuki 柴田 智行 SHIBATA Tomoyuki 田中 孝 TANAKA Takashi

近年の、ICT（情報通信技術）分野でのクラウドサービスの普及と同時に、IoT（Internet of Things）におけるエッジコンピューティングも重要性を増している。

東芝デジタルソリューションズ(株)は、東芝コミュニケーションAI“RECAIUS”を、クラウドサービスとエッジデバイス向けミドルウェア製品の両面から提供しており、これらを更に連携させるハイブリッド化構想の取り組みを開始した。この構想に基づいて、IoTプラットフォームと連携した顔認証によるセキュリティーの二重化、及びエッジ・フォグ・クラウドコンピューティング層の構成による、複数カメラ間での人物対応付けという二つの事例に取り組み、相補的・相乗的な効果の創出に有効であることを確認した。

With the wide dissemination of cloud services in the field of information and communication technology (ICT) in recent years, edge computing has become essential for the Internet of Things (IoT).

"Toshiba Communication AI RECAIUS" provided by Toshiba Digital Solutions Corporation have already been supplying cloud services and middleware products for edge computing. We have now launched a hybridization project to make these services and products work in tandem. Based on this hybridization concept, we have worked on two programs: double security based on face authentication using an IoT platform, and person matching using images captured by multiple cameras through coordination of the edge, fog, and cloud computing layers. From the results obtained, we have confirmed the effectiveness of this approach for creating complementary and synergetic effects.

1. まえがき

ICTの進展とインターネットの浸透により、インターネットを基盤にした情報処理であるクラウドコンピューティングは、個人の生活、企業の活動、更には社会の制度などの大きな変化を支える存在になっている。一方で、センサーやデバイスがより小さく、安価に手に入る時代となったことを背景に、PC（パソコン）やスマートデバイスだけでなく、あらゆるものがインターネットにつながるIoTの社会が実現し始め、現場におけるエッジコンピューティングの重要性も増してきている。例えば、各社からスマートスピーカーがリリースされ、また、カーナビや家電などでの音声操作が普及し始めている。一般的に、これらを用いたサービスの処理は、家庭内や現場に設置されているエッジデバイスか、エッジデバイスがインターネットを介してつながっているクラウドシステムかの、いずれかで行われている。

東芝デジタルソリューションズ(株)は、長年取り組んできた音声認識や、音声合成、画像認識の機能を、東芝コミュニケーションAI“RECAIUS”として、クラウドサービス

(WebAPI : Web Application Programming Interface) とエッジデバイス向けミドルウェア製品の両面から提供している。当社は、このようなRECAIUSの特長を踏まえ、クラウドシステムとエッジデバイスを組み合わせて相補的・相乗的にこれらを活用するハイブリッド化により、それぞれを単独で提供するよりも、より良い価値を生み出し、新たなビジネスを創出することを目指した取り組みを進めている。

ここでは、まず、RECAIUSにおけるクラウドサービスとエッジデバイス向けミドルウェアのハイブリッド構想について説明し、続いて、この構想に基づいて実際に取り組んでいる、RECAIUSハイブリッドサービスの事例について述べる。

2. クラウドサービスとエッジデバイス向けミドルウェアのハイブリッド構想

2.1 クラウドサービス及びエッジデバイス向けミドルウェア製品のメリットとデメリット

RECAIUSは、クラウドサービス(表1)と、エッジデバイス向けミドルウェア製品(表2)の両面からの提供が可能である。

表1. RECAIUSクラウドサービス一覧

List of RECAIUS cloud services

分類	サービス名称	機能概要
音声認識	音声認識サービス	音声(日, 英, 中, 韓)を認識するPaaS
音声合成	音声合成サービス	多様な話者と感情表現でテキストを音声化するPaaS
音声翻訳	同時通訳サービス	音声(日, 英, 中, 韓)を同時通訳するPaaS
音声対話	音声対話サービス	曖昧な問い掛けにも応える音声対話を構築できるPaaS
知識処理	知識探索・活用サービス	テキスト資料を分析し, 知識として情報推薦・検索を可能とするPaaS
画像認識	人物ファインダー	人物の検出, カウント, 年齢や性別, 笑顔などの属性, 混雑度を判定するPaaS

日: 日本語 英: 英語 中: 中国語 韓: 韓国語
PaaS: Platform as a Service

*2018年7月時点

表2. RECAIUSミドルウェア製品一覧

List of RECAIUS middleware products

分類	サービス名称	機能概要
音声認識	音声認識ミドルウェアボイストリガー	音声でのトークスイッチ代替。特定のキーワードだけを高速で検出する組み込み音声認識エンジン
	音声認識ミドルウェアグラマー認識	あらかじめ認識する内容を定義しておき, 認識精度を向上させることができる組み込み音声認識エンジン
音声合成	音声合成ミドルウェアToSpeak	カーナビ, ゲーム, スマートフォン, 家電など身近な機器への組み込み音声合成ソフトウェア

*2018年7月時点

クラウドサービスは, 大量の辞書やデータを保持することができ, 連続的な発話の音声認識や, 感情表現などを含めた音声合成, 知識処理などが可能となるメリットがある。しかし, デメリットとして, インターネット接続が必須であることやレスポンス性能が問題となることがある。

ミドルウェア製品は, インターネット接続が不要でレスポンス性能が高いことがメリットである。一方, 大きな辞書を持つことができないため, 音声認識できる単語数が限定されたり, 音声合成の話者が固定されたりするデメリットがある。

2.2 ハイブリッド化による効果

2.1節で述べたような特徴をそれぞれ持つ, クラウドサービスとエッジデバイス向けミドルウェアを, 相補的・相乗的に活用することで, 以下のような効果が見込まれる(図1)。

- (1) 継続的な高性能・新機能の提供を実現 クラウドシステム側からエッジデバイス側へ, 最新の辞書データや, 最新のアルゴリズム, 新しい機能などを随時提供することで, エッジデバイス側での高い性能や新しい機能の提供を継続的に実現できる(効果①)。
- (2) より強力なAI機能を提供 クラウドサービスとエッジデバイス向けミドルウェアで, 処理量・処理内容に

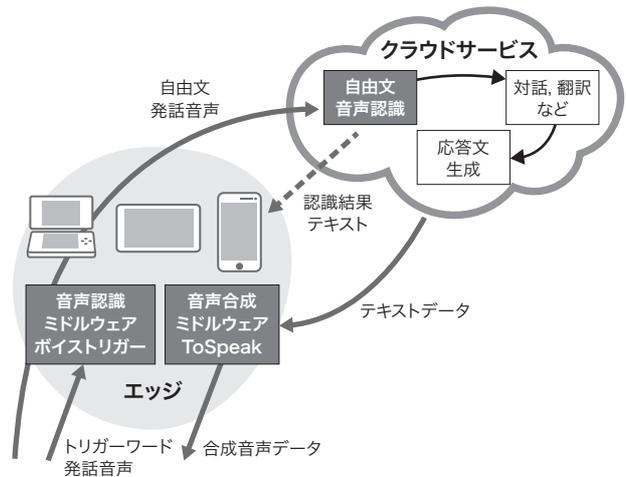


図1. クラウドサービスとエッジのハイブリッド化による効果

ハイブリッド化により, 継続的に, 最新で強力なAI機能を提供でき, きめ細かいサービスを実現できる。

Effects of hybridization

よってAI処理の役割を最適に分担することで, クラウド処理やエッジ処理単体では実現困難な強力なAI機能を実現できる(効果②)。

- (3) きめ細かい付加価値機能を状況に応じて提供 エッジデバイス側のユーザーの使用状態や発話情報などをクラウドシステム側に上げることで, エッジデバイス側の使用状態に合わせ, 付加価値の高い機能や情報をクラウドシステム側から供給でき, より賢く, かつきめ細かいサービスをユーザーに提供できる。これまで, エッジデバイス側での発話情報などは, リアルタイムで処理されるので, 事後活用をしていなかったが, クラウドサービスと組み合わせることで, 高付加価値化に向けた活用ができる(効果③)。

2.3 ハイブリッド化で創出される新たなユーザー価値

ハイブリッド化によって, ユーザーが新たに得られる価値の例を, 以下に述べる。

2.3.1 ユーザー環境音の収集による音声認識辞書の自動学習

エッジデバイス側で音声や環境音を自動収集してクラウドシステム側に集約する(効果③)。処理の重い音声認識辞書の作成処理をクラウドシステム側で実施して辞書を更新し(効果②), エッジデバイス側にアップデート版を提供する(効果①)。これまで, 案件ごとに環境音や言語モデルの録音・学習をオフラインで実施していたが, 音声認識辞書を自動学習できるようになり, 使っているうちに賢くなる

(認識精度が良くなる) AIサイクルが構築できる。

2.3.2 パーソナライズ支援

ユーザー指定の辞書をクラウド側からエッジデバイス側に自動配信する(効果①)ことで、ユーザーが所望のトリガーワードで機器を音声操作したり、音声合成によって、ユーザーが指定した声優や好きなタレントなどの音声でガイドしたりするパーソナライズが可能になる。

3. ハイブリッド構想に基づいた取り組みの事例

3.1 IoTプラットフォームと連携した顔認証とセキュリティの多重化

IoTデバイス製品を、米国や、中国、欧州、日本といった世界各国へ展開するためのIoTプラットフォームの一つに、Ayla Networks Inc.が提供するIoTプラットフォームがある⁽¹⁾。当社は、東芝コミュニケーションAI RECAIUSとAyla Networks Inc.のIoTプラットフォームを連携させた、新たな画像ソリューションの提供に向けた協業を進めている。新しいハイブリッドサービスの概念実証(PoC: Proof of Concept)を通して、IoTプラットフォームにRECAIUSを実装するときの課題を抽出している。

具体的には、IoT製品としてのスマートロックシステムと、画像認識エンジンを、Ayla IoTプラットフォームに組み合わせた、新たな個人認証の枠組みを実装する(図2)。スマートロックシステムとは、宿泊施設を予約したエンドユー

ザーのスマートフォンに、現地のドアを開けるためのロックID(識別番号)と呼ばれる暗証番号がクラウドシステムから事前にメール送付され、エンドユーザーが、宿泊時に施設の入り口ボタンでロックIDを入力することで解錠して入室できるシステムである。トラブルが発生しやすい、鍵の物理的な受け渡しが不要になり、民泊などでの活用が伸びている一方で、エンドユーザー間でのロックIDの流用リスクがあった。

この問題に対し、画像認識エンジンによる顔照合での個人認証をAyla IoTプラットフォームのスマートロックシステムと組み合わせることで、顔認証と暗証番号認証の二重認証システムが実現でき、ロックIDの流用を防ぐことが可能になる。しかし、実際には、エッジデバイス側に顔照合機能を実装するだけでは、エッジデバイス側への顔の事前登録が困難であるため、次々とユーザーが変わるような、民泊などでの入室セキュリティとしての運用は難しい。

このような事例では、クラウドシステムとのハイブリッド化が有効である。すなわち、スマートフォンを通じて、クラウドシステム側で顔を登録し、その情報をエッジ側に配信する構成にすることで、エッジデバイス単体では実現が困難な民泊などでの顔認証を可能とし、これを併用した二重認証システムの適用(効果②)が可能になる。

3.2 複数カメラ間での人物対応付け

安全・安心な社会づくりに向け、高度な防犯画像ソリューションサービスを実現するため、複数のカメラの視野範囲にまたがって移動する複数人物の移動経路を、少ない計算量で精度良く認識する技術を開発した⁽²⁾。この技術は、各カメラ近くのエッジコンピューティング層、複数カメラを束ねたフォグコンピューティング層(一つの施設に一つを想定)、複数の施設のデータを収集するクラウドコンピューティング層という三つの階層をハイブリッド化した構成により実現する。

以下では、ハイブリッド化での効果②を生むための、処理量・処理頻度に応じたエッジ・フォグ・クラウドコンピューティング層の最適な構成について説明する。

3.2.1 複数カメラ間での人物対応付けの全体構成

複数カメラ間の人物対応付け技術によって、施設内の各人物における移動軌跡及び属性情報を得ることで、例えば30代男性が立ち止まる傾向にある場所や、時間帯別の施設内滞留状況など、属性や、時間、場所などに基づく統計的な分析が可能になる。

複数台のカメラ映像から各人物の移動経路や属性情報を

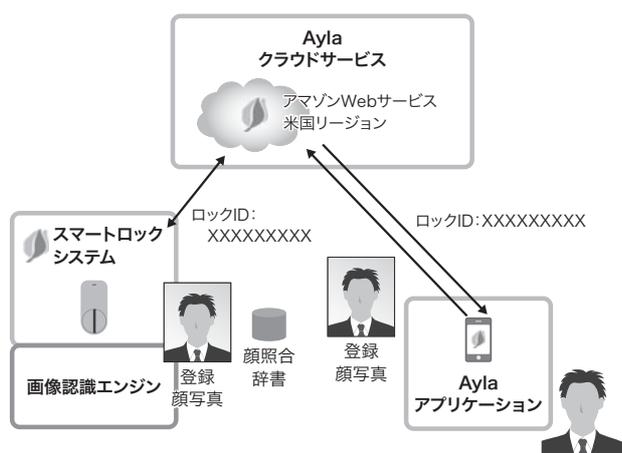


図2. IoTプラットフォームと画像認識エンジンを連携させたアーキテクチャーの例

IoTプラットフォームを活用することで、ローカルネットワーク外からの設定環境や課金サービスをそろえた高度なAIサービスを提供する。

Example of architecture linking IoT platform and RECAIUS image recognition engine

取得するには、①各カメラ映像から人物を検出・追跡し、②得られた追跡結果について施設内の複数台カメラ間で同一人物の追跡結果を統合し、③各施設の各人物の追跡結果から統計的な分析を行う、という処理が必要である。

これらの処理の実現形態として、インターネット上のクラウドシステム環境にカメラ映像をアップロードし、全ての処理をクラウドシステム上で行うことは、応答時間や通信コストの観点から現実的ではない。そこで、近年注目されているエッジコンピューティングやフォグコンピューティングの考え方を導入し、処理を分散することで、データ通信量と処理量を最適化する。

まず、各処理に必要な処理量の目安を概算するため、表3に、入力されるデータ量とその頻度、及び必要となる装置数を示す。

①の処理は、1台のカメラ映像中に写る人物を検出・追跡するため、カメラ映像のフレームレートよりも高速に処理し続ける必要がある、各カメラに付属したコンピューター若しくは物理的に近い場所にあるローカルネットワーク内のサーバーで処理することが望ましい（エッジコンピューティング層）。

②の処理は、①の処理により得られた追跡結果を、施設内の複数台のカメラから収集し、同一人物の追跡結果を統合するため、一定時間分の追跡結果をため込んで処理する。また、追跡結果のデータ量は、カメラ映像に比べて少なく、通信頻度も少ないため、ローカルネットワーク外のサーバーでも処理は可能である。更に、人物画像を必要とする特徴抽出を行うので、ローカルネットワーク上で処理することで、カメラに写る人物の個人情報保護やプライバシーの問題にも対応できる（フォグコンピューティング層）。

③の処理は、②で得られた各施設の結果をクラウドサーバーで分析してユーザーに提供するものであり、インターネット上のクラウドサーバーで処理すればよい（クラウドコンピューティング層）。

エッジ・フォグ・クラウドコンピューティング層でのハイブリッド化の具体的な処理フローを図3に示す。

表3. 各処理における入力データ量、入力頻度、及び必要装置数
Throughput required for each process

処 理	入力データ量	入力頻度	必要な装置数
① カメラ内人物検出・追跡	多い(映像)	5-30 fps	カメラ台数分
② 各カメラ間の対応付け	中程度	数分から数時間に1回	施設数分
③ 統計分析	少ない	半日から数日に1回	1台

fps : Frames per Second

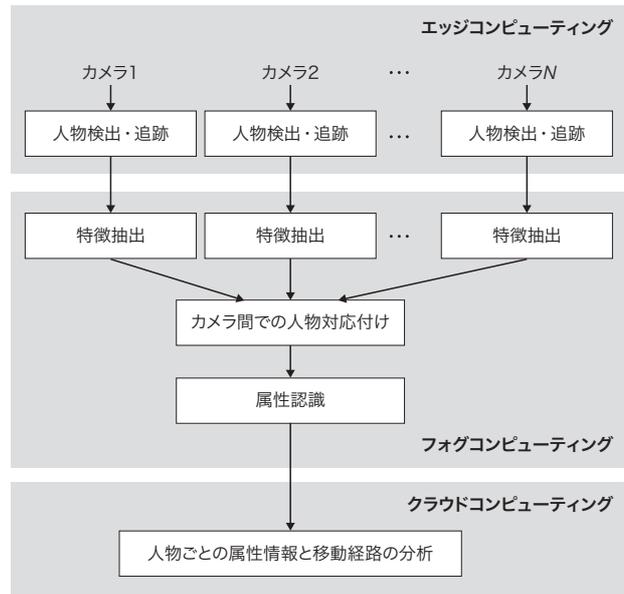


図3. 複数カメラ間での人物対応付けの処理フロー

人物検出・追跡はエッジコンピューティング層で、カメラ間の人物対応付けはフォグコンピューティング層で、追跡結果からの統計的分析はクラウドコンピューティング層でそれぞれ行う。

Flow of processing to match images of person from multiple cameras

3.2.2 複数カメラ間の人物対応付けの分散処理

フォグコンピューティング層で実装するカメラ間の人物対応付け処理は、対象とするカメラ台数の増加に伴い必要な計算量が増加する。そのため、遅延がないシステムを実現するために、対象とするカメラ台数に応じて処理を分散できる技術も必要となる。

各カメラを解析して得られた各人物の追跡結果から、同一人物を判定するための特徴量を抽出する。この特徴量は、同一人物同士は近く、他人とは遠くなる特徴空間でのベクトルである。対応付けする全ての特徴量について、特徴空間上で近接する特徴量を表す部分グラフだけと比較するグラフマッチングをベースとしたクラスタリング手法(k-NN (Nearest Neighbor) kernel shift法)を用いることで、対象のカメラ台数が増加しても、高速性を維持した処理が可能となる。また、対象のカメラを複数のグループに分割し、それぞれのグループを独立処理し、得られた結果を再度統合する分散処理が可能である。これにより、フォグコンピューティング層が対象とする処理は、導入するサーバー台数に応じて応答速度をある程度コントロールできる。

4. あとがき

当社は、人を想い(おもい)、人を支えるRECAIUSを、クラウドサービスとエッジデバイス向けミドルウェア製品の両面から提供している。

今回は、これらを連携させて相補的・相乗的な効果を生むハイブリッド構想を具体化するとともに、この構想に基づいて取り組んでいるIoTプラットフォームと連携した、顔認証によるセキュリティーの二重化と、エッジ・フォグ・クラウドコンピューティング層構成による複数カメラ間での人物対応付けという二つの事例について述べた。

今後も、顧客ヒアリングやPoCの成果を生かせるポテンシャルカスタマーの発掘を進めながら、ハイブリッド化により、価値あるサービスを提供し、新たなビジネス機会を創出していく。

文 献

- (1) Ayla Networks. "Connect. Virtualize. Transform.". Ayla Networks. <<https://www.aylanetworks.com/jp/>>, (accessed 2018-05-31).
- (2) 柴田智行. 複数のカメラで撮影された人物の対応付けと移動経路の推定技術. 東芝レビュー. 2018, **73**, 3, p.60-61. <http://www.toshiba.co.jp/tech/review/2018/03/73_03pdf/r01.pdf>, (参照2018-05-31).



松本 信幸 MATSUMOTO Nobuyuki
東芝デジタルソリューションズ(株)
RECAIUS 事業推進部 営業部
電子情報通信学会会員
Toshiba Digital Solutions Corp.



柴田 智行 SHIBATA Tomoyuki
研究開発本部 研究開発センター
メディア AI ラボラトリー
電子情報通信学会・情報処理学会会員
Media AI Lab.



田中 孝 TANAKA Takashi
東芝デジタルソリューションズ(株)
RECAIUS 事業推進部 営業部
Toshiba Digital Solutions Corp.