

コミック画像のコマ分割処理における制御パラメータに関する検討

石井 大祐[†] 河村 圭[†] 渡辺 裕[†]

[†] 早稲田大学大学院国際情報通信研究科

〒 367-0035 埼玉県本庄市西富田 1011

E-mail: †d-ishii@fuji.waseda.jp

あらまし コミックのコマ分割では高速かつ高精度な処理が必要とされている。我々はこれまで高速性と正確性の両立を目的とした高速高精度コマ分割手法 (FAFSM) の提案を行ってきた。コマ分割では、処理プロセスの低減だけでなく、画像解像度の縮小により処理の高速化を行うことが可能である。しかしながら、画像解像度の縮小率が高くなるにつれてコマの分割精度は低下する。本稿では、実際にコマ分割実験を行い、コマ分割における分割精度と処理時間との相対的な評価を行った。そして、FAFSM における画像解像度変換によるコマ分割への影響に関して検討を行った。キーワード コミック, 画像解析, メタデータ

A Study on Control Parameters of Frame Separation Method for Comic Images

Daisuke ISHII[†], Kei KAWAMURA[†], and Hiroshi WATANABE[†]

[†] Graduate School of Global Information and Telecommunication Studies,
Waseda University, 1011 Nishitomida, Honjo, Saitama, 367-0035 Japan.

E-mail: †d-ishii@fuji.waseda.jp

Abstract Comic frame separation is need to high speed and high accuracy. We have proposed a fast and accurate frame separation method (FAFSM), which purpose is to separate comics with high accuracy and fast processing time. Fast processing in comic frame separation according to not only reduce redundant process but also image shrink. However, frame separation accuracy decreases as image shrink ratio rises. In this paper, we evaluated the relation between accuracy and processing time by frame separation experiment. Moreover, we examined the influence on comic frame separation by image shrink.

Key words Comics, Image analysis, Metadata

1. はじめに

近年、電子書籍市場の発展は顕著である。電子書籍閲覧端末はパーソナルコンピュータをはじめ、電子ペーパー端末や携帯電話などに広がっている。携帯電話向けの電子書籍配信事業はここ数年で非常に大きな伸びを示しており、中でも、電子コミックは作品数、利用者ともに多く、電子書籍の主要コンテンツである。

電子コミックの閲覧にあたり、携帯電話をはじめとする小型の端末では、画面サイズが非常に小さく、紙面全体を表示した際に、テキストの可読性が損なわれる問題がある。実際のコミック配信では、コマと呼ばれる枠の並びによって定義されている紙面内の読み順に沿って画面のスクロールを行うことで、文字の可読性を保ちつつ小型画面端末におけるコミック画像の

閲覧を実現している。

本稿ではスクロール表示を行うための読み順を、ページ上における視線遷移情報と呼ぶ。現在、この視線遷移情報は人手で抽出され、配信データに重畳されている。コミック雑誌をはじめとする多量のページをもつコンテンツの配信を実現するため、視線遷移情報抽出における人的負担を低減する技術が求められている。人手による視線遷移情報の抽出に対して、前処理として、画像を高速に解析し、かつ高い精度でコマを抽出することができれば、修正作業の労力を削減可能である。

これまでに、文書画像を対象とした画像内のレイアウト解析手法 [1-3]、及び任意図形抽出処理 [4, 5] が提案されている。しかしながら、コミック画像は、各オブジェクト間が明確に分離されていない場合や、コマ形状が矩形に限定されていない等の特徴があるため、これらの手法では正しく認識が行えない。

コミック画像に特化した手法として田中らによるコマ分割処理の研究がある。コマの枠線に沿った画素では、コマの枠線に起因するエッジが存在するという特徴がある。そのため、コマの枠線に沿って引かれた直線上では、多くの点で高い濃度こう配値が得られる。この特徴を利用し、田中らは画像上の濃度こう配を基準に2分割を繰り返し適用する手法を提案している [7]。さらに、これを改良し、コミックの画像的特徴を利用することでさらに高精度なコマ分割手法を提案している [8]。本稿ではこの手法を T-method と呼ぶ。

我々はこれまでに、コミック画像に特有の画像的特徴を考慮した手法として、分割線候補の検出にロバスト性を持たせ、分割線上の濃度こう配方向を用いた分割線適合検査を導入した高精度コマ分割手法、およびこれに分割線候補検出順序の最適化、画像解像度の変更及び冗長な分割線候補検出処理の削減を導入した高速高精度コマ分割手法 [10] を提案している。以下、高速高精度コマ分割手法 (Fast and accurate frame separation method) を FAFSM と記す。

しかしながら、これまでに、処理速度及び処理精度を総合した画一的な性能評価は行われていない。本稿では、実際にコマ分割実験を行い、コマ分割の性能と、分割にかかる時間との相対的な評価を行った。さらに、画像解像度変換によるコマ分割への影響に関して検討を行った。

2. コマ分割手法

2.1 T-method

本稿では、画像の2辺を結ぶ、幅1画素の直線を検出線と呼ぶ。T-method では、以下に示す手順でコマ分割処理を実行する。

まず、分割に用いる重み付けされた濃度こう配値を決定する。画像全体に対して縦横に Sobel フィルタを適用し画像上の各点における最大濃度こう配方向及びその値を決定する。次に検出線のパラメータ (ρ, θ) をそれぞれ1刻みで変化させ画像の全探索を行う。ここで、各検出線上の、検出線に対する垂直方向の濃度こう配合計値を求める。この値に対して、画像中心と検出線の距離 ρ に応じてガウス関数により重みづけを行う。以上で得られた値を $Aw(\rho, \theta)$ とする。

次に、コマ間の白画素連続領域長比率 WPR に関する調査を行う。WPR は検出線の長さ、検出線に沿った検出線の近隣における最長の白画素連続長との比率であり、 $WPR = \text{近隣最長白画素連続長} / \text{検出線長}$ となる。

続いて、T型隣接に対する重みづけを行う。T型隣接は、コマ間の白画素領域がT字型となるようなコマの並びを示す。ここで、縦方向、横方向双方の分割線候補の交点を見つける。そして、交点近隣の 40×40 画素において、検出線による濃度こう配の調査を行い、その縦方向、横方向の最大値をそれぞれ T_v および T_h とする。 $T_v > T_h$ の場合には縦方向の分割線候補における TR を $TR = T_h / T_v$ 、 $T_h > T_v$ の場合には横方向の分割線における TR を $TR = T_v / T_h$ とする。これ以外の場合 $TR = 1$ とする。

以上の処理をまとめた式が以下となる。

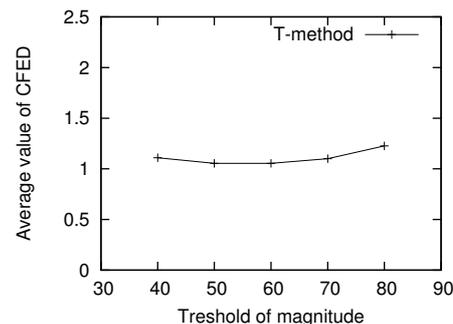


図1 しきい値による編集距離の変化

Fig. 1 Difference of Success rate by threshold value

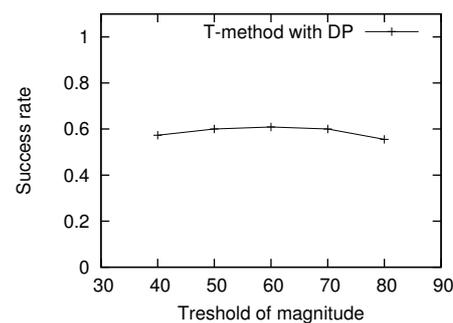


図2 しきい値による成功率の変化

Fig. 2 Difference of Success rate by threshold value

$$Hw(\rho, \theta) = TR \times WPR \times Aw(\rho, \theta) \quad (1)$$

この $Hw(\rho, \theta)$ が分割線候補中で最大となるものについて、その値が h_{th} 以上であればこれを分割線とし、分割処理を行う。そして、全ての検出線上の濃度こう配合計値が小さい状態、つまり、分割対象画像内に分割線が存在しない状態となるまで、以上の処理を再帰的に適用することで、各コマへの分割処理を行う。T-method では、この h_{th} のみを任意のパラメータとして分割処理を制御する。さらに、前処理としてダグラスポーカー法による四角形の塗りつぶし処理を行うことで、コマ内の線による影響低減が試みられている。

ここで、予備実験として実際に h_{th} を変化させコマ分割を行い、3章で述べる評価手法を用いて各 h_{th} における分割精度を評価した。結果を図1, 2にそれぞれ示す。図1では値が小さいほど、図2では値が大きいほど良好な結果を示す。以上より、今回 $h_{th} = 60$ を分割パラメータとして採用した。

3. FAFSM とコマ分割処理の評価法

3.1 FAFSM

我々はこれまでに高精度かつ高速なコマ分割処理を行うことを目的として、FAFSMの提案を行ってきた。FAFSMでは、以下の6つの処理を導入することでコマ分割の高精度化及び高速化を実現している。

- 帯による分割線候補検出処理
- 分割線適合検査
- コマ内外検査
- 分割線候補検出順序の最適化

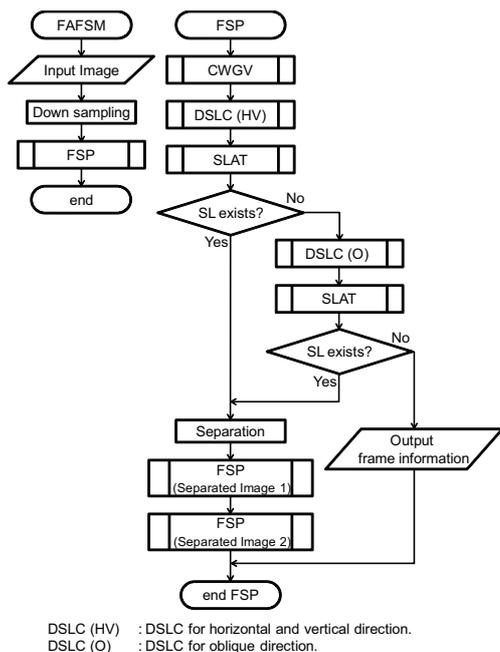


図 3 高速高精度コマ分割手法 (FAFSM) の処理フロー
 Fig. 3 A flow of FAFSM

- 冗長な分割線検索の削減
- 画像解像度の変換

3.1.1 FAFSM の処理手順

FAFSM の処理フローを図 3 に示す。

以下では、分割線候補の検出を DSLC、分割線適合検査を SLAT、適合分割線を SL、および画像を指数とする分割処理のためのプロシーダを FSP とそれぞれ記す。FAFSM では入力された画像に対してまず画像解像度変換により低解像度化を行う。そして低解像度化された画像を FSP に渡す。FSP では、分割線候補検出順序の最適化及び冗長な分割線検索の削減に基づいた手順により、分割線の検出及び画像の分割処理を実行する。

ここで、CWGV は中心に重みづけされた画像の濃度こう配値を求めるプロシーダである。まず画像に対して Sobel フィルタにより各画素における濃度こう配値を求め、さらにガウス関数を用いて画像中心に値の重みづけを行うことで、重みづけされた濃度こう配値を取得する。

3.2 コマ分割処理の評価法

3.2.1 コマ編集距離

本稿では、分割結果として出力される、コマの情報及び視線遷移情報となる順序の精度を評価するために、コマ分割に対して編集距離 [9] の概念を導入する。編集距離はある文字列 A に対して、文字の挿入、削除及び置換の操作を行い、他の文字列 B と同一とするために要する操作の数である。

我々はコミック画像におけるコマを文字と同様にとらえることで編集距離の概念をコマ分割結果に対して適用する。以下ではコマ編集距離を Comic frame editing distance (CFED) と呼ぶ。ここで、コマは画像の一部領域であり、画像上の全てのコマはある順序で並べられるものとする。CFED 取得時における

表 1 計算機環境

CPU	Intel Core2Quad Q6600 2.4GHz
Memory Size	4GB
OS	Linux

表 2 画像サイズ及び画像枚数

	Image Size[pixel×pixel]	Numvers of Pages
Comic-A	740×1200	20
Comic-B	760×1200	30
Comic-C	1024×1536	30
Comic-D	840×1200	30

操作を以下に定義する。

- 「コマの挿入」任意のコマ、あるいは画像上の任意の範囲を新たなコマとして定義し、任意の順番に挿入する
- 「コマの削除」任意のコマを削除する
- 「コマの領域変更」任意のコマの領域を変更する

以下に CFED と操作の例を示す。ここで、あるコミック画像において、A,B,C,D,E という順番で 5 つのコマが並んでいるとする。

例 1 D のコマを過分割し、A,B,C,D1,D2,E という順番でコマが得られた場合。

操作 1:D2 を削除。結果、A,B,C,D1,E。

操作 2:D1 を D に領域変更。結果、A,B,C,D,E。

よって CFED=2 となる。

例 2 C、D のコマを分割しなかったため、A,B,CD,E という順番でコマが得られた場合。

操作 1:CD を C に領域変更。結果 A,B,C,E。

操作 2:C と E の間に D を挿入。結果 A,B,C,D,E。

よって CFED=2 となる。

例 3 C,D,E を誤った形状 F,G に分割した為、A,B,F,G という順番でコマが得られた場合。

操作 1:F を C に領域変更。結果 A,B,C,G。

操作 2:G を E に領域変更。結果 A,B,C,E。

操作 3:C と E の間に D を挿入。結果 A,B,C,D,E。

よって CFED=3 となる。

以上のように定義した CFED を用いることでコマ単位における分割精度の評価を実現する。

3.2.2 成功率

CFED は画像中のコマ数が多いほどその値が大きくなる。コマ数が少ない画像では、ほとんど分割が行えていない場合であっても、その値は小さくなる。そこで、ページ単位での成功率についても同時に評価を行う。ここで、コミック画像 1 ページ内に含まれる全てのコマが規定の順序どおりに全て正しく得られた場合に成功であるとする。上記条件はコマ編集距離が 0 となる場合に満たされる。成功率は全ての対象ページのうちコマ編集距離が 0 となったページの割合である。

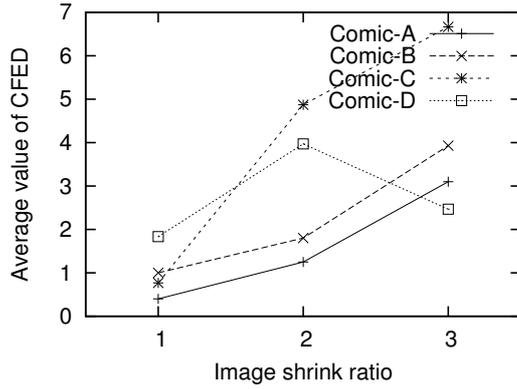


図 4 T-method における画像解像度変換による CFED の変化
Fig. 4 Difference of Levenshtein distances by Image shrinkage in T-method

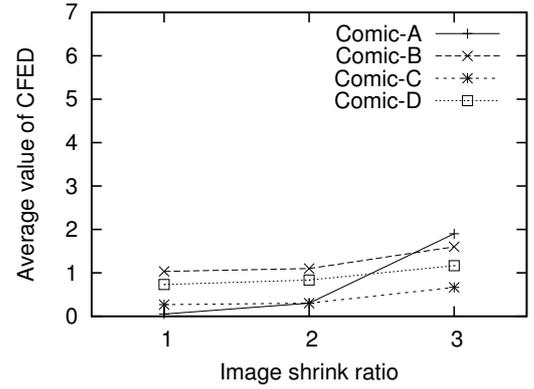


図 7 FAFSM における画像解像度変換による CFED の変化
Fig. 7 Difference of Levenshtein distances by Image shrinkage in FAFSM

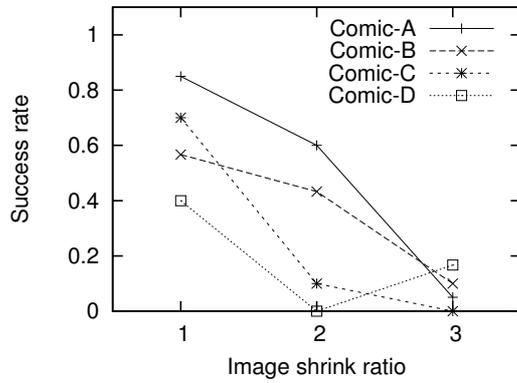


図 5 T-method における画像解像度変換による成功率の変化
Fig. 5 Difference of success rates by Image shrinkage in T-method

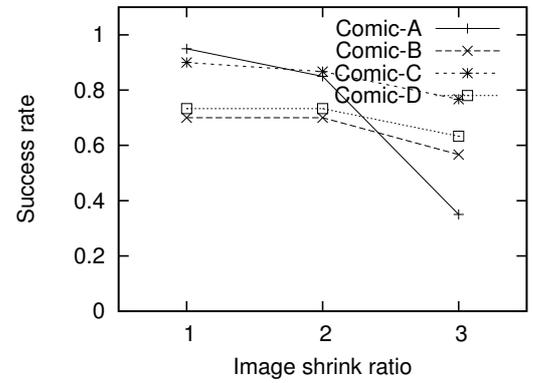


図 8 FAFSM における画像解像度変換による成功率の変化
Fig. 8 Difference of success rate by Image shrinkage in FAFSM

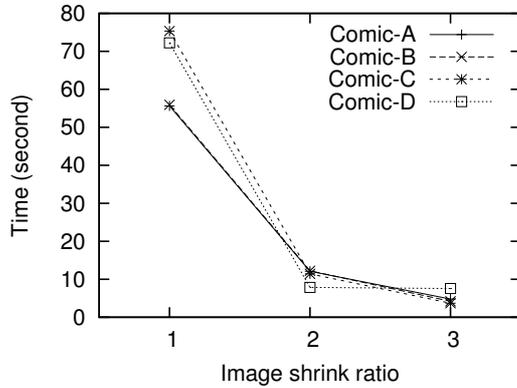


図 6 T-method における画像解像度変換による処理時間の変化
Fig. 6 Difference of processing time by Image shrinkage in T-method

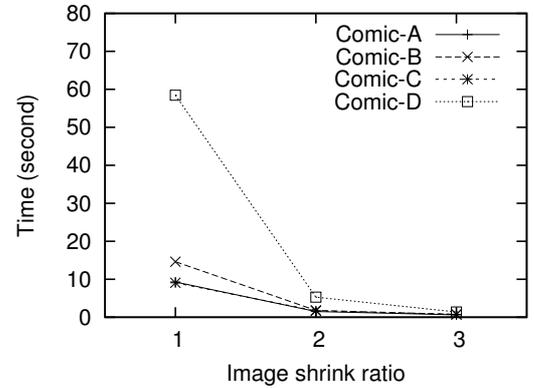


図 9 FAFSM における画像解像度変換による処理時間の変化
Fig. 9 Difference of processing time by Image shrinkage in FAFSM

4. 実験

4.1 実験条件

実験には以下に示す 4 作品合計 110 枚の画像を用いた。実験に用いた計算機環境を表 1, 入力画像のサイズおよびページ数を表 2 にそれぞれ示す。

分割性能の評価では, コマの CFED 及びページ単位成功率を用いた。また, 処理速度は画像の入出力, 解像度変換, 前処理等も含めた総合時間を直接計測した。

本実験では FAFSM のパラメータは [10] よりそれぞれ $A = 20, B = -2, E = 10, \rho = 20, \delta = 40, n = 7, m = 2$ を用いる。また, 今回 FAFSM に対してしきい値 M_{th} を設け, その線上の平均輝度勾配値が, 設定したしきい値以上となるもののみ分割線候補として検出するようにした。 M_{th} の値は実験により最適な値である 60 とした。 FAFSM の画像解像度変換には $R \times R$ サイズの平均値フィルタを用いてダウンサンプリングを行う。

T-method では高速化のアプローチがとられていない。そこ

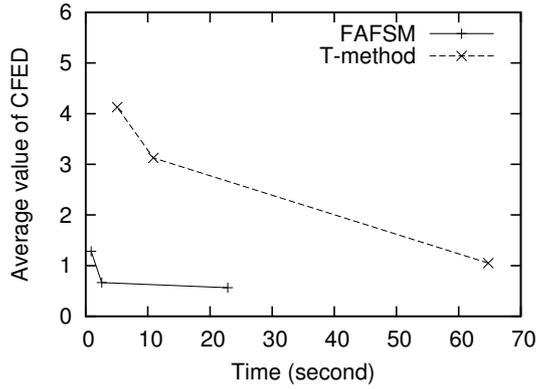


図 10 各手法における処理時間と CFED

Fig. 10 CFED with processing time in each methods

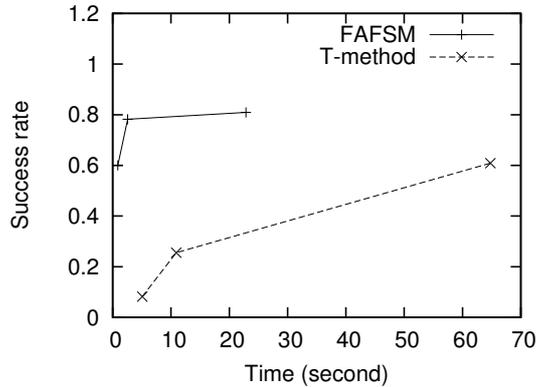


図 11 各手法における処理時間と成功率

Fig. 11 Success rate with processing time in each methods

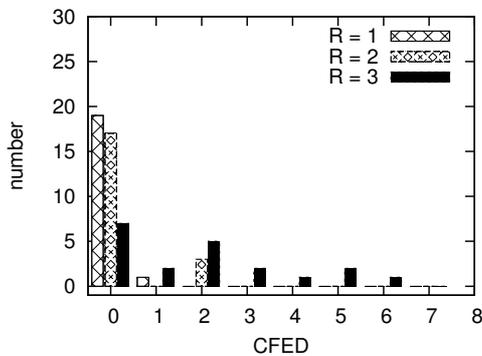


図 12 FAFSM における画像解像度変換による CFED の変化 Comic-A

Fig. 12 Difference of CFED by Image shrinkage in FAFSM Comic-A

で FAFSM と同様の処理により画像解像度を縮小することで、実質的な処理時間を削減した。ただし、 $R = 1$ の場合には画像解像度変換は行われず、通常の T-method と同一となる。

4.2 最適分割パラメータに関する評価

4.3 コマ分割速度および分割精度に関する評価

高速な分割を行うために、画像解像度を縮小する。ここでは、画像解像度変換が分割精度に与える影響を調査し、各手法における分割精度と分割時間の関係を示す。

T-method 及び、FAFSM を用いて、各シーケンスにおける

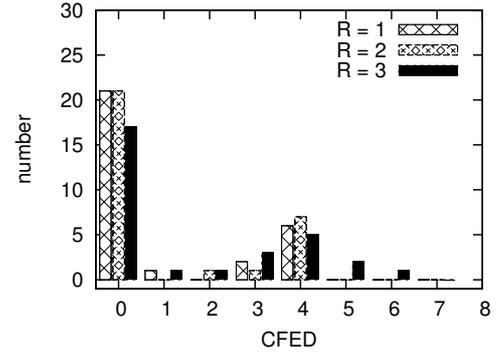


図 13 FAFSM における画像解像度変換による CFED の変化 Comic-B

Fig. 13 Difference of CFED by Image shrinkage in FAFSM Comic-B

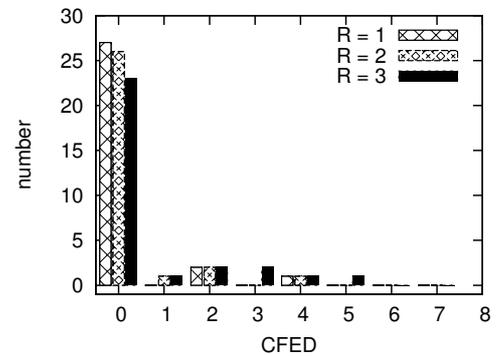


図 14 FAFSM における画像解像度変換による CFED の変化 Comic-C

Fig. 14 Difference of CFED by Image shrinkage in FAFSM Comic-C

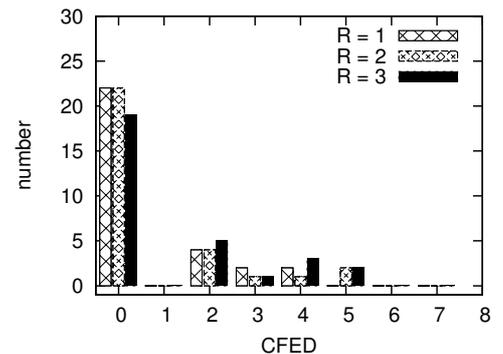


図 15 FAFSM における画像解像度変換による CFED の変化 Comic-D

Fig. 15 Difference of CFED by Image shrinkage in FAFSM Comic-D

画像解像度変換の影響について調査を行った。

T-method を用いた場合の CFED および成功率の変化を図 4,5 に示す。T-method では、 $R=2$ の時点で Comic-B, Comic-C の CFED が急激に、その他のシーケンスではやや緩やかに値が上昇している。その後 $R=3$ の時点では Comic-C はそのまま上昇を続け、逆に Comic-B では値が回復している。その他のシーケンスはこの時点で値が急激に上昇する結果となった。

成功率も同様に Comic-B は急激な低下の後回復し, Comic-C は $R=2$ の時点で急激に低下し, その後 0 となった. その他のシーケンスは $R=3$ の時点で成功率が大幅に低下した.

全てのシーケンスにおいて $R=2$ の時点から分割精度が低下した原因として, T-method では, 検出線に幅 1 画素の線を用いるため, 検出線とコマの枠線に若干のずれがある場合には検出が行えない事があげられる. ここで, 検出線のパラメータである ρ 及び θ の変化刻み幅が一定のまま, 画像解像度を小さくしたために, 検出線とコマの枠線が一致しない割合が高くなり, 分割線の検出率が急激に低下したと考えられる. また, Comic-B の場合には, $R=2$ ではまったく一致のとれなかったコマの枠線が, $R=3$ の場合に上手く検出線と一致のとれる位置に来たため, CFED の値が上昇したと考えられる.

FAFSM を用いた場合の CFED および成功率の変化を図 7,8 に示す. FAFSM では R が大きくなるにつれて CFED は上昇し, 成功率は低下することが確認された. CFED について, 特に Comic-A では R が 2 から 3 になる時点で他のシーケンスと比較して急激に値が上昇している. 同様に, R が 2 から 3 になる時点で Comic-A の成功率が急激に低下した. この原因は Comic-A では他のシーケンスに比べてコマ同士の間隔が狭く, 画像解像度を下げるに従って, すき間が埋まってしまふコマが多く存在したためと考えられる.

両手法おける分割に要した平均処理時間を図 6,9 にそれぞれ示す. どちらの手法においても $R=1$ から 2 の間では処理時間が大幅に削減された. 一方 $R=3$ の時点ではあまり処理時間は削減出来ていない. FAFSM では R が 1 から 2 に変化する際の Comic-D の処理時間削減率が特に顕著である. これは Comic-D の画像解像度が他のシーケンスと比較して高いため, 低解像度化による画素数の減少量が大きいことが原因である.

次に, 各手法にを用いて分割を行い, 分割時間とその際の平均編集距離および成功率を調査した. 結果を図 10 及び図 11 にそれぞれ示す. 図 10 では左下が最も高速かつ高精度であり, 右上に近づくにつれて低速かつ低精度となる. 図 11 では左上が最も高速かつ高精度であり, 右下に近づくにつれて低速かつ低精度となる.

T-method では, 処理時間を短くするに従って CFED は大幅に上昇し, 成功率は大幅に低下している. これに対して FAFSM では 5 秒未満の高速処理でも平均 1 未満の CFED, 及び 8 割程度の成功率を維持している. この原因として, FAFSM では帯による分割線検出を行っているため, 画像解像度を小さくした場合においても, 分割線を検出することができるのに対し, T-method では分割線の検出率が低下しやすいことが原因であると考えられる. グラフより, 今回調査を行った全ての処理時間において FAFSM が T-method よりも CFED 及び成功率が高い結果となった.

加えて, FAFSM における各シーケンスにおける画像解像度ごとの CFED の分布に関して調査を行った. 結果を図 12-15 にそれぞれ示す. 全てのシーケンスに共通して, R が 3 となった時点でより大きな CFED の数が増加する特徴がみられる. これは, 全く分割が行われないシーケンスが増加するためである.

この原因として, コマの境界自体が埋まってしまう場合がある事, 及び分割線とのマッチングがうまく行えない可能性が高くなる事が挙げられる.

Comic-B ではすべての R において CFED=3, 4 の数が多い結果となった. この原因は, Comic-B は他のシーケンスと比較して, コマからのみ出し, 四角形でないコマの数, コマ内の直線成分等が非常に多く, これが原因となって解像度変換を行わない場合であっても, 未分割もしくは過分割が多く発生したためである.

全てのシーケンスを通じて, $R=2$ では CFED の分布は $R=1$ と比較してあまり変化していない. これは, 分割結果をコマ単位でみた場合においても, ほぼ同等の分割結果が得られていることを示している. よって, コマ分割を行う際に $R=2$ としても分割性能に与える影響は小さいといえる.

以上の実験より, FAFSM において $R=2$ の場合であれば, $R=1$ の場合とほぼ同等の分割精度を保ちつつ, より高速な分割処理が行えることが確認できた.

5. おわりに

本稿では, 高速かつ高精度なコマ分割を行うことを目的として, 実験により各手法の分割処理時間と分割精度の関係を明らかにした. さらに, FAFSM 手法が分割速度及び分割精度双方で優れていることを明らかにした. また, FAFSM における画像解像度変換の各シーケンスに対する影響を明らかにし, $R=2$ 程度までは得られる結果がほぼ同等であることを示した. 結果として, FAFSM において $R=2$ とした場合に高い分割精度を保ちつつ高速な分割処理が行えることを確認した.

文 献

- [1] 辻 三郎, “スプリット検出法による文書画像構造解析,” 信学論 (D-II), vol.J74-D-II, no.4, pp.491-499, April 1991.
- [2] K.Y. Wong, R.G. Casey, and F.M. Wahl, “Document Analysis Systems,” IBM J. Res. Dev., vol.26, no.6, pp.647-656, November. 1982.
- [3] K. Kise, A. Sato, and M. Iwata, “Segmentation of page images using the area Voronoi diagram,” Computer Vision and Image Understanding, vol.70, no.3, pp.370-382, 1998.
- [4] D.H. Ballard, “Generalizing the Hough Transform to Detect Arbitrary Shapes,” Pattern Recognition, vol.13, no.2, pp.111-122, 1981.
- [5] 木村彰男, 渡辺考志, “高速一般化ハフ変換-相似変換不変な任意図形検出法,” 信学論 (D-II), vol.J81-D-II, no.4, pp.726-734, April 1998.
- [6] 山田雅之, 鈴木茂樹, ラフマツト ブディアルト, 遠藤 守, 宮崎慎也, “携帯電話を利用したコミックの閲覧システムとその評価,” 芸術科学学会論文誌, vol.3, no.2, pp.149-158, June 2004.
- [7] 田中孝昌, 東海林健二, 外山 史, “マンガ画像のコマ割り構造解析,” 2006 信学総大, D-12-89, March 2006.
- [8] T. Tanaka, K. Shoji, F. Toyama, and J. Miyamichi, “Layout Analysis of Tree-Structured Scene Frames in Comic Images,” 20th International Joint Conference on Artificial Intelligence (IJCAI- 07), pp.2885-2890, Hyderabad, India, Jan. 6 - 12, 2007.
- [9] LEVENSHTEIN, V. “Binary Codes Capable of Correcting Deletions, Insertions and Reversals,” Soviet Physics Doklady 10, 8, 707-710, 1966.
- [10] 石井大祐, 河村 圭, 渡辺 裕, “コミック画像におけるコマ分割処理の高速化に関する検討,” 情処全大, 1P-2, March 2007.