

バレットタイムブック：多視点映像提示を用いた書籍の視覚情報の拡張

Qiu XinYi† 宋戸 英彦‡ 亀田 能成‡ 北原 格‡

†筑波大学 大学院システム情報工学研究科 ‡筑波大学 計算科学研究センター

1. はじめに

映像メディア技術の発展により書籍媒体の表現・閲覧法が多様化している。従来の雑誌や書籍では文字や画像を用いた表現が主だったが、電子書籍の登場により、紙媒体では困難な新たな表現・閲覧法が実現されている。また、画像・映像情報の提示法に関する研究開発も活発に行われている。バレットタイム映像は、被写体の周りに複数台のカメラを配置し、多視点映像を順番に切り替えながら観察することにより、運動視差に基づいた3次元情報を提示する手法である。書籍に含まれる写真や図の表現・閲覧にバレットタイム映像を導入することで、2次元画像では表現しにくい3次元情報の表現が可能となる。

QRコードは、コードをスキャンすることにより埋め込まれたマルチメディア情報へのリンクを容易に取得することができるが、マーカー印刷により本のデザインが損なわれるという問題が存在する。書籍に印刷された図や写真をそのままマルチメディア情報へのリンク情報として利用することにより、上述した問題の解決が可能である。例えば、Magic Book[1]では、書籍を撮影した画像を用いて、対応するリンク情報を検出し、拡張現実提示を実現している。

本稿では、書籍を撮影した画像中に含まれる図や写真などを認識し、その結果を用いて、データベースに登録されたバレットタイム映像を検索・再生する方式を提案する。

2. バレットタイムブック (BT-Book)

図1に示すように、提案するBT-Bookは、書籍とタッチパネルインタフェースを装備した計算機(タブレット端末)で構成される。画像特徴量を用いて書籍を撮影した画像中に含まれる図や写真を認識し、その結果からデータベースに登録されたバレットタイム映像を検索し、タブレット端末で再生する。

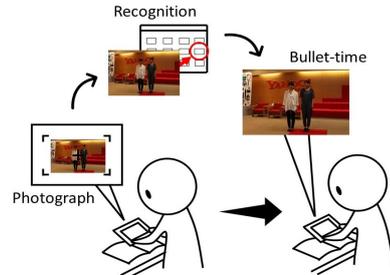


図1 提案する書籍視覚情報の拡張法

BT-Bookは、大きく”撮影”,”前処理”,”認識処理”,”閲覧”の四つの処理からなる。まず、ユーザが書籍画像を撮影する(Step1)。前処理では撮影された画像に対して矩形検出を行い、認識処理の対象領域(写真領域)を特定する。認識処理では、AKAZE[2]を用いて写真領域の画像特徴を検出し、登録された画像の画像特徴とのマッチング処理によって、対応するバレットタイム映像を認識する(Step2)。認識処理が成功した場合、バレットタイム映像を呼び出す。認識に失敗した場合は、撮影状態に戻る(Step3)。バレットタイム映像提示は、Kitamuraら[3]のマルチタッチインタフェース閲覧方式によって実行する(Step4)。

3. 書籍画像撮影と図の認識

3.1 撮影と前処理

本研究では、タブレット端末としてMicrosoft社のSurface Pro 4を用いる。Surface Pro4には8メガピクセルの背面カメラと5メガピクセルの正面カメラが搭載されているが、本研究では、背面カメラを用いて1920画素x1080画素の映像を撮影し、認識処理の計算コストを抑えるため0.25倍に縮小した画像を保存する。撮影画像の一例を図2(a)に示す。

前処理として保存画像から矩形領域を検出する。まず、図2(b)に示すように、インターフェース画面上で設定されるターゲットフレーム領域(図2(a)中の破線矩形)が切り出される(1回目の切り出し処理)。

画像2(b)に対して矩形検出処理を適用し、図2(c)に示す検出領域を切り出す(2回目の切り出し)。その際、画像2(b)に対してローカル閾

“Bullet-Time Book: Augmentation of visual information in figures by bullet time video display”

†QIU XINYI, Hidehiko Shishido, Yoshinari Kameda, Itaru Kitahara

‡Graduate School of Systems and Information Engineering, University of Tsukuba

値法による二値化処理を施すことにより、切り出し精度を向上させる。矩形領域の輪郭線形状、および、切り出し領域サイズに基づいて、2回目の切り出しの成否を判断する。2回目切り出しが失敗したと判断された場合は、画像 2(b)を切り出し矩形領域として認識処理に進む。

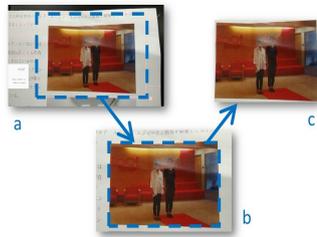


図2 写真領域の切り出し処理

3.2 認識処理

本研究では画像特徴量として検出精度と処理速度を両立しつつ画像のスケール、回転、照明変化にロバストな AKAZE (Accelerated-KAZE) を用いて特徴点検出を行う。

AKAZE の画像特徴点毎に算出される feature descriptor を用いてマッチング処理を行う。撮影した画像とデータベースに登録された画像の間でマッチングすると判断された特徴点の個数を加算し、それが一番多いデータベースを認識結果として出力する。

4. バレットタイム映像の生成提示

被写体の周りに複数台のカメラを配置して多視点映像を撮影する。本方式では、4K 解像度の映像を毎秒 30 枚撮影する。各カメラのカメラパラメータは、Structure from Motion を用いて推定する。本方式では、オープンライブラリの VisualSFM を採用する。

バレットタイム映像生成時には、閲覧者がタブレット端末を用いて注視点（カメラ切り替わり時に中心となる点）を入力する。上述したカメラパラメータを用いて注視点画面中央で観測されるような2次元射影変換を施した映像を、タブレットのストローク操作による提示カメラ切り替え入力に応じて切り替えることにより、バレットタイムが生成提示される。

5. BT-book インターフェース

図 3 に本研究で実装した BT-Book インターフェースを示す。閲覧スタートボタン (Button1)、閲覧終了ボタン (Button2)、メッセージボックス、提示画面 (ウィンドウ全体) で構成される。提示画面中には、写真領域の切り出し処理で利用する四角形フレームを重畳表示する。撮影時

には、ターゲット（書籍の図や写真）がフレーム内に収まるようユーザに指示する。



図3 BT- Book インターフェース

閲覧システムが起動すると、タブレット画面には、図 4 左上に示すような撮影画面が表示される。閲覧スタートのボタンを押すと、撮影、前処理に続いて認識処理が行われる。その後、認識処理の間、待ち合わせ画面が表示される。認識作業が成功した場合、待ち合わせ画面を閉じ、データベースから対応するバレットタイム映像データを読み出す。認識に失敗した場合、撮影画面に戻り、再撮影を促す。

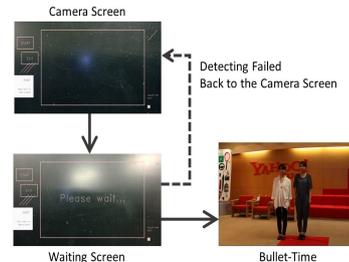


図4 BT-Book インターフェースの流れ

6. おわりに

本研究では、多視点映像提示技術を用いて書籍の視覚情報を拡張する手法を提案した。タッチパネルを用いたバレットタイム映像閲覧インターフェースを実装することで、2次元画像では表現しにくい3次元情報表現が可能で、直感的な視点切り替え操作が可能な閲覧方式を実現した。本研究は科研費 (17H01772) の助成を受けたものである。

参考文献

[1] Mark Billinghurst, Hirokazu Kato, Ivan Poupyrev, "The MagicBook - moving seamlessly between reality and virtuality", IEEE Computer Graphics and Applications, p6-8, 2001
 [2] Nobuyuki Kitamura, Yoshinari Kameda and Itaru Kitahara, "A Method to Switch Multiple-View Videos Using Multi-Touch Interface", Asia-Pacific workshop on Mixed-Reality 2016 (APMR2016), 2 pages, 2016
 [3] Pablo Fernández Alcantarilla, Jesús Nuevo, Adrien Bartoli, "Fast Explicit Diffusion for Accelerated Features in Nonlinear Scale Spaces", BMVC, 2013