2016 年 2 月 4 日

平成 27 年度　情報・通信工学科　コンピュータサイエンスコース　卒業研究発表会

OpenCV-GPU モジュールを用いた数字専用 OCR

情報・通信工学科　学籍番号:1211015　成見研究室　井出拓弥

1　はじめに

OCR(Optical Character Recogniton/光学文字認識) とは、手書きの文字や印刷された文字を読み取り、データと照らし合わせて文字を判断しコンピュータに認識可能な文字コードに変換する装置である。昨今、ビジネスの場では多くの注文書や請求書などの書類が扱われている。それらの管理を効果的に行うため、書類を電子テキスト化する OCR 技術が用いられている。しかし、画像処理は重い処理のため相応の時間がかかる。そこで業務の効率化を図るため、OpenCV-GPU モジュールを用いて、数字専用の高速な OCR システムを開発する。

2　OpenCV と GPGPU

2.1　OpenCV

OpenCV とは、Intel によって開発された画像認識に関する機能のライブラリである。正式名称を”Intel Open Sorce Computer Vision Library”と言い、オープンソースの C/C++, java, Python のライブラリ集で、コンピュータビジョンに必要な各種機能がパッケージされている。OpenCV を用いることによって、物体の認識、パターン認識、動きの認識等の高度な画像処理機能を容易に扱うことができる。

2.2　GPGPU

GPGPU(General Purpose Computing on GPU) はグラフィック描写に用いられる GPU（Graphics Processing Unit）の高い計算性能を通常の計算にも利用することを目的とした技術である。GPU はグラフィックス処理の高速化のために演算器の並列度を向上させているという特徴がある。この特徴が GPU 並列処理に強い計算ハードウェアとして注目される要因となった。GPU にも得意不得意があり、全ての計算が高速化されるわけではない。一般的に CPU は連続的な計算が得意であり、GPU は並列的な計算が得意とされている。そのため、並列化するメリットがないようなプログラムでは GPU の性能は発揮されない。

3　既存研究

3.1　GPU による高速画像処理 [1]

GPGPU 技術に注目し、画像処理アルゴリズムを高速化した例。この研究は、CUDA を用いて基本的な画像処理アルゴリズムの実装を行った研究である。GPU 上にテンプレートマッチングを実装し、高速化も行っている。また、回転やスケールに対して頑健な特徴である SIFT を GPU 上で実装している。

本研究では OpenCV の GPU モジュールを用いている点で異なっている。OpenCV の API を用いることで CUDA プログラミングの記述を行わずに GPGPU を扱うことができる。

3.2　S-OCR[2]

数字の読み取りに限定した OCR アルゴリズムを考案した研究である。隠れモルコフモデル、ニューラルネットワークといった複雑なアルゴリズムを用いずにシンプルなアルゴリズムのみで OCR を実現している。「数字の背景にある白の画素数」、「数字の重心の x 座標と y 座標の比」、「数字のカーブを描く輪郭の長さ」、これらのパラメータを用いることでフォント形式 Arial Unicode MS の数字の認識率は 96.16%を達成した。本研究はテンプレート画像を追加することでどのフォントにも対応できるという点で異なる。

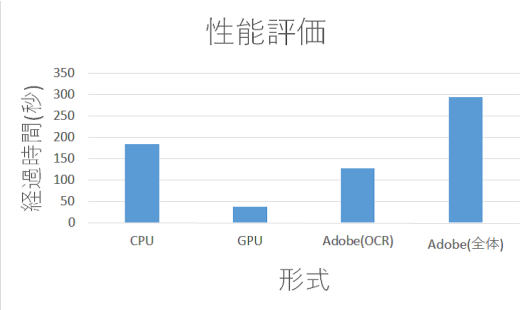

図 1: 性能評価

4　システム概要

0〜9 のそれぞれの数字に対応する画像とのテンプレートマッチングを行うことで数字専用の OCR を実現する。matchTemplate は画像内から発見したテンプレートの座標を返すので 0-9 それぞれの座標をひとつの配列に格納して画像内の順になるようにソートすることで認識結果を得る。matchTemplate にはＣＰＵ版と GPU 版の 2 種類がある。パフォーマンスの比較を行うために、それぞれ用いたシステムを作成する。

4.1　実験

4.1.1　実験方法

実際に OCR を業務に活用すると想定して、OCR を終えるまでにかかる実時間の計測を行う。本システムの実用性を測るため、既存の OCR ソフトと本システム双方を用いて OCR 実行時間の比較を行う。既存の OCR ソフトには Adobe acrobat DC Pro[3] を用いる。本システムの GPU を用いたものと Adobe acrobat DC Pro を用いて 100 ページの pdf ファイルを対象に OCR を行い、実時間を計測する。出力先はどちらもテキストファイルである。

4.2　実験結果

実験結果は図 1 のようになった。

4.3　評価・考察

本システムは既存製品である Adobe の約 3.4 倍の速度を示した。CPU のみの場合は Adobe の約 0.7 倍の速度を示している。本システムは 100 ファイルの OCR に約 36 秒、既存製品である Adobe は約 126 秒の時間を要した。その差は約 1 分半である。OCR するファイル数が膨大になってくると要する時間も多くなる。OCR するべきファイルが 10000 ファイルあったとすると、処理を終えるまでにかかる時間の差は約 2 時間半に及ぶ。2 時間半短縮可能であると考えると、十分業務の効率化につながると言える。

5　今後の課題

認識に用いた matchTemplate は画像の回転やゆがみといった変化に弱く、認識精度に不安が残るものである。フォントが変化するだけで認識できなくなってしまうため、多くのフォントを認識しようとするとテンプレート画像が膨大な数になってしまう。数字専用にすることで高速化を図っているが、テンプレートが膨大になりすぎれば速度は低下する。また、数字に限定したことにより数字以外の情報も必要とする場面では役に立たないというデメリットもある。今後の課題は特徴抽出、標準パターンなどを用いて認識精度を向上させ、対応できる文字の種類を増やすことである。

参考文献

[1] 出口 大輔, 井手 一郎, 村瀬 洋, ”GPU による高速画像処理”, 名古屋大学大学院情報科学研究科

[2] Jenil Shah, Viral Gokani, ”A Simple and Effective Optical Character Recognition System for Digits Recognition using the Pixel-Contour Features and Mathematical Parameters”, Department of Electronics & Communication, DharmsinhDesai university

[3] Adobe Acrobat Pro DC, https://acrobat.adobe.com/jp/ja/products/acrobat-pro.html