

ものづくりの現場で文字情報の活用を支援する 情景文字認識技術

Scene Text Recognition Technology to Obtain Information in Manufacturing Processes

古畑 彰夫 小野 聡一郎 馮 思萌

■ FURUHATA Akio

■ ONO Soichiro

■ FUU Shimou

ものづくりの現場では、データ分析での活用や品質基準への対応などの目的から、材料の使用履歴を蓄積管理していくことが求められている。このような管理すべき情報はバーコードなどを用いて入力されることが多いが、ロット番号のように文字でしか記載がない情報も中には存在し、これらの取得が課題になっている。

東芝デジタルソリューションズ(株)は、東芝グループが長年培ってきた文字認識技術を応用した情景文字認識技術により、このような情報取得の支援を行っている。製造現場の状況は多様であるため、ユースケースに応じてメリット・デメリットを考慮して、既存のモジュールの利用やパラメーターの適正化によりカスタマイズを行うことで課題を解決し、実用水準を達成した。

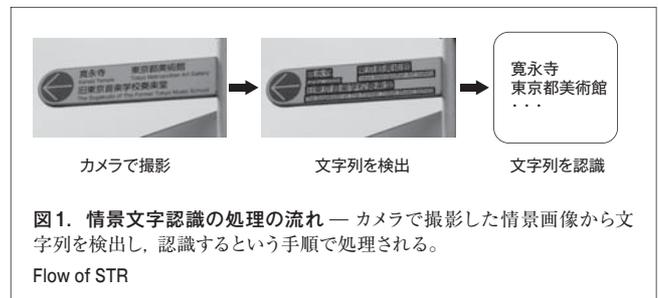
With the accelerating movement toward the certification of traceability and analysis of big data from processing steps in recent years, the acquisition of information on materials directly and indirectly used in manufacturing processes in factories and other production facilities has become increasingly important. Such data are often acquired from barcodes, etc. However, some information, such as lot numbers and expiry dates, is shown only in the form of characters. The acquisition of such text-based data is one of the main problems in this field.

Toshiba Digital Solutions Corporation has been developing a scene text recognition (STR) technology to support the acquisition of information from factory scenes applying optical character recognition (OCR) technology, a field in which the Toshiba Group has accumulated long experience. Since scenes at manufacturing facilities vary greatly according to the situation, we have customized our STR technology to adjust to the environment by selecting and assembling known modules and optimizing their parameters in accordance with the restrictions specific to the site. We have confirmed that our new STR technology has now reached the practical level.

1 まえがき

近年、ものづくりの現場では、ビッグデータを活用した業務改善や品質基準への対応のためのトレーサビリティ確保などが重要視されており、製造過程において、製品がいつ、どの機械を通り、そこで使用された材料は何だったか、という詳細な履歴情報を、全工程にわたって蓄積し、管理していくことが求められるようになってきている。管理すべきこのような情報は、多くの場合、バーコードやビーコンなどを用いて入力されるが、製品にこれらの付与が難しい場合や、ロット番号のように、製品個体ごとに情報が異なる場合などでは、文字でしか記載がない重要な情報もある。現状、これらの情報は自動取得が困難であり、多くの場合、目視確認や手書きでの記録にとどまったり、手入力コストが多く掛かったりして、十分な活用ができていないことが課題となっている。

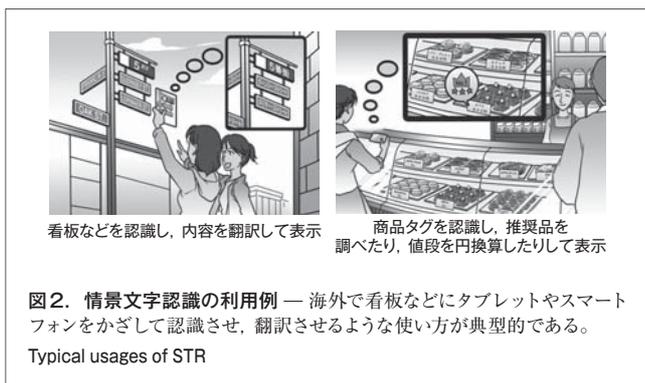
情景文字認識は、このような課題の解決に有効な技術である。例えば、タブレットや、スマートフォン、眼鏡型ウェアラブル端末などに搭載されたデジタルカメラで撮影した風景画像の中から、文字の書かれた部分を検出し、認識できる。情景文字認識では、図1の手順に従って画像中の文章をテキスト化する。まず、カメラで撮影した情景画像の中から文字らしさなどの情報を用いて文字列を抽出する¹⁾。次に、抽出された各文字列



を認識し、結果のテキストを得る。情景文字認識は、認識して得たテキストを利用するサービスの入力機能として用いられることが多い(図2)。最近ではカメラを文字にかざし、翻訳して表示するスマートフォン用のアプリケーションが有名である。

しかし、ものづくりの現場での文字認識は、多くの場合、撮影環境や文字自体の品質などの影響で、通常の文字認識に比べて低精度であり、高精度な認識を実現するには、ユースケースに応じたカスタマイズが必要である。

東芝デジタルソリューションズ(株)は、東芝グループが長年培ってきた文字認識技術を応用し、文字でしか存在しない重要な情報でも簡単な操作で取得可能とする情景文字認識技術を開発し、カスタマイズすることにより実際の案件に適用した。ここでは、情景文字認識技術の概要と適用事例、今後の展望



について述べる。

2 情景文字認識の主な課題と解決策

2.1 課題

ものづくりの現場における、情景文字認識技術を活用する上での主な課題には、以下の二つが挙げられる。

- (1) 認識対象の書式 多くのコンシューマー向けアプリケーションは、通常の文章を対象とし、言語知識を活用して認識する。しかし、ものづくりの現場で認識対象とする文字列は型番などであり、通常の文章ではない。文字列は、多くの場合定まった書式に従って書かれ、通常の文章とは言語的な性質が大きく異なる。このため、コンシューマー向けアプリケーションをそのまま用いても、高い認識精度は期待できない。
- (2) 画像の品質 現場の構造や予算などの理由から、カメラの性能や設置位置に制約があり、綺麗な画像を撮影することが難しく、画像の品質が低い場合が多い。典型的な低品質画像の例を図3に示す。撮影系の問題でピントずれした画像や、光が反射した画像、円筒形の瓶に貼



られたラベルのように非平面状の対象物の画像、撮影角度のばらつきなどによって文字が変形した画像などが挙げられる。

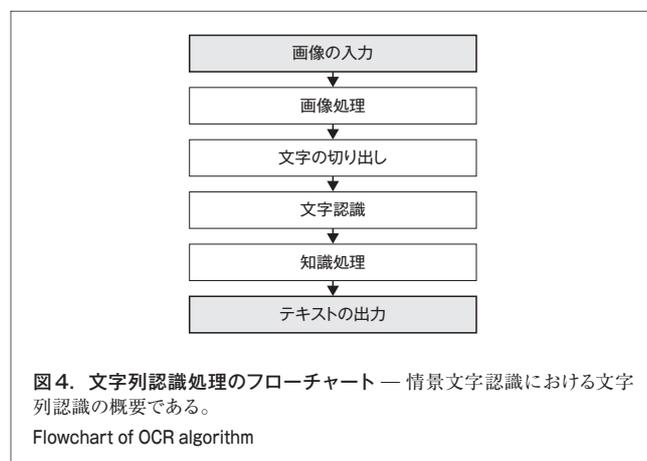
書式や撮影画像の品質はユースケースごとに大きく異なり、同じアルゴリズムやパラメーターセットで全てのケースに対応するのは現実的に困難である。このため、情景文字認識を用いる際には、ユースケースごとに生じ得る課題とその範囲を特定し、絞り込んで解決するといったような適切なカスタマイズが必要になる。

2.2 解決策

以下では、情景文字認識技術のうち、図1に示したような文字列の認識処理（以下、文字列認識と略記）におけるカスタマイズについて述べる。文字列認識のフローチャートは、概略、図4に示すとおりである。文字の切り出し及び知識処理においては主に書式情報の利用、画像処理及び文字認識においては主に撮影画質への対応で、それぞれカスタマイズを行う。

2.2.1 書式情報の利用 ここで言う書式とは、“数字2桁、区切り記号、英字2桁、数字3桁、区切り記号、数字2桁”というような、対象文字列内での文字の出現規則である。文字の切り出しには、比較的正確に検出しやすい区切り記号を最初に検出し、その位置を基準として残りの文字の切り出しを行う手法や、文字が黒で製品地が白であることを前提として濃淡情報を用いる手法などを用いる。これらの手法により、文字同士の接触が多い場合でも比較的正確に文字の切り出しを行うことができる。知識処理では、与えられた書式情報に基づいて認識結果を補正する。例えば、縦長楕円（だえん）形の文字に対する認識結果を、その文字の字種が数字なら“0（ゼロ）”、英字なら“O（オー）”とする。書式が複雑な場合、決定性有限オートマトンを用いた記法を用いることもできる⁽²⁾。

2.2.2 撮影画質に対応するためのカスタマイズ 画像処理では、鮮鋭化フィルターなどを用いて組み合わせたアルゴリズムとパラメーターの最適化や、斜め撮影画像の正面撮影への補正などを行う。文字認識では、主に撮影条件に起因す



る画像劣化をモデル化し、この劣化モデルに基づいて文字認識の学習を行う。ただし、劣化の程度が大きい場合には、実環境から画像データを収集して学習に用いる。

3 情景文字認識の適用事例

3.1 製品及び材料の個体管理

ものづくりの現場では、製品や材料の個体管理が重要な課題である。多くの場合、これにはバーコードやRFID（無線ICタグ）などが用いられるが、対象物の性質上、それらの添付が困難な場合、あるいは作業員の目視確認が重要な場合には、管理番号を印刷して管理することがある。そこで、情景文字認識技術を適用して、工程の各段階で製品に印刷された管理番号を認識し、正しい製品が来ているかを確認した。

大型の製品を製造する工場の場合には、近距離からの撮影は難しいため、カメラを遠距離に設置する必要があった。また、正確に同じタイミングで撮影することはできず、撮影対象との距離や角度に撮影ごとのずれが生じた。このためピントがずれ、ややぼけた画像が認識対象となった。また、対象物が大型で印刷が難しいため、多くの場合で字形の崩れやドットの欠けが発生した（図5）。このように、文字の劣化が著しいため人間でも正確な判読が難しく、機械と人間のダブルチェックで精度を上げることが目的であった。

この事例では文字劣化の度合いが非常に大きいため、劣化モデルを仮定して学習データを生成すると、例えば“6”の右上の隙間が完全に閉じて“8”に類似した字形となった画像のように、適切でない劣化画像が大量に生成されてしまう。そこで、前述したように実際の文字画像を大量に収集して学習に用いる方式を採用した。また、文字接触が非常に多いため、比較的正確に検出しやすい区切り記号を最初に検出してから文字の切り出しを行う方式を採用した。更に、書式情報としては、桁数と、各桁の出現文字種を与えた。これらのカスタマイズにより、画像内全ての文字が正しく認識できている確率を表す全文字一致率が80%以上を達成し、実用化することができた。

3.2 車両のナンバープレート認識

大規模な建設作業所においては、主として渋滞や騒音など



図5. 製品上に刻印された文字の例 — 実際に収集し2値化処理された画像のサンプルである。

Examples of characters printed on target products

の周辺への影響を避ける目的で、トラックなどの工事車両を迅速かつ正確に誘導管理することは、重要な課題である。しかし、車両情報の管理は手書きで行われていることが多く、記録作業自体の手間や、時間、情報共有に課題があった。

当社は、大成建設（株）と共同で、これらの作業を効率化する入退場車両管理システムを開発した。情景文字認識技術を用いて撮影した画像のナンバー情報を電子化し、クラウドシステム上に置いた管理台帳へ自動的に登録すると同時に、リスト照合を行い車両情報を迅速かつ正確に作業員に通知する（図6）。

この事例では、撮影画像の品質が大きな課題であった。建築作業所の出入り口は大型車両が通行するため幅が広く、入退場時の車両の角度や通行範囲は様々である。このため、単一の固定カメラで撮影し認識することは困難だが、出入り口の全範囲をカバーする数のカメラを設置するのはコスト上困難である。また、屋外のため日照の影響を受けやすく、反射光などで判読不能な画像になることがある。

これらの課題を解決する手段として、眼鏡型ウェアラブル端末を導入した。この端末の採用により、装着者が容易に撮影する位置や角度を調整することができ、また、作業員が手を自由に使えるという利点も得られる。

しかし、眼鏡型ウェアラブル端末搭載のカメラは、固定カメラに比べ解像度が低い上、ピントや、角度、撮影距離のずれが生じやすい。そこで、撮影角度に対応する3次元アフィン変換、撮影距離に対応する拡大縮小、ピントずれに対応するぼかし処理の組み合わせで劣化モデルを作成し、これに従って生成した文字画像を学習して、文字認識辞書をカスタマイズした。また、ナンバープレートは各項目の書式、出現語などが決まっているため、これらの情報も認識に用いた。このようなカスタマイズの結果、ほぼ誤りなくナンバーを認識することができるようになった。

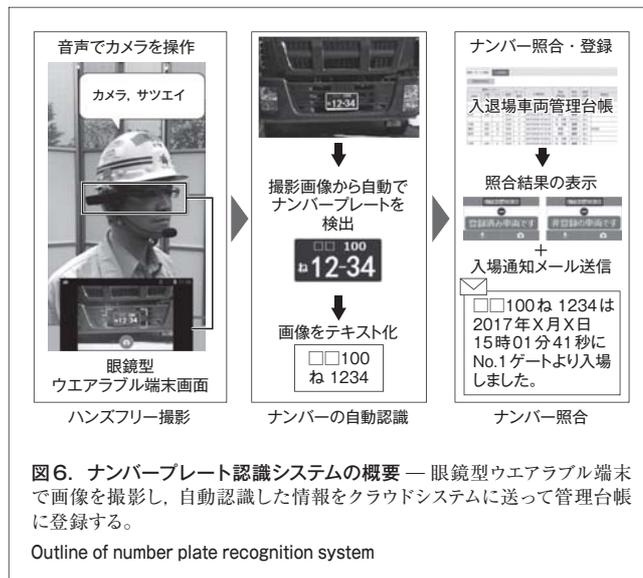


図6. ナンバープレート認識システムの概要 — 眼鏡型ウェアラブル端末で画像を撮影し、自動認識した情報をクラウドシステムに送って管理台帳に登録する。

Outline of number plate recognition system

この入退場車両管理システムでは、音声によるコマンド入力
で撮影を行う機能を実装することで、シャッターを押した際に
生じるぶれを防止している。音声入力は、カメラ撮影ができな
い状況での車両ナンバー入力機能としても用いている。認識
に使われた画像や音声は、車両情報にひも付いて車両管理台
帳に保存され、必要に応じて確認することができる。

4 今後の展望

4.1 カスタマイズの効率化

ユースケースに応じたカスタマイズで考慮すべき点は非常に
多く、誰でもがカスタマイズを容易に行うことは、現状できて
いない。今後これを誰でも行えるようにするには、二つの方法
が考えられる。

その一つは、データを大量に集め、これに基づいた機械学
習を行う方法である。ディープラーニング(深層学習)技術をは
じめ、AI技術の発展に伴ってこのような方法に注目が集まって
いるが、十分な学習データの収集及び整備が困難であること
や、学習パラメーターの調整が容易ではないことなどの問題が
残っている。しかし、最近では仮想学習データを生成するGANs
(Generative Adversarial Networks)⁽³⁾など、これらの課題を
解決する技術の研究開発も進んでいる。

もう一つは、画像処理や文字認識などで用いられるアルゴ
リズムの組み合わせ、あるいはパラメーターを自動的に最適化す
る方法である。この方法に必要な学習データはディープラー
ニング技術に比べて少なく、また、過去の知見を生かしてアルゴ
リズムやパラメーターの候補を絞り込めるので、早期に最適解
を見つけやすい利点がある。当社はこのような最適化機能を持
つ認識基盤を開発し⁽⁴⁾、カスタマイズの省力化に向けた環境
整備を進めている。

4.2 オンサイトでの認識精度向上

ものづくりの現場の環境は日々少しずつ変わっており、情景
文字認識も環境の変化に自動的に追従することが望ましい。
そのような場合に活用できる技術として、当社は半教師あり学
習の研究開発を進めている。教師あり学習では、収集した
データに人間が教示してから学習するが、半教師あり学習で

は大部分が未教示のままで学習に用いることができるため、
運用中に取得した画像を自動的に取り込んで利用環境に最適
化していくことができ、学習データへの教示数を減らしてカス
タマイズコストを削減することが可能である(図7)。

半教師あり学習では、通常、事前に学習パラメーターの決
定が必要だが、当社はこれを簡略化し、より効率的に最適化
するための研究開発を行っている⁽⁵⁾。

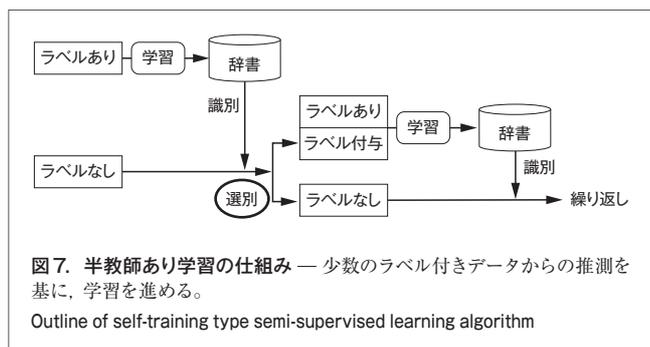
5 あとがき

ものづくりの現場で様々な情報を取得するためには情景文
字認識が有効であるが、ユースケースに応じたカスタマイズが
必要である。

今後も、ユースケースに向けたカスタマイズを進めるととも
に、カスタマイズ自体を容易にする技術、利用環境に自動的に
適応する技術を開発して、誰でも容易に利用できるサービスと
して提供していく。

文 献

- (1) Tonouchi, Y. et al. "A hybrid approach to detect texts in natural scenes by integration of a Connected-Component Method and a Sliding-Window Method". Computer Vision - ACCV 2014 Workshops, Singapore, Singapore, November 1-2, 2014, Revised Selected Papers, Part II. Jawahar, C. V.; Shan, S. eds. Springer International Publishing, 2015, p.106-118.
- (2) 鈴木智久, 中島康裕. 決定性有限オートマトンと文字候補ラテイスの照合によるフレキシブルOCR 知識処理. 東芝レビュー. 2015, 70, 4, p.46-49.
- (3) Goodfellow, I. J. et al. "Generative Adversarial Nets". Advances in Neural Information Processing Systems 27 (NIPS 2014). Ghahramani, Z. et al., eds. Neural Information Processing Systems Foundation, 2014, p.1-9. <<https://papers.nips.cc/paper/5423-generative-adversarial-nets.pdf>>, (accessed 2017-06-12).
- (4) 三宅達也, ほか. 画像認識技術の製品搭載を加速するソフトウェアプラットフォーム. 東芝レビュー. 2014, 69, 12, p.35-39.
- (5) 田中遼平, ほか. 識別境界付近のデータを選択的に用いる半教師あり学習. 信学技報. 2017, 116, 528, p.1-6.



古畑 彰夫 FURUHATA Akio

東芝デジタルソリューションズ(株) ソフトウェア& AIテクノ
ロジーセンター 知識・メディア処理技術開発部
電子情報通信学会会員
Toshiba Digital Solutions Corp.



小野 聡一郎 ONO Soichiro

東芝デジタルソリューションズ(株) ソフトウェア& AIテクノ
ロジーセンター 知識・メディア処理技術開発部
情報処理学会会員
Toshiba Digital Solutions Corp.



馮 思萌 FUU Shimou

東芝デジタルソリューションズ(株) ソフトウェア& AIテクノ
ロジーセンター 知識・メディア処理技術開発部
Toshiba Digital Solutions Corp.