

— 見分ける・理解する —

情報社会を支える 画像認識・文書解析技術の最新動向

Latest Trends in Recognition Technology of Image and Document to Support Information-oriented Society

永崎 健 川股 幸博

Nagasaki Takeshi Kawamata Yukihiro

日立グループは、画像・文書の利活用に向けたさまざまな技術を開発し、情報インフラとして社会に提供している。昨今は、新しい撮像デバイスや学習技術の導入により、画像認識や文字認識の適用場面が増えてきた。画像認識に関しては、天候や昼夜による照明変動、路面標示のかすれといった悪環境においても誤動作なく認識が行える技術が開発され、交通事故防止を目的とした車載カメラへの適用が進んでいる。また、大量文字パターンの統計学習によって文字の誤読を低減するOCR技術や、海外とやり取りする仕様書から業務に必要な情報を簡便に抽出する技術を開発し、OCR製品や文書業務支援などで使われ始めている。

1. はじめに

日立グループが30年以上にわたって研究開発を続けてきた画像の自動認識技術は、さまざまな分野で実用化されている。例えば、工業分野・ビジネス分野においてはFA（Factory Automation）やOCR（Optical Character Recognition）などが典型的な事例である。

近年では、スマートフォンやドライブレコーダなど安価な撮像機器が普及してきたことに伴い、新しい市場が開拓されつつある。例えば、車載カメラを使った人物認識や障害回避機能は、交通社会を支える新しいインフラ技術として広く使われ始めている。また、文字認識においても、従来からある紙文書の読み取りだけでなく、文書作成ソフトなどで作られた電子文書をDB（Database）へ変換する文書解析技術や、カラー文書画像処理の進展により、その適用場面を広げつつある。

日立グループは、自動認識技術においてハードウェアからアプリケーションまで幅広い技術ラインアップを有している。

ここでは、画像認識技術と文書解析技術に関する最近の動向、およびハードウェア・ソフトウェアを含めた応用事例について述べる。



図1 | 日立グループの画像処理技術

30年以上にわたるFA、セキュリティ、自動車向けの画像処理のノウハウを日立画像処理ライブラリ、日立画像処理アクセラレータに蓄積し、日立グループ内の画像処理応用アプリケーションに横断的に活用している（ステレオカメラは富士重工株式会社と共同開発）。

2. 画像認識技術

2.1 画像認識技術の取り組み

日立グループはFA、セキュリティ、ITS (Intelligent Transport System) 分野で培った技術およびノウハウを相互に活用することで、高性能かつロバストな画像処理アプリケーションを展開している (図1参照)。

FA分野では、配線設計データと検査画像を照合することで微細加工に起因する配線不良を検出する半導体検査装置、セキュリティ分野では人間のさまざまな動きや多様な顔をロバストに抽出して侵入者などの異常を検知する監視システム、ITS分野ではさまざまな天候や、昼夜などの異なる屋外環境で、車両、歩行者、レーンを検知する車載カメラシステムを製品化している。

ここで培われた技術、ノウハウは日立画像処理ライブラリあるいは日立画像処理アクセラレータとして蓄積し、日立グループのさまざまなアプリケーションに横断的に活用している。

2.2 日立画像処理ライブラリ

日立グループは、交通事故防止を目的とした画像認識技術を応用した車載カメラの開発を推進している。天候、昼夜を問わず屋外を走行する車載向けのカメラであるため、照明変動、路面標示のかすれといった悪環境においても誤動作なく画像認識を行うことが必須である (図2参照)。

そこで国内の路面標示を認識する画像処理ライブラリを開発した^{1), 2)}。この画像処理ライブラリは、廉価なLSI (Large-scale Integration) でも動作可能にするため画像認識の処理時間や使用メモリを可能な限り節約した構造を持つことを特徴としている。

3. 文書解析技術

3.1 文書解析技術の取り組み

情報をやり取りする手段として、文書は数多く使われて



図2 | 車載カメラの認識対象

車載カメラで認識する日本国内の路面標示を示す。画像認識ライブラリは路面標示のかすれ、影、路面の照り返しなど屋外環境におけるノイズに対応している。

いる。しかし、その媒体は、従来からある紙文書に加え、PC (Personal Computer) で作成された電子文書や電子メールなど多様化している。この動向を踏まえて、日立のOCR技術もまた、紙文書や帳票類の読み取りから電子文書の構造化までをカバーする総合的な文書解析技術へと発展してきた。画像中の文字を読み取るだけでなく、文字列の表記やレイアウト配置を解析して文字列の意味を解釈する包括的な文書解析が可能である。この解析機能を支える要素技術として、文字認識/学習技術、カラー文書画像処理、意味解釈の三つがある。

3.2 文書解析技術の特長

(1) 文字認識/学習技術

金融向けデータエントリ業務では、さまざまな品質の文字・文書を安定して認識できることと、読み取った結果に誤読が少ないことの2点が要求される。従来は誤読を抑制するために、文字パターンの形状の可否を人間が定めたルールに従って判定する手法が使われてきた。しかし、ルールベースの手法はさまざまな変形パターンへの対応が難しいという課題があった。

そこで、文字の屈曲や積率などの局所的・大局的特徴を抽出する約50種のパターン特徴抽出器を、大量データを使った統計学習によって効果的に組み合わせることで誤読を低減する技術を開発した。大量サンプルで評価したところ、安定した認識率を維持しつつ従来よりも誤読を80%低減した。さらに、高次元空間上での認識のふるまいを解析し、曲率の高い判別面が精度に影響を与えるという現象に着目して判別面曲率軸を使った次元圧縮方式を開発し、これによる高速かつ安定した認識性能を実現した^{3), 4)}。

(2) カラー文書画像処理

手元の文書を手軽に撮像できる機器としてスタンド型スキャナがある。このスキャナは撮像部が外部に開放されているため、外光や照明変動の影響を受けやすい。こういった課題に対処するのがカラー文書画像処理である。その一機能である陰影補正技術は、色クラスタリングを用いて帳票の背景色を推定することで、キャリブレーション作業なしに、照明変動や物体の写し込みで生じた影を検知・除去できる。営業店端末システムなど、外光の影響を受けやすい窓口での文書受付業務を支援するシステムには欠かせない技術となっている (図3参照)。

(3) 意味解釈

OCRや文書解析システムの導入ネックの一つに、帳票定義作業の煩雑さがある。さまざまな会社が発行した帳票や仕様書から必要な情報を手軽に抽出するための技術が意味解釈である。その一例として金融機関向け定義レス帳票

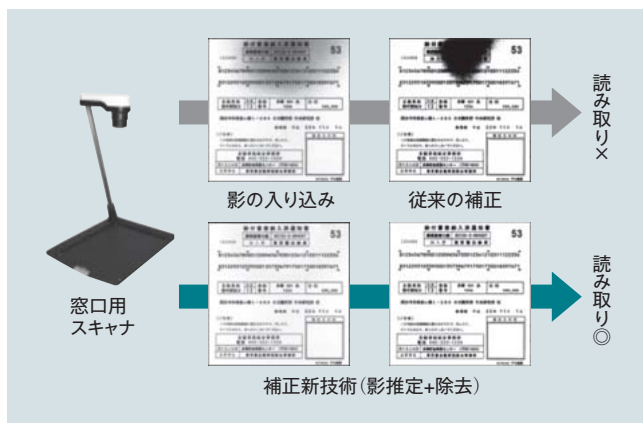


図3 カラー文書画像処理

外光の影響を受けるオープン型スキャナでも、キャリブレーションの手間なしに、陰影補正によって高品質な画像が得られる。

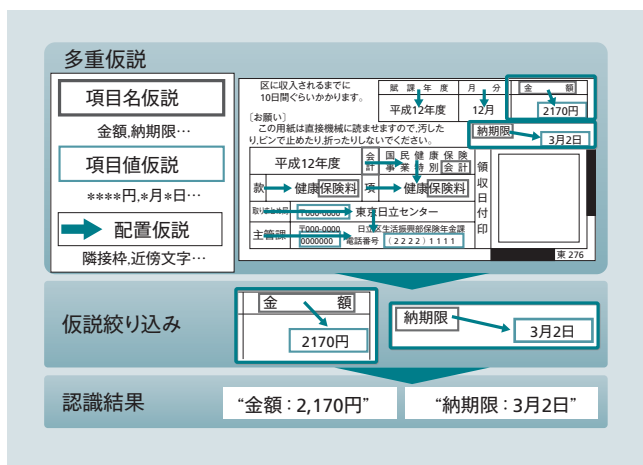


図4 意味解釈技術による定義レス帳票認識

給与報告書・納付通知書などさまざまな帳票（非整列枠形式）を、文書のレイアウト構造を事前に定義する必要もなく読み取ることができる。

認識がある。読み取りたいデータを、項目名（「金額」や「納期限」といったキャプション）と項目属性（数値列、年月日などデータ表記の形式）のペアで定義すれば、当該データの記述箇所を自動的に見つけて読み取る。読み取り項目の位置は定義する必要がないのでシステム導入時の手間を大きく軽減することが可能となる^{5), 6)}（図4参照）。

4. 適用事例

4.1 自動車事故防止のための画像処理応用

日立グループは、自動車向け画像認識として主に近傍センシング用モニタカメラとステレオカメラ向けの画像認識技術の開発を行っている。適用例について以下に述べる。

(1) リアカメラを用いた路面標示認識

駐車する際の後方のモニタ用として多くの車に搭載されているリアカメラの映像を流用して、前述した画像処理ライブラリを用いて路面標示の種別および車両との相対位置を算出し、地図DBに記録した路面標示の位置と比較することでカーナビゲーションシステムの自車両位置測定情報の高精度化を実現した（図5参照）。

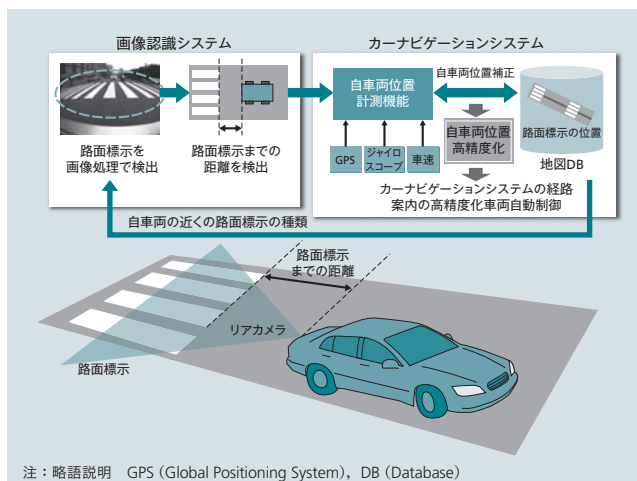


図5 リアカメラによる高精度位置計測

路面標示の位置を記録した地図DB、リアカメラ映像で認識した路面標示、車両の相対位置によって高精度に自車両位置を測定、カーナビゲーションシステム経路案内を高精度化した。

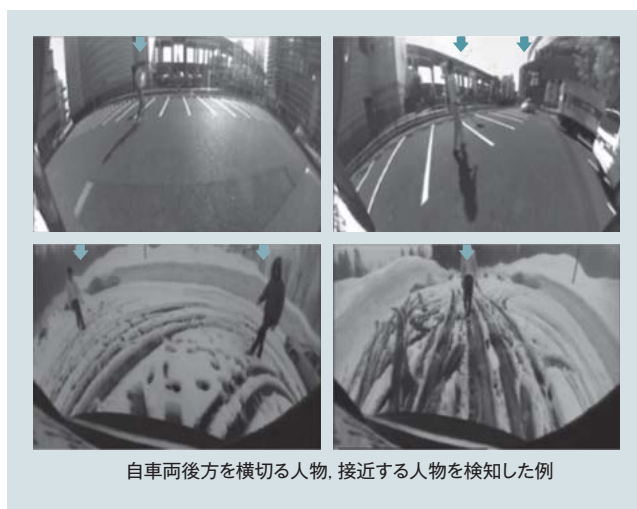


図6 周囲監視カメラによる移動体検知

自車両前方や後方を横切る人物、接近する人物・車両を検知し、ドライバーに警報を発する。

(2) 周囲監視カメラによる移動体検知

車両の移動状態を考慮することで周囲監視カメラの映像から自車両周囲の歩行者などの移動体を高精度に検出し、ドライバーに警報するシステムを実現した⁷⁾（図6参照）。

(3) ステレオカメラによる予防安全

日立グループは、富士重工業株式会社との共同開発によりステレオカメラの製品化を行うとともに、将来的なアプリケーションとして、カーブ前での減速制御を目的としたカーブ道路線形の推定、道路逸脱防止制御を目的とした道路端認識アルゴリズムの開発に取り組んでいる⁸⁾（図7参照）。

4.2 文書認識の応用

(1) 領収印の読み取り

カラー文書処理技術の応用例の一つが領収印読み取りである。従来技術では、あらかじめ押印や帳票の色を指定する必要があったのに対し、この技術では人間の知覚に近い

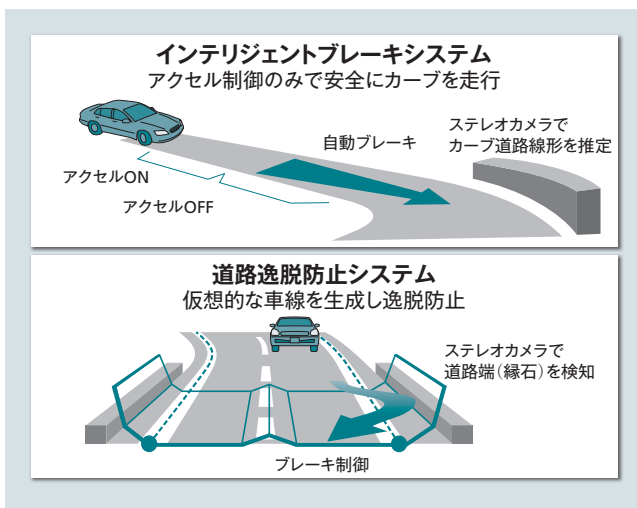


図7 | ステレオカメラによる将来の予防安全アプリケーション
カーブ検知によってカーブ前減速を行うインテリジェントブレーキシステム、道路端認識による道路逸脱防止システムの概要を示す。

カラー空間を用いた色彩の区分けによって押印色と帳票印刷色を自動的に分離して文字認識ができる。低品質な押印に対しても読み取れるため、税納付の期限チェックなどの用途で利用されている（図8参照）。

(2) 仕様データ点検支援ツール

ビジネスのグローバル化が進むにつれて、海外ベンダーや顧客が作成した英文の仕様書の処理負荷が増えている。このツールは、電子文書や紙の文書から仕様データを抽出し⁹⁾、自社製品の仕様項目と適合するかどうかを点検する作業を支援する。英文仕様書にはさまざまな様式が存在するが、文書の構成要素(章、節、文、キャプション、表)の解釈を多重化し、チェックしたい仕様項目を定めたオン・オフロジックを解釈する文書パーサにより、仕様データの抽出精度を高めている。予備実験ではデータ入力・点検にかかる作業時間を約半減した(図9参照)。

5. おわりに

ここでは、画像認識技術と文書解析技術に関する最近の

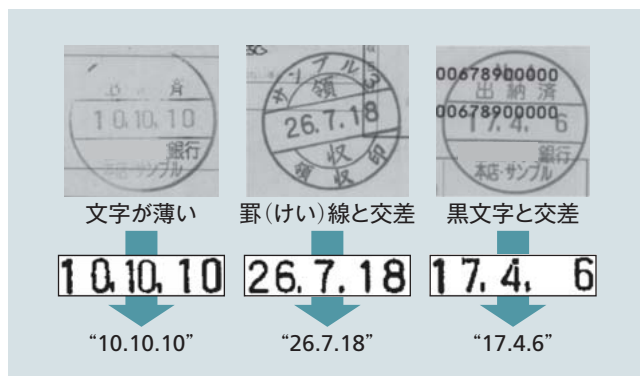


図8 | 領収印の自動検知と読み取り

振込伝票の上のさまざまな場所に押されたスタンプ（任意色、任意方向）を検知し、日付データを読み取る。

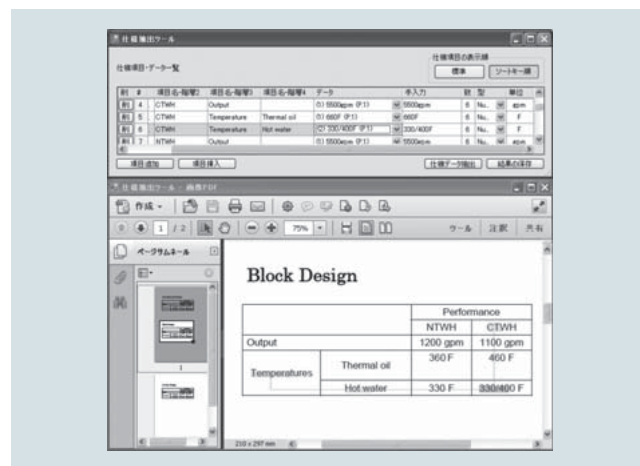


図9 | 仕様データ点検支援ツールの画面例

英文仕様書に書かれた仕様データを自動抽出し、自社製品のスペックに適合しているか否かを点検できる。

動向、およびハードウェア・ソフトウェアを含めた応用事例について述べた。

今日の情報化社会では日々膨大な画像・文書データが生成・蓄積されている。より高度な画像認識・文書解析技術の開発を通して、こうした大量にある画像・文書データの利活用に必要となる情報インフラ環境が提供できる。

参考文献

- 1) T. Shima, et al. : A Vision-based road marking detection system for car navigation using an on-board rear view camera, 15th World Congress on Intelligent Transport Systems (2008)
- 2) 樋口、外 : リアカメラの画像を用いた路面標示認識システムの開発, ViEW2008 ビジョン技術の実利用ワークショップ, 精密工学会 (2008)
- 3) T. Miyoshi, et al. : Simplified Polynomial Network Classifier for Handwritten Character Recognition, ICPR (2008)
- 4) T. Miyoshi, et al. : Character Normalization Methods using Moments of Gradient Features and Normalization Cooperated Feature Extraction, CCPR (2009)
- 5) M. Fujio, et al. : Logical Structure Understanding of PDF Documents for Generating Business Document Templates, DAS (2008)
- 6) 平山、外 : 仮説検証型アプローチを用いた定義less帳票認識技術, FIT (2010)
- 7) 清原、外 : 車周辺監視のための移動体検出技術の開発, ViEW2011 ビジョン技術の実利用ワークショップ, 精密工学会 (2011)
- 8) 樋口、外 : ステレオカメラによる立体物情報をを用いた道路形状推定, ViEW2009 ビジョン技術の実利用ワークショップ, 精密工学会 (2009)
- 9) M. Seki, et al. : Information Management System Using Structure Analysis of Paper/Electronic Documents and Its Applications, ICDAR (2007)