

(4) 工事写真からの写真管理項目の抽出とその認識に関する研究

Research on Extraction and Recognition of
Photography Management Categories from Image Data of Construction北川悦司¹・田中成典²・安彦智史³・若林克磨⁴・姜文淵⁵

Kitagawa Etsuji, Tanaka Shigenori, Abiko Satoshi, Wakabayashi Katsuma, and Jiang Wenyan

抄録：近年、建設分野では、国土交通省の推進によって、様々な成果品の電子納品が実施されている。その中の一つである工事写真は、工事名や工種などの写真管理項目の記入が義務付けられている。しかし、写真管理項目の記入には、工事写真中の黒板に記入されている文字を目視で確認する必要があるため、手間がかかる問題がある。そのため、工事写真中の黒板の印刷文字と手書き文字を抽出し、それらを OCR で認識することで、写真管理項目の記入内容を取得する手法が求められている。そこで、本研究では、様々な前処理、OCR による文字認識と自然言語処理を用いた誤り訂正を行うことで、高精度な写真管理項目の抽出と認識を実現する。

キーワード：CALS/EC, 工事写真, デジタル写真管理情報基準, 写真管理項目

Keywords : CALS/EC, Image Data of Construction, Standard of Digital Photography Management Information, Photography Management Categories

1. はじめに

近年、建設分野では、国土交通省の推進によって、様々な成果品の電子納品¹⁾が実施されている。その中の一つである工事写真は、工事名や工種などの写真管理項目²⁾の記入が義務付けられている。しかし、写真管理項目の記入には、工事写真中の黒板に記入されている文字を目視で確認する必要があるため、手間がかかる問題がある。そのため、工事写真中の黒板の印刷文字と手書き文字を抽出し、それらを OCR (Optical Character Reader)³⁾⁴⁾で認識することで、写真管理項目の記入内容を取得する手法が求められている。しかし、黒板に記入された文字を OCR で認識するには、次の三つの課題を解決する必要がある。

一つ目は、工事写真中の黒板領域を特定することである。工事写真には、黒板以外にも様々な物体が存在する。そのため、黒板に記入された文字を認識するには、黒板領域を特定する手法が必要となる。黒板領域の特定に関する関連研究としては、物体の色情報を用いる手法⁵⁾⁶⁾がある。しかし、これらの手法では、日照や照明の影響によって、工事写真中の黒板領域に明度変化が生じる場合、物体を正しく特定できない問題がある。そのため、物体のエッジ情報を用いる手法⁷⁾や勾配方向ヒストグラムを用いる手法⁸⁾がある。しかし、これらの手法では、黒板が 3 次元的に傾いている場合、黒板領域を正しく特定できない問題がある。

二つ目は、3 次元的な文字の傾きを補正することである。3 次元的に傾いている文字を OCR で認識する場合、文字特徴が正しく取得できず、誤認識してしまう。そのため、文字を正しく認識するには、3 次元的な文字の傾きを補正する必要がある。3 次元的な文字の傾き補正に関する関連研究としては、再帰的な文字の傾斜補正を行う手法⁹⁾や文字の並びから回転角度を推定する手法¹⁰⁾がある。しかし、それらの手法では、大きさが等しい文字が規則正しく並んでいることを前提としているため、大きさにばらつきがみられる手書き文字を対象とする場合、回転角度の推定に失敗してしまう問題がある。

三つ目は、黒板領域と文字領域を正しく分離することである。黒板領域に存在するノイズによって文字領域の形状が変化する場合、文字の認識精度が低下する問題がある。そのため、文字を正しく認識するには、黒板領域と文字領域を適切に分離する手法が必要となる。黒板領域と文字領域の分離に関する関連研究としては、色情報のクラスタリングを用いる手法¹¹⁾や輝度ヒストグラムを用いる手法¹²⁾がある。しかし、これらの手法では、日照や照明の影響によって、黒板領域や文字領域の単色性が損なわれた場合、領域を正しく分離できない問題がある。

これらの課題を解決するために、本研究では、黒板領域の特定、3 次元的な文字の傾き補正、及び、黒板領域と文字領域の分離を行う。そして、分離した文字を

1 : 正会員 情博 阪南大学 准教授 経営情報学部
(〒580-8502 大阪府松原市天美東5丁目4番33号, Tel :072-332-1224, E-mail : kitagawa@hannan-u.ac.jp)

2 : 正会員 工博 関西大学 教授 総合情報学部

3 : 学生会員 関西大学大学院 総合情報学研究科

4 : 非会員 関西大学大学院 総合情報学研究科

5 : 非会員 関西大学 総合情報学部

OCR で認識することで、様々な環境で撮影された工事写真から写真管理項目の記入内容を取得するシステムを開発する。

さらに、本研究では、CALS/ECの推進に伴い、黒板に記入される工事名が入札情報としてWeb上で公開されていること、及び、工種が限定されていることに着目し、工事名欄と工種欄における文字認識結果に対し、自然言語処理を用いた誤り訂正を行うことで、実環境に適応可能なシステムを開発する。

2. 研究の概要

本研究では、まず、工事写真中の黒板を抽出する。次に、抽出した黒板の正面画像を生成する。そして、正面画像から文字を抽出する。さらに、抽出した文字に対し、OCRによる文字認識と自然言語処理を用いた認識結果の誤り訂正を行い、写真管理項目の記入内容を取得する。本システムの概要を図-1に示す。本システムは、1) 黒板領域抽出機能、2) 正面画像生成機能、3) 文字領域抽出機能と4) 文字認識機能で構成される。入力データは、黒板が写る工事写真と黒板のテンプレート画像とし、出力データは、写真管理項目の記入内容とする。

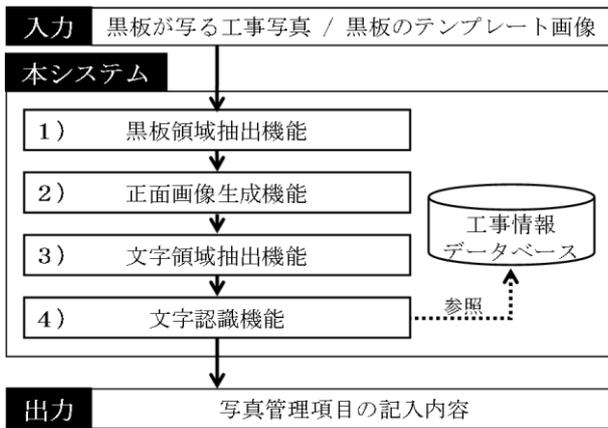


図-1 本システムの概要

(1) 黒板領域抽出機能

本機能では、文字が存在する領域を特定するために工事写真から黒板領域を抽出する。工事写真は屋外での撮影が想定されるため、高精度に黒板領域を抽出するためには、明度変化や3次元的な傾きに対して頑健に対応する必要がある。そこで、本研究では、まず、工事写真を日照や照明の影響に頑健な2値画像の生成手法であるNAT (Noise Attribute Thresholding) 法¹³⁾を用いて領域ごとに分離する。次に、分離した領域に対してSURF (Speeded Up Robust Features)¹⁴⁾を用いて特徴点を発生させる。SURF特徴を用いることで、スケール変化、3次元的な傾きと明度変化に頑健な特

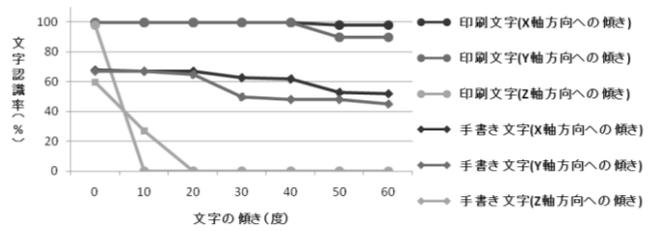


図-2 傾きに伴う文字認識率の変化(補正前)

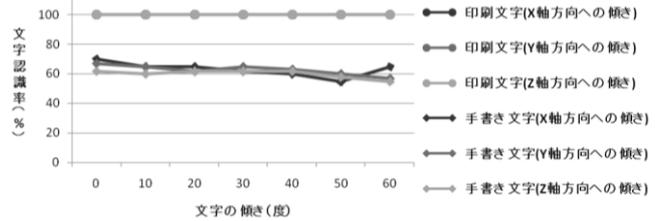


図-3 傾きに伴う文字認識率の変化(補正後)

徴点を発生させることができる。最後に、黒板のテンプレート画像から事前に発生させたSURF特徴と最も類似度の高い領域を黒板領域として抽出する。これにより、明度変化や3次元的な傾きに頑健な黒板抽出を実現する。

(2) 正面画像生成機能

本機能では、3次元的な文字の傾きを補正するために、抽出した黒板領域から黒板の正面画像を生成する。黒板領域に3次元的な傾きが生じている場合、黒板に記入された文字を正しく認識できない問題がある。本研究では、3次元的な文字の傾きがOCRによる文字認識にどのように影響するかを確認するため、事前実験を行う。実験では、黒板をX軸、Y軸とZ軸方向に0度から60度まで10度ずつ回転させて撮影した21枚の工事写真から印刷文字と手書き文字を手作業で抽出し、それらをOCRで認識する。文字の認識には、事前実験の結果、文字認識精度が最も高かったA社のOCRを使用する。事前実験の結果を図-2に示す。結果から、傾きによって印刷文字と手書き文字の認識精度が低下することがわかる。この原因としては、傾きによって文字が潰れてしまい、文字の形状が大きく変化してしまったことが考えられる。本研究では、この問題を解決するため、黒板領域の正面画像を生成することで、3次元的な文字の傾きを補正する。まず、黒板領域の輪郭線追跡を行い、輪郭線を抽出する。次に、抽出した輪郭線に対してハフ変換を行い、黒板の外枠を構成する直線を算出する。そして、それらの交点を角点として抽出する。最後に、角点から平面射影変換行列¹⁵⁾を算出することで黒板の正面画像を生成する。

本手法の有用性を確認するために、事前実験で使用した21枚の工事写真中の黒板領域から正面画像を生成し、それらの正面画像から抽出した文字をOCRで認識する。認識結果を図-3に示す。結果から、本手法

を用いることで、3次元的な文字の傾きが補正され、文字認識率が向上していることがわかる。

(3) 文字領域抽出機能

本機能では、OCRで認識する文字を取得するために、黒板の正面画像から記入欄ごとの文字領域を抽出する。工事写真は屋外での撮影が想定されるため、黒板領域と文字領域を適切に分離するためには、明度変化に対して頑健に対応する必要がある。そこで、本研究では、まず、黒板の正面画像を日照や照明の影響に頑健な2値画像の生成手法であるNAT法を用いて領域ごとに分離する。次に、ラベリングを行い、面積が最大であるラベルを黒板の罫線として抽出する。そして、黒板の罫線を基に記入欄を特定することで記入欄ごとの文字候補領域を抽出する。最後に、文字候補領域間の重心の距離を基に統合処理を行い、黒板の記入欄ごとの文字領域を抽出する。

(4) 文字認識機能

本機能では、OCRによる文字認識と自然言語処理を用いた認識結果の誤り訂正を行うことで、写真管理項目の記入内容を出力する。手書き文字は、個人による形状の変化が見られるため、OCRで認識する場合、認識精度が低下する。例えば、「高槻工事」という単語の場合、「高槻土事」と認識するなど形状に近い文字を誤認識する問題がある。そこで、筆者らは、工事名欄と工種欄に記入された文字の認識結果に対し、自然言語処理による誤り訂正を行う。本研究では、CALS/ECの推進に伴い、公共機関がWeb上に工事の入札情報を公開していることに着目し、入札情報に含まれる情報から工事情報データベースを構築する。そして、OCRによる認識結果と工事情報データベースに含まれる候補単語の類似度を比較することにより、誤り文字の訂正を実現する。尚、類似度の比較では、認識結果を構成する単語群と工事情報データベースに含まれる単語の部分一致率が閾値以上である単語を候補文字とする。

本手法の有用性を確認するために、文字列の誤り訂正の実験を行う。実験では、0%から90%まで10%ずつの割合で任意の文字を空白に置換した文字列に対し、本手法を用いて誤り訂正を行った際の文字認識率を確認する。工事名欄向けの入力データには、2010年度における大阪府、京都府と兵庫県の建設工事の入札情報として公開されている全6646件の工事名に対し、任意の文字を空白に置換したものを使用する。また、工種欄向けの入力データには、新土木積算体系に定義されている全166件の工種に対し、任意の文字を空白に置換したものを使用する。実験結果を図-4と図-5に示す。結果から、文字列に50%の割合で誤りが含まれる場合において、工事名の誤りを約60%の精度で、工種の誤りを約80%の精度で訂正できることがわかる。

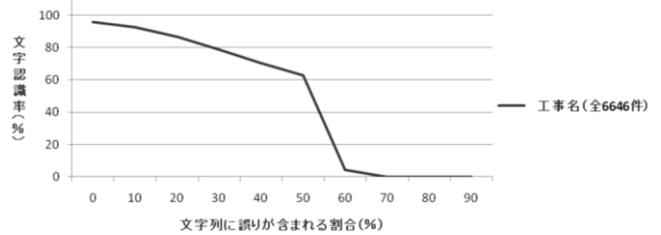


図-4 認識結果の誤り訂正の実験結果(工事名)

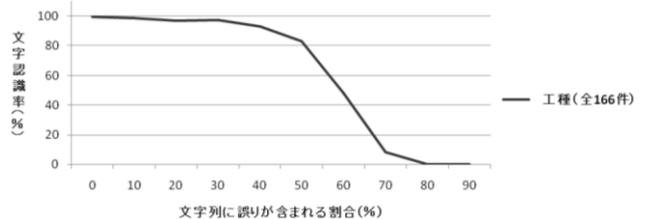


図-5 認識結果の誤り訂正の実験結果(工種)



図-6 実証実験で使用する工事写真の一部

3. システムの実証実験

実証実験では、黒板の抽出精度、正面画像の生成精度、文字の抽出精度と文字認識結果の誤り訂正精度を検証し、本システム全体の有用性を確認する。

(1) 実験内容

本実験では、様々な環境で撮影された工事写真から黒板領域を正しく抽出できることを実証するために、黒板領域にスケール変化、3次元的な傾きや明度変化が存在する全100枚の工事写真(図-6)を用意し、黒板の抽出精度を検証する。次に、黒板領域から正しく正面画像を生成できることを実証するために、正しく抽出できた全90枚の黒板に対する正面画像の生成精度を検証する。そして、黒板領域から文字領域のみを正しく抽出できることを実証するために、正しく生成できた全90枚の正面画像に対する文字の抽出精度を検証する。最後に、誤り訂正によって文字認識精度が向上することを実証するために、正しく抽出できた全2261文字をOCRで認識する。そして、工事名欄と工種欄における文字認識結果に対し、自然言語処理を用いた誤り訂正を行うことで、それらの誤り訂正精度を検証する。文字認識には、事前実験の結果、文字認識精度が高かった3社のOCRを使用する。

(2) 結果と考察

黒板抽出の実験結果を表-1に示す。結果から、90.00%の精度で黒板を抽出できた。しかし、本手法では、黒板の白い外枠と背景領域を分離することで、黒板領域を抽出するため、標尺などが外枠を遮蔽する場合、黒板を正しく抽出できなかった。この解決策としては、黒板の外枠に加えて、黒板の色情報を用いて背景領域と黒板領域を分離することが考えられる。次に、正面画像生成の実験結果を表-2に示す。結果から、100.00%の精度で正面画像を生成できた。そして、文字抽出の実験結果を表-3に示す。結果から、96.25%の精度で文字を抽出できた。しかし、隣り合う文字同士が接触している場合、正しく文字を抽出できなかった。この解決策としては、自然言語処理を用いた文字候補領域の適切な分割と統合が考えられる。最後に、誤り訂正の実験結果を表-4と表-5に示す。結果から、誤り訂正によって文字認識率が向上したことから、本手法の有用性を確認できた。しかし、本手法では、工事情報データベースに正解データが存在しない場合、誤り訂正できない。この解決策としては、言語モデルを用いて認識結果の誤りを部分的に訂正することが考えられる。

表-1 黒板抽出の実験結果

	工事写真数	抽出成功	抽出率
本手法	100枚	90枚	90.00%

表-2 正面画像生成の実験結果

	黒板数	生成成功	生成率
本手法	90枚	90枚	100.00%

表-3 文字抽出の実験結果

	総文字数	抽出成功	抽出率
本手法	2349文字	2261文字	96.25%

表-4 文字認識の実験結果(誤り訂正前)

	総文字数	認識成功	認識率
OCR(A社)	2261文字	1920文字	84.92%
OCR(B社)	2261文字	1838文字	81.29%
OCR(C社)	2261文字	1707文字	75.50%
平均	2261文字	1822文字	80.58%

表-5 文字認識の実験結果(誤り訂正後)

	総文字数	認識成功	認識率
OCR(A社)	2261文字	2144文字	94.83%
OCR(B社)	2261文字	2116文字	93.59%
OCR(C社)	2261文字	2004文字	88.63%
平均	2261文字	2088文字	92.35%

4. おわりに

本研究では、黒板領域の特定、3次元的な文字の傾き補正、及び、黒板領域と文字領域の分離を行った。そ

して、分離した文字に対し、OCRによる文字認識と自然言語処理を用いた認識結果の誤り訂正を行うことで、様々な環境で撮影された工事写真から写真管理項目の記入内容を取得するシステムを開発した。そして、実証実験から本手法の有用性を確認した。今後の予定としては、黒板の外枠の遮蔽に頑健な黒板領域抽出手法、自然言語処理を用いた文字候補領域の適切な分割と統合による文字領域抽出手法と文字認識結果の部分的な誤り訂正手法を考案することで、写真管理項目の抽出、及び、認識精度の向上を目指す。

参考文献

- 1) 国土交通省: CALS/EC アクションプログラム 2008, 2008年.
- 2) 国土交通省: デジタル写真管理情報基準, 2010年.
- 3) 中野康明: 文字認識・文書理解の最新動向[I], 電子情報通信学会論文誌, 電子情報通信学会, Vol.83, No.1, pp.64-68, 2000年.
- 4) 坂井邦夫, 入江文平, 水谷博之: 文字・文書の認識と理解, 情報処理, 情報処理学会, Vol.44, No.11, pp.1123-1129, 2003年.
- 5) 内村圭一, 中村国章, 富永裕之, 胡振程, 脇阪信治, 有田秀和: 道路情報整備のための案内標識の位置特定, 電子情報通信学会論文誌, 電子情報通信学会, Vol.J88-A, No.2, pp.122-130, 2005年.
- 6) 森稔, 柏野邦夫: 適応的領域分割と初期しきい値推定によるテンプレートマッチングの高速化, 電子情報通信学会論文誌, 電子情報通信学会, Vol.94-D, No.5, pp.881-892, 2011年.
- 7) Bo, W., and Nevatia, R.: Detection of Multiple, Partially Occluded Humans in a Single Image by Bayesian Combination of Edgelet Part Detectors, Proceedings on the Tenth IEEE International Conference on Computer Vision, Vol.1, pp.90-97, 2005.
- 8) 山口晃一郎, 内藤貴志: 歩行者検出のためのこう配方向の二次元ヒストグラム特徴, 電子情報通信学会論文誌, 電子情報通信学会, Vol.J94-D, No.1, pp.365-373, 2011年.
- 9) 林信吾, 志久修, 黒田英夫, 玉井敬一, 福本博文: 射影ひずみを受けた活字文字列認識の一方法, 電子情報通信学会論文誌, 電子情報通信学会, Vol.J91-D, No.2, pp.512-515, 2008年.
- 10) Wang, K., and Kangas, J.: Character location in scene images from digital camera, Proceedings on Pattern Recognition, Vol.36, No.10, pp.2287-2299, 2003.
- 11) 芦田和毅, 永井弘樹, 岡本正行, 宮尾秀俊, 山本博章: 情景画像からの文字抽出, 電子情報通信学会論文誌, 電子情報通信学会, Vol.J88-D-II, No.9, pp.1817-1824, 2005年.
- 12) 山口拓真, 丸山稔: 階層型識別器を用いた情景画像からの文字抽出手法, 電子情報通信学会論文誌, 電子情報通信学会, Vol.J88-D-II, No.6, pp.1047-1055, 2005年.
- 13) Hon-Son, D.: A noise attribute thresholding method for document image binarization, International Journal on Document Analysis and Recognition, Vol.4, Issue.2, pp.131-138, 2001.
- 14) Herbert, B., Andreas, E., Tinne, T. and Luc, G.: Speeded Up Robust Features (SURF), Computer Vision and Image Understanding, Elsevier, Vol.110, No.3, pp.346-359, 2008.
- 15) 杉本茂樹, 奥富正敏: ステレオ画像を用いた高速な平面パラメータ推定法, 情報処理学会論文誌, 情報処理学会, Vol.48, No.SIG1(CVIM17), pp.24-34, 2007年.