

セルビア郵政公社 郵便自動処理システム

Integrated Letter Sorting System for PE Post of Serbia

関根 弘義

入江 文平

金森 一也

■SEKINE Hiroyoshi

■IRIE Bunpei

■KANAMORI Kazuya

セルビア郵政公社では、東芝がターンキー契約で三つの主要な郵便区分局に納入した郵便自動処理システムが2015年12月から実運用に入った。

全体のシステムは、郵便区分機4台、小包・大型郵便区分機1台、OCR(光学式文字読取装置)及びBCR(Bar Code Reader)の認識処理とVCS(Video Coding System)を集中化した統合OCR(V)システム(Optical Character Recognition and Video Coding)システム、稼働データや運用計画を管理するIT(情報技術)システム、及び大口事業者から受け入れる大量の郵便物の数量を計測して料金の妥当性を検証する収益保護システムから構成される。今回のプロジェクトは、セルビア初の郵便自動処理システムの導入であったことから、システム運用開始後の運用方法や組織体制などの計画を早期の段階で作成し、セルビア郵政公社と連携して入念な準備作業を進めた。また、最新のディープラーニング(深層学習)手法を用いたOCR技術を投入することで目標の認識率である92%以上を達成した。更に、三つの郵便区分局に分散した区分機器からの情報は統合OCR(V)システムを設置した1局で集中管理することとし、各局をWAN(広域通信網)で接続する方式により実現させた。

Toshiba has developed a new integrated letter sorting system and delivered its constituent equipment to three mail sorting centers of PE Post of Serbia under a turnkey project. The system commenced operation in December 2015.

This system incorporates four letter sorting machines (LSMs), a mixed mail sorter (MMS), a unified optical character recognition and video coding (OCR(V)) system for integrated management of optical character reader (OCR) and barcode reader (BCR) recognition processing and video coding systems (VCSs), information technology (IT) systems for operation planning and management, and revenue protection systems for verification of the validity of charges based on the measured number of mail items. Since this is the first deployment of such a letter sorting system in Serbia, we carefully prepared the operational plan and organizational framework at the conceptual design stage in cooperation with PE Post of Serbia. We have confirmed that, through the application of the latest OCR technology using deep learning, the system achieves an address recognition rate of more than 92% as stipulated in the target specifications. Centralized management of the equipment installed at the three mail sorting centers has also been realized using the unified OCR(V) system installed at one mail sorting center connected to the other centers via a wide area network (WAN).

1 まえがき

郵便自動処理システムは、手紙や小包などの郵便物を宛先に従って自動で区分する区分機器を中核として、郵便物の宛先を読み取るOCR、BCR、及びVCSと、それらを集中管理する統合OCR(V)システムやITシステムなどから構成される総合ソリューションである。東芝は、郵便自動処理システムの開発を推進しており、同様なシステムの設計、調達から試運転までを一括して行うターンキー契約で進め、シンガポールポストへ納入した実績を持つ⁽¹⁾。

セルビア郵政公社が導入した郵便自動処理システム(図1)とこのプロジェクトの主な特徴は、次のとおりである。

- (1) セルビア初の郵便自動処理システムの導入
- (2) 最新OCR技術の採用
- (3) 複数局の区分機器と統合OCR(V)システム間をWANで接続し、情報を1局で集中管理



図1. セルビア郵政公社へ納入した郵便区分機(一部) — セルビア初の郵便自動処理システムであり、2015年12月から稼働している。

Part of letter sorting machines for PE Post of Serbia

ここでは、これらの特徴を中心に、プロジェクトの概要について述べる。

2 初となる郵便自動処理システムの導入に向けて

2.1 導入上の課題

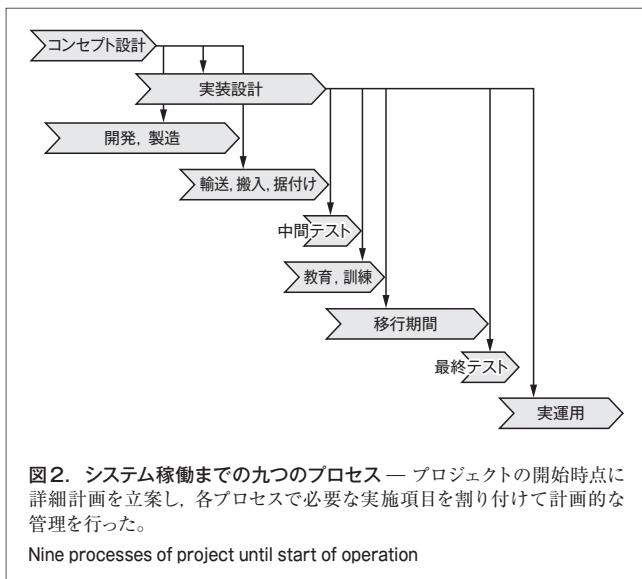
セルビア郵政公社では、これまで郵便物の区分を全て手作業で行っていたため、区分機器を使用する運用方法やノウハウがなかった。例えば、局舎内の区分機器のレイアウトや処理フロー、オペレーターや保守員の組織体制や教育訓練、停電などの障害発生で長時間にわたり区分機器が運用できない状況が発生した際の対応などである。

今回は、運用方法やノウハウの提供が契約に含まれており、区分機器や統合OCRシステムなどの機材納入に加えて、郵便自動処理システムの多岐にわたる運用方法やノウハウの提供と習得のための教育訓練が求められた。そこで当社は、コンセプト設計並びに実装設計を早い段階で行い、セルビア郵政公社と設計結果を共有しながら準備作業を共同で進めた。またその過程で、社外コンサルタントを活用することで種々の課題を解決していった。

2.2 コンセプト設計と実装設計

このプロジェクトの開始時点で、郵便自動処理システムが最終的な稼働に至るまでに必要な実施項目を洗い出し、九つのプロセスに割り付けた詳細計画を立案した(図2)。プロジェクトの実行においては、この計画に従って国際標準のPMBOK (Project Management Body of Knowledge) ガイドに準拠したプロジェクトマネジメント手法を用いた管理を行った。

初めての導入であることから、まずコンセプト設計で自動処理に向けた基本的なシステムを設計し、次の実装設計で具体的な運用レベルの計画も作成することにした。これら二つのプロセスでは、機器やシステムの仕様、輸送用機器の仕様、局舎内のレイアウト、区分機器の区分計画、処理フロー、局間・局内輸送、組織体制、教育訓練、リスク対応計画、中間・



最終テストの試験計画、実運用に向けた移行計画など、合計31の文書体系を中間テスト開始時まで作成した。

これらの文書をセルビア郵政公社と共有することで、郵便自動処理システムの運用開始に向けて早い段階から計画的な準備を進めていった。

2.3 社外コンサルタントの運用ノウハウ活用

プロジェクト開始時点で当社は、区分機器の運用ノウハウは持っていたが、郵便自動処理システム全体の運用に関してはまだノウハウ蓄積が十分ではなかった。そこで、郵便事業を運営する日本郵便(株)と、米国やカナダを中心に郵便自動処理システムの実運用に詳しくコンサルティング実績も豊富な Program management, LLCの2社とコンサルタント契約を結び、主にコンセプト設計と実装設計、並びに移行期間の各プロセスにおいて運用面でのノウハウの提供を受けた。

これにより、2014年12月に中間テストに合格した後、移行期間において徐々にシステムを稼働させ、2015年12月には最終テストに合格した。引き続き実運用に入り、その後順調に稼働している。

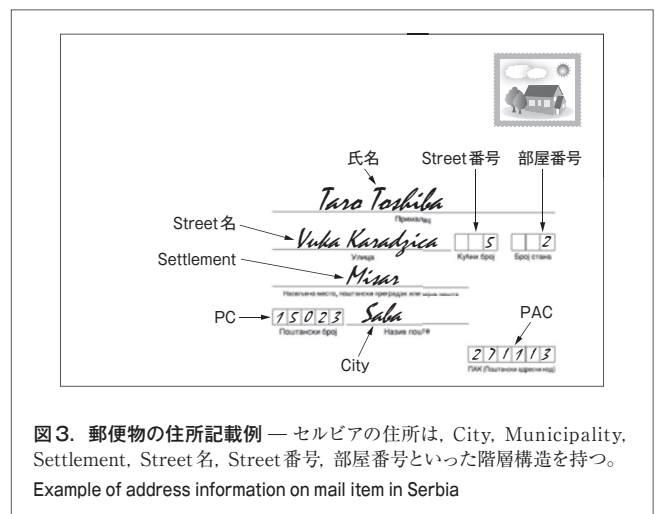
3 最新のOCR技術

住所認識上の問題点と、それらを解決するための技術について述べる。これらの新規開発・改良技術を組み合わせることで、仕様の住所認識率である92%以上を達成することができた。

3.1 住所認識上の問題

区分機器でのOCRの役割は、スキャナで取得された郵便物の画像データから宛先住所を認識することである。セルビアの住所情報は、City, Municipality, Settlement, Street名, Street番号, 部屋番号といった階層で表される。郵便物への住所記載例を図3に示す。

郵便物の区分が全て手作業で処理されていたため、郵便自



動処理システムの導入に際して次のような問題があった。

- (1) 5桁の郵便番号PC (Postal Code) のうち下3桁が0で埋められたり、誤記載されたりすることが多い。また、新制度の6桁の郵便番号PAC (Postal Address Code) は記載率が非常に低い。
- (2) Street名が個人名に由来するケースが多いため、しばしば変更されたり、異なった場所に同じStreet名が多くあったりする。
- (3) MunicipalityやSettlementなどの中間階層が省略された記載が多い。

(1)の結果としてPCは上2桁の情報しか利用できないことが多く、また(2)や(3)の結果として一つの住所に絞り込めないことが多い。これらの問題に対処するため、次に述べる各種技術の新規開発と改良を行った。

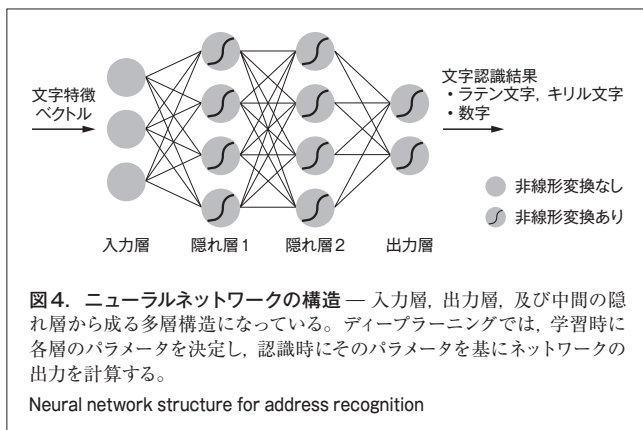
3.2 ディープラーニング手法の導入

3.1節で述べたような問題を解決する共通の基本技術は、一文字一文字を認識する文字認識エンジンである。そこで、近年パターン認識の分野で注目されているディープラーニング手法を利用したニューラルネットワークによる文字認識エンジンを新たに開発した(図4)。

ニューラルネットワークは生物の神経細胞をモデル化した素子を接続した神経回路網で、ディープラーニングは多層に接続されたニューラルネットワークを自動学習させるためのアルゴリズムである。今回、実際の郵便画像から切り出して正解を教示した文字の画像を様々に変形させた大量の学習データを用いて、短期間の学習で従来方法を上回る高い性能を実現した。

3.3 住所決定優先度ルールの実装

3.1節の(3)で述べたような中間階層が省略された住所表記では、宛先が一意に定まらない場合がある。この問題について手作業による区分の運用方法をヒアリングした結果、同じStreet名でも、都市部の場合は上位階層のSettlementやMunicipalityが省略され、地方の場合は省略されない、といったような暗黙の優先度ルールがあることがわかった。そ



こで、地名ごとに適用されていたルールをデータ化し、認識用住所データベース(DB)に取り込んで判定を行うロジックを開発して実装した。

3.4 複数のサブモジュールと最終判定方式の採用

3.1節の(1)で述べたような複数の郵便番号体系及び記載の不正確さに対処するため、住所認識サブモジュールや、PAC認識サブモジュール、PC認識サブモジュールなどの専用サブモジュールを用意し、各モジュールの結果を基に最終認識結果を判定する方式を採用した。

4 統合OCRシステムでの分散拠点情報の集中管理

4.1 分散拠点環境での課題

セルビアではベオグラード市、ノビサド市、及びニシュ市に郵便区分局を設けている。郵便自動処理システムの導入に伴い、これら3拠点に区分機器を導入することになった。統合OCRシステムでは複数の機器の情報を集中管理するため、このような分散拠点環境では拠点間を接続する必要がある。今回のケースでは、拠点間をWANで接続する方式で実現することとなった。

分散拠点環境に統合OCRシステムを導入するのは当社にとって初の試みであり、次のような課題があった。

- (1) 各拠点間で処理した全ての郵便物画像データをWAN経由でセンターサーバと連携して統合OCR処理を行う方式では、ネットワークトラフィックが大きくなり、WANで使用できる帯域幅の制限を超過して処理遅延が発生する。
- (2) 各拠点でそれぞれ処理された郵便物に対して1か所で集中エントリを実施するには、オペレーション管理を行うために各拠点の状況をリアルタイムで確認して作業調整を行う必要がある。

これらの課題に対処するため、システム構成を検討して各種カスタマイズを行うことで、分散拠点環境における統合OCRシステムの導入を実現した。各課題に対する解決策を次に述べる。

4.2 WANの帯域幅制限への対策

統合OCRシステムでは、各郵便区分機で取得した郵便物の画像データを使用してOCR及びVCSで処理を実施している。4.1節の(1)で述べたWANの帯域幅制限に対処するため、次の2点の対策を実施した。

- (1) OCRでは認識率を向上させるため高解像度かつ大容量の画像データを使用する。そこでOCRについては、OCR制御サーバとOCR処理部を区分機器と合わせて各拠点に設置し各拠点内で処理を済ませることにした。

ベオグラード区分局にセンターサーバ(WebアプリケーションサーバとDBサーバ)を設置し、WANによりノビサド区分局とニシュ区分局に設置したOCR制御サーバと

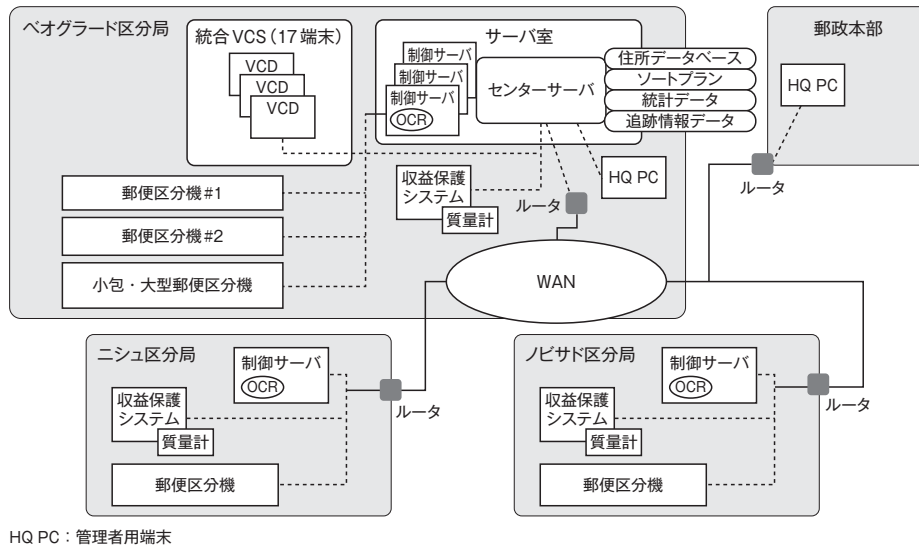


図5. 郵便自動処理システムの構成 — 3拠点をWANで接続し、ベオグラード区分局で各局の区分機器の情報を統合OCRV処理する方式とした。
Configuration of integrated letter sorting system

OCR処理部を接続する構成とした(図5)。これにより、OCR用の画像データをセンターサーバと連携して処理する必要がなくなり、ネットワークトラフィックを軽減させることができた。

(2) OCRで読み取れなかった画像データはVCSに送信される。VCSで使用する画像データはWebアプリケーションで処理されるため、WANを経由してセンターサーバに集約する必要がある。当初、VCSの処理対象となった画像データだけをOCRで使用した画像データから圧縮して作成し、センターサーバと連携させることでネットワークトラフィックを軽減させることにしていた。

しかし顧客要求で、全ての画像データを圧縮して上位のITシステムと連携させることに変更された。そこで、VCSで使用する画像データ作成時の圧縮率を、VCSエンタリ時の可読性を顧客と確認しながら調整し、その結果ネットワークトラフィックを軽減させることができた。

4.3 VCS処理の効率化

4.1節の(2)で述べた集中エンタリでは、VCS処理の効率化が求められる。VCSオペレーターはVCD (Video Coding Desk) 端末のモニタに表示された情報を確認して判定作業を行う。統合OCRVシステムでは、VCS状態監視機能により、各拠点における区分機器のVCS要処理データの蓄積状況及びVCSオペレーターの処理対象となる郵便区分機や、対応時間、処理件数などの情報を、一元的にリアルタイムで確認できるようにした。

これにより、オペレーション管理者は各拠点の要処理データの残量を確認することができ、17台のVCD端末に対して処理対象を最適に設定したり、必要に応じて変更したりすることが可能になった。

5 あとがき

郵便自動処理システムは2015年12月から実運用に入り、セルビア郵政公社での郵便物の処理業務に大きく貢献している。当社は、セルビア郵政公社の強い意向を踏まえ、国外の郵政事業者向けに同様の郵便自動化・情報システムの共同提案を適宜検討し実施していくこととし、2016年1月にMOU(了解覚書)を締結した。

今回のプロジェクトの経験を生かし、今後も海外向け郵便自動処理システムの受注活動を推進していく。

文献

- (1) 渡辺 徹 他. シンガポールポスト 郵便自動処理システム. 東芝レビュー, 70, 4, 2015, p.10-13.



関根 弘義 SEKINE Hiroyoshi

インフラシステムソリューション社 セキュリティ・自動化システム事業部 物流・郵便機器システム営業部。郵便自動化機器の商品企画及び開発に従事。
Security & Automation Systems Div.



入江 文平 IRIE Bunpei

インフラシステムソリューション社 小向事業所 SAソフトウェア設計部シニアエキスパート。郵便・物流自動化機器システムの開発・設計に従事。電子情報通信学会会員。
Komukai Complex



金森 一也 KANAMORI Kazuya

東芝ソリューション(株) ソリューションセンター 官公ソリューション第一部主任。郵便・物流関連のITシステムの開発に従事。
Toshiba Solutions Corp.