

適応的撮像に基づく自動・高速・高精細書籍電子化システムの開発と評価

Automatic, High-speed and High-resolution Book Digitization System Based on Adaptive Capture

○ 野口 翔平(東大) 正 溜井 美帆(東大) 山田 雅宏(東大)
正 渡辺 義浩(東大) 正 石川正俊(東大)

Shohei NOGUCHI, Univ. of Tokyo, Shohei_Noguchi@ipc.i.u-tokyo.ac.jp

Miho TAMEI, Univ. of Tokyo, Masahiro YAMADA, Univ. of Tokyo,

Yoshihiro WATANABE, Univ. of Tokyo and Masatoshi ISHIKAWA, Univ. of Tokyo

This paper proposes a new book digitization system that enables to obtain high-resolution document images while the pages are flipped automatically. This system provides two distinctive features including the adaptive capture and the automatic pager turner. This adaptive capture requires observing the state of the flipped pages at high speed and with high accuracy. In order to meet this requirement, we newly propose a method to obtain the 3D shape of the book, track each page and evaluate the state. We explain the details about the proposed high-speed book digitization system and the effective method of the real-time 3D image processing for the adaptive capture. Also the verifications of the developed system are reported.

Key Words: Book digitization, High speed 3D sensing, Automatic page turner machine, Adaptive capture

1. はじめに

近年，電子化された書籍を読む機会が増えてきている．これとともに，紙媒体で保管されている各種書籍の電子化による保全が注目されている．しかし，従来技術の速度はこのようなニーズに対して十分な性能を有していないと考えられる．

このような背景の下，高速かつ非破壊かつ手軽な新たな電子化の方式として，書籍のめくり動作中にその動きを止めることなく連続的に書籍情報を取り込む「ブックフリッピングスキャニング」が提案された [1]．

本稿では，ブックフリッピングスキャニングの構想の下，高速性，高精細，及び自動化のすべてを達成する書籍電子化システムについて報告する．具体的には，独自に開発しためくり機 [2] と高速性と高精細を両立するための適応的撮像タイプ [3] の2つの統合を開発方針のベースとしている．本稿では，新しく開発したシステムを報告するとともに，適応的撮像のための書籍の形状・状態認識手法を新たに提案する．実験では，300 ページ/分程度の速度でめくられる書籍をターゲットとして，システムの性能評価を行った．

2. 自動めくり機を搭載する高速書籍電子化システム

現在，様々な書籍電子化システムが開発・市販化されている．しかし，多くのシステムは書籍の種類を限定しても，20～50 ページ/分程度であり，十分な速度を有していなかった．3次元計測を行い，書籍電子化を行うタイプのシステム [4, 5, 6] も報告されているが，高速化は難しかった．また，本システムと同じように，速度と解像度が異なるカメラを複合的に利用するシステム [7, 8] も提案されているが，高品質な書籍電子化への適用は難しいと考えられる．

本システムは，めくり機によって書籍を1ページずつ自動でめくり，各ページの書籍画像をその変形の動作に応じて自動的に取り込むことができる．システムの動作の概要を下記に述べる．まず，めくられた書籍のページの変形は，構造化光法の原理の下，高速な3次元センシングによってリアルタイムに取得される．取得された3次元形状は，各ページに分割

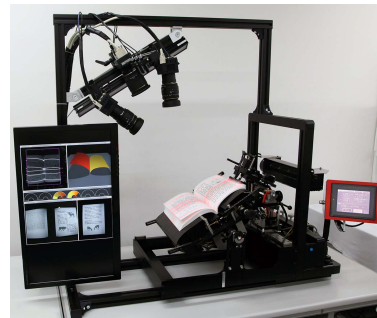


Fig. 1 High-speed document digitization system with automatic page turner machine

され，その変形の状態が評価される．この評価の結果，高解像度カメラの視点で，書籍電子化のために最適となるタイミングが検出される．検出されたタイミングにおいて，高解像度カメラによって，高精細な書籍画像が適応的に撮像される．さらに手法 [9] によって，取得された形状と高解像度カメラの撮像画像を用いて，歪みを補正している．なお，本システムは，めくり機と処理部の間で連携を取ることなく，正確なめくり動作と電子化を達成することができる．

開発したシステムの写真を図 1 に示す．サイズは，130cm×60cm×110cm である．本システムの高速度カメラのフレームレートは，各フレーム間での紙面の縦じり方向を軸とする回転が 300 ページ/分のめくり速度で 1 度以内であるように，400fps のものを用いた．また解像度は 1,280 × 1,024 である．一方，高解像度カメラのフレームレートは，書籍の限界的なめくり速度に対して，1 ページ 1 回撮像できれば良い．本システムでは，目標ページめくり速度を 300 ページ/分としており，2 台で分担して撮像する場合には 2.5fps 以上が要求される．書籍は均等なタイミングでめくられるわけではないため，本システムでは 5fps のカメラを用いた．また，解像度は電子化の一般的な要求解像度である 600dpi 程度であれば良い．本システムでは，A4 サイズまでの書籍を対象として，解像度 6,584 × 4,384 のカメラを用いた．

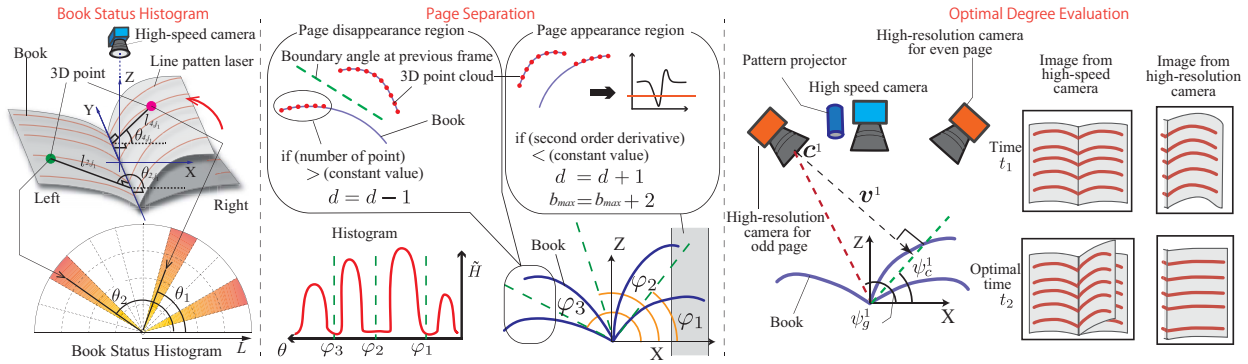


Fig. 2 Overview of the proposed method.

さらに、その他のシステムの構成要素として、3次元センシング用の構造化光照射プロジェクタ、高解像度カメラ撮像用のストロボ照明、処理のための計算機、及び各デバイスを信号制御する組込みボードがある。本システムでは、構造化光として5本の赤色ラインパターンレーザを用いている。レーザの仕様は、波長660nm、投光角60度である。使用した計算機の構成は、CPU Xeon X5675、メモリ56GBである。

3. 適応的撮像のためのリアルタイム3次元書籍状態認識

3.1 書籍形状の特徴抽出

本手法では、構造化光としてマルチラインパターンを投影して3次元の形状を取得する。本設計では、書籍が変形する範囲を予め計測し、撮像画像上の輝点がどのラインかを一意に決定できるように投影位置と向きを調整した。これによって、高速かつ正確に3次元点群 $P = \{p_{i,j} : i = 1, \dots, \alpha, j = 1, \dots, \beta_i\}$ を計算している。図2の左に示されるように、書籍の綴じ部と平行にY軸を設定し、Z軸を綴じ部の中心点から高速カメラへの向きとする座標系を設定した。ここで添え字 i はM本のラインパターンのうち i 番目のラインパターンに属していることを示し、 j はその β_i 点のなかで j 番目の点を示している。

適応的撮像では、撮像に最適なタイミングを知るために時間的に高分解能な書籍状態の認識が必要となる。また、同認識は計測ノイズにロバストであることが重要である。従来は、高速に計算可能な曲面モデルを用いていたが、同モデルでは様々な紙面変形を正確に表現できないため、誤認識を起こす問題があった[3]。

そこで、本手法では、各3次元点 $p_{i,j}$ について、図2左に示されるように綴じ部を通る直線と点の距離 $l_{i,j}$ と、綴じ部を軸とした回転角度 $\theta_{i,j}$ の2変量を利用する。全ての点群に対して計算された2変量を用いて、Book Status Histogram (BSH)

$H(s, t)$ を計算する。点 $p_{i,j} = (x, y, z)$ は下式によって計算される階級 (s, t) に含まれる。

$$s = \left\lfloor \frac{\sqrt{x^2 + z^2}}{L/N_L} \right\rfloor, t = \left\lfloor \frac{\arctan(z/x)}{\pi/N_\theta} \right\rfloor \quad (1)$$

ただし、距離方向の階級数を N_L 、階級幅を紙面の横幅 L を用いて、階級幅を L/N_L とする。また、角度方向の階級数を N_θ とし、階級幅を π/N_θ とする。なお、 $\lfloor \cdot \rfloor$ は天井関数である。このBSHによって、書籍形状の状態を認識する。

3.2 ページトラッキング

提案するページ分割手法の概要を図2の中央に示す。本手法では、図中の右下部で示した点線のように綴じ部からの直線によって点群を各ページを分割する。

まず、各時刻の観測形状に対して、高速カメラから観測されるページ数 d と、これらのページのうち最大のページ番号 b_{max} を算出する。観測されるページ数 d は、ページ分割のために必要となる。最大のページ番号 b_{max} は、ページ番号の割り当てのために必要となる。ページ番号は、通常の書籍のページ番号と同様に、紙面の表裏にそれぞれ割り当てる。今回は、表側が奇数ページ、裏側が偶数ページとする。

ここで、書籍は右から左にページがめくられていると仮定する。このとき、新しいページは、書籍右端から現れる。この端の領域をページ出現領域と呼ぶ。また、ページは左端で消失する。この領域をページ消失領域と呼ぶ。この2つの領域内の形状を計測することで、 b_{max} と d を算出している。

次に、算出されたページ数とBSHを用いて、3次元点群をページごとに分割する。前節で算出した $H(s, t)$ を用いて、1次元ヒストグラム $\hat{H}(t) = \sum_s H(s, t)$ を定義する。 $\hat{H}(t)$ は、図2の中央の左下に示される赤曲線のような形になる。そこで、同図の破線のようにページ数 d に分割し、ページの境界角度 $\{\varphi_u : u = 1, \dots, d-1\}$ を算出する。この境界角度 φ_u と φ_{u+1} に対して、書籍のページ番号は連続的であるため、この範囲内にある点群のページ番号 b_u は下式となる。

$$b_u = b_{max} - 2u + k \quad (k = 0 \text{ or } 1) \quad (2)$$

ただし、 k は0のときに偶数ページ、1のときに奇数ページを表す。この偶奇は、面の法線を利用することで判定することができる。しかし、システム的设计上、各々の高解像度カメラが偶数、または奇数それぞれのページのみを最適なタイミングで捉えるようになっているため、適応的撮像後に割り当てることが可能であり、同判定をこの段階で行うことは必須ではない。

3.3 ページの状態評価

本システムでは、適応的撮像のための高速カメラと高解像度カメラを図2右のように配置した。例えば、同図のように、高速カメラの視点で大きく変形している場合でも、高解像度カメラの視点では歪みの少ない紙面の状態が捉えられている瞬間となる。このため、変形の状態やページの遮蔽関係を高解像度カメラを基準に考える必要がある。

まず、書籍の偶奇ページを撮像する高解像度カメラの位置を c^k 、向きを v^k とする。添え字 k は前述のとおり、ページの偶奇を表す。このとき、各ページの高解像度カメラに対して最適な撮像タイミングへの適合度 $f_{b_u}^k(H)$ を式(3)によって算出する。

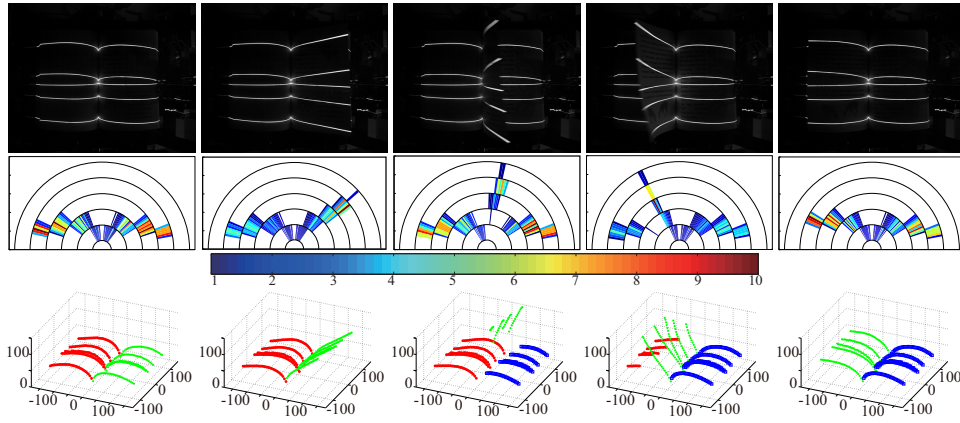


Fig. 3 Results of page tracking. Top: Images from high speed camera. Middle: Book status histogram. Bottom: 3D point cloud.

$$f_{b_u}^k(H) = \frac{\sum_s \sum_t A_{b_u}^k(s, t) \times H(s, t)}{\sum_s \sum_t H(s, t)} \quad (3)$$

この適合度が閾値 γ^k に対して、 $f_{b_u}^k(H) > \gamma^k$ になったとき、制御ボードから高解像度カメラに撮像指令を送り、適応的撮像を行う。なお、適合度は偶数ページ、奇数ページに対する値を個別に計算している。

式(3)で $A_{b_u} \in \mathbb{R}^{N_L \times N_\theta}$ をページ状態マスクと呼ぶ。まず、図2右のように、偶数ページ用の高解像度カメラの撮影最適角度 ψ_c^k と撮影範囲 ψ_g^k を求める。最適撮像角度として、本システムでは高解像度カメラに対して垂直になる角度を採用した。これは高解像度カメラに対して、紙面が垂直に位置しているときが解像度の面で有利と考えられるためである。また、識別されたページ番号を利用して、高解像度カメラで各ページを一度のみ撮像する。

奇数ページについては、 $\varphi_u < s\pi/N_\theta < \varphi_{u+1}$ かつ $s\pi/N_\theta < \psi_g^1$ ならば、下式とする。偶数ページについては、 $\varphi_u < s\pi/N_\theta < \varphi_{u+1}$ かつ $s\pi/N_\theta < \psi_g^0$ ならば下式とする。

$$A_{b_u}^k(s, t) = a_k \exp\left(\frac{(t - \lceil \psi_c^k N_\theta / \pi \rceil)^2}{2(\sigma(1 - s/2N_L))^2}\right) \quad (4)$$

a_k はページ状態マスクの倍率を指定する任意定数である。なお、 σ は最適撮像角度に対する適合度の厳密度を判別するためのパラメータである。 σ が小さいほど、式(4)の値が大きい範囲が狭くなる特性を備えている。この値を小さくすることで、適合度の遷移が急伸的になり、より最適なタイミングでのみ撮像するように調整できる。しかし、簡便な閾値処理では高解像度カメラの撮像に失敗する問題がある。パラメータ σ はこれらのトレードオフを調整して、予め決定される。

また、奇数ページについて、 $\varphi_{u+1} < s\pi/N_\theta$ かつ $s\pi/N_\theta < \psi_g^1$ 、または、偶数ページについて、 $\varphi_u > s\pi/N_\theta$ かつ $s\pi/N_\theta > \psi_g^0$ ならば、高解像度カメラから見たときに注目しているページを遮蔽する可能性のある領域なので下式とする。

$$A_{b_u}^k(s, t) = -a_2 \quad (5)$$

a_2 は遮蔽領域重みを表す任意定数である。

それ以外の領域については $A_{b_u}^k(s, t) = 0$ とする。

本手法は、従来手法[3]と比較して、モデルによって表現できる形状に制約がなく、ページ分割もロバストに処理でき、ページめくり動作に対しても強い制約を必要としないという利点がある。

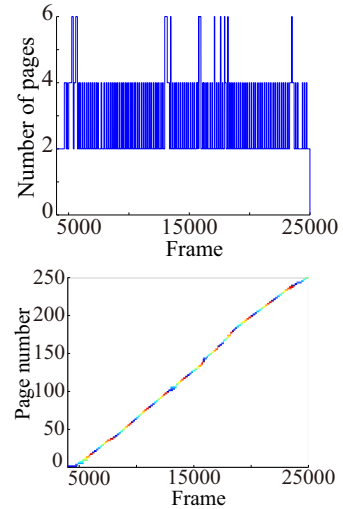


Fig. 4 Result of page number and number of pages in each frame.

4. 実験

4.1 実験条件

対象として 81g/cm^2 の紙で作成された 250 ページの B5 書籍を用いた。自動めくり機で 312 ページ/分の速度で書籍をめくり、前節の提案手法によって適応的撮像による書籍の電子化を行った。このとき、1 フレームの処理に要した時間は平均で 2.3ms であった。

処理時のパラメータは、 $a^0 = a^1 = 3$ 、 $a^2 = 2$ 、 $\sigma = 3$ 、ヒストグラムの分割数は長さ方向に $N_L = 5$ 、角度方向に $N_\theta = 100$ とした。評価値の閾値は奇数ページについては $\gamma^0 = 0.35$ 、偶数ページについては $\gamma^1 = 0.3$ とした。3 次元点群のサンプリング間隔は画像上で水平方向に 15 ピクセルごととした。

4.2 ページトラッキング結果

図3に本手法を用いたページのトラッキング結果を示す。上段は高速カメラで撮像された画像、中段は計測された 3 次元点群から作成された BSH、下段は 3 次元点群をページごとに分割したものである。最も左の図の瞬間を時刻 0 としたとき、左から順に 137.5ms、155ms、182.5ms、260ms 後の様子である。BSH について、点群数と色の濃度の関係は図3中段に示している。なお、可視化されたヒストグラムで白色の領域は点群数が 0 である。ページ分割の図ではトラッキングされた各ページが異なる色とマーカーで示されている。図に示されるように、本手法によってリアルタイムにページをトラッキングできていることが確認できる。

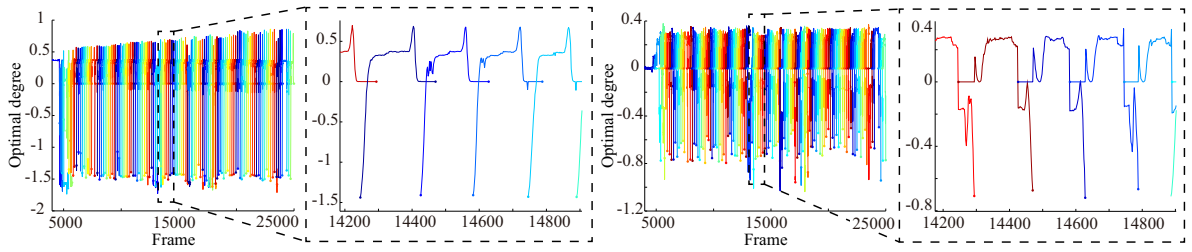


Fig. 5 Result of optimal degree of odd and even pages. Left: All the pages of a book. Right: Enlarged view of 2 second.

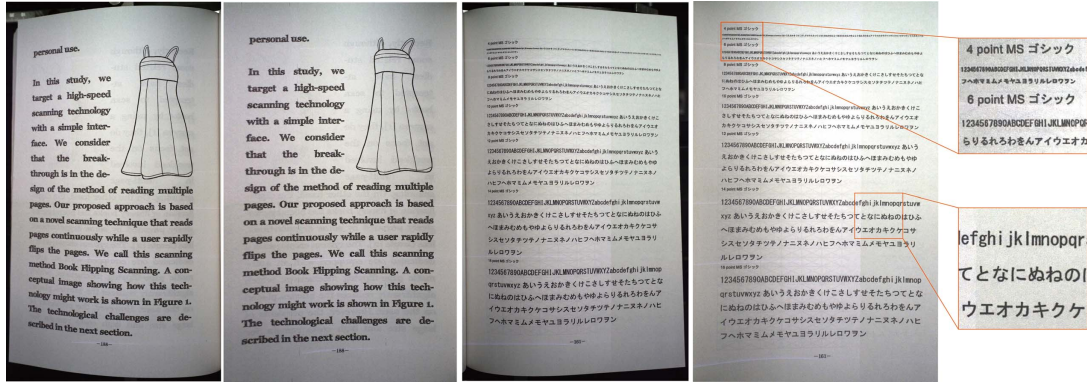


Fig. 6 Results of captured image with adaptive approach.

また、図4にページ番号とページ数の遷移を示す。それぞれ、横軸をフレーム番号、縦軸をページ数とページ番号とした。

4.3 評価関数と適合度

図5に各時刻における偶奇ページ用の高解像度カメラから見たページの適合度の遷移を示す。各図において、左は書籍1冊分の適合度の推移、右はそのうち約2秒間の遷移を切りだした拡大図となっている。各曲線は異なるページで異なる色を用いて表している。図に示されている通り、自動めくり機を導入したことにより、比較的安定して書籍がめくられていることがわかる。しかし、拡大図によれば、各ページは少しずつ異なる紙面変形していることがわかり、またピークのタイミングも均等に並んでいない。そのため、高解像度カメラの単純な均等タイミングでの連写では、書籍電子化に十分な解像度で画像が得られないと考えられる。

図6に適応的撮像によって得られた高解像度カメラ画像と、その画像を同時刻に取得された3次元形状から可展面を推定して補正する手法[9]によって平面に展開補正した画像の例を示す。左から順に1番目は偶数ページを高解像度カメラで適応的撮像したものであり、2番目はその画像を展開補正した画像である。同様に、右側の画像は奇数ページの例を示す。図のように、最適な瞬間に撮像することで、高解像度な書籍画像を取得し、展開補正によって高品質な書籍電子化が実現できた。

4.4 繰り返しに対する精度評価

紙質の異なる2種類(64g/m², 84g/m²)の250ページのB5書籍を用いて、繰り返し実験を行った。撮像ができなかった場合と、遮蔽によって対象のページのうち1文字でも確認できない場合の2種類を撮像失敗と定義した。繰り返し回数は各書籍5回ずつとした。撮像失敗ページの総数はそれぞれ9, 16ページで、撮像成功率は99.28, 98.72%であった。また平均のページめくり速度はそれぞれ324.8, 298.6ページ/分であった。以上より、2種類の紙質で同様の高い精度で書籍電子化できていることがわかる。

5. まとめ

本稿では、高速・高精細を両立した完全に自動な書籍電子化システムを提案した。本システムは、ハイブリッドな3眼のマルチカメラ構成によって高速・高精細を両立するとともに、高速かつ自動のめくり操作を有しており、高速書籍電子化の総合的機能を搭載している。特に、高速で計測ノイズにロバストな適応的撮像のための手法を報告した。本手法は、書籍状態ヒストグラムと呼ぶ新たな特徴量を利用することで、書籍ページのトラッキングや撮像タイミングに関する適合度を400fpsの速度でリアルタイムに計算可能にしたものである。実験の結果、めくり速度約300ページ/分に対して、正確に適応的撮像ができることを確認した。

謝辞

本研究の一部は戦略的情報通信研究開発推進制度(SCOPE)の下で行われた。

文献

- [1] T. Nakashima, Y. Watanabe, T. Komuro, and M. Ishikawa. Book flipping scanning. In *22nd Symposium on User Interface Software and Technology*, pp. 79–80, 2009.
- [2] 溜井美帆, 山田雅宏, 渡辺義浩, 石川正俊. 高速書籍電子化のための高速書籍自動めくり機の設計と評価. 第30回日本ロボット学会学術講演会講演論文集, pp. 4G3–8, 2012.
- [3] 有間英志, 糸山浩太郎, 山田雅宏, 小室孝, 渡辺義浩, 石川正俊. 高速3次元センシングによる適応的撮像を行う高精細書籍電子化システムの提案. 第17回画像センシングシンポジウム講演論文集, pp. IS2–17, 2011.
- [4] A. Yamashita, A. Kawarago, T. Kaneko, and K. T. Miura. Shape reconstruction and image restoration for non-flat surfaces of documents with a stereo vision system. In *Proceedings of ICPR*, pp. 482–485, 2004.
- [5] Z. Zhang, C. L. Tan, and L. Fan. Estimation of 3D shape of warped document surface for image restoration. In *Proceedings of ICPR*, pp. 486–489, 2004.
- [6] M. S. Brown and C. J. Pisula. Conformal deskewing of non-planar documents. In *Proceedings of CVPR*, pp. 998–1004, 2005.
- [7] Y. D. Wang, I. Ishii, T. Takaki, and K. Tajima. An intelligent high-frame-rate video logging system for abnormal behavior analysis. *Journal of Robotics and Mechatronics*, Vol. 23, No. 1, pp. 53–65, 2011.
- [8] 重本倫宏, 星川章, 長原一, 岩井儀雄, 谷内田正彦, 鈴木俊哉. 時間的・空間的分解能の異なる複合センサカメラシステム. 情報処理学会論文誌, Vol. 47, No. SIG 5, pp. 35–44, 2006.
- [9] Yoshihiro Watanabe, Kotaro Itoyama, Masahiro Yamada, and Masatoshi Ishikawa. Digitization of deformed documents using a high-speed multi-camera array. In *The 11th Asian Conference on Computer Vision*, 2012.