

GAによる機械加工メモの自動要約に関する予備検討

Preliminary Study on Automatic Summarization of Machining Technical Documents with GA

立林 裕太郎* 藤田 充洋**
Yutaro Tatebayashi* Mitsuhiro Fujita**
*豊田工業大学
*Toyota Technological Institute

古谷 克司* 佐々木 裕*
Katsushi Furutani* Yutaka Sasaki*
**(株)豊田中央研究所
**Toyota Central R&D Labs., Inc.

1 研究背景

遺伝的アルゴリズム(GA)は大規模な解空間において強力な最適化手法の一つであることが示されている[1,2]. しかし, GAは解の全探索が計算量的に困難である場合に近似解が得られる点で有用であるが, GAを適用するだけで要約問題が解けるようになるわけではない. 著者らが知る限り, GAを使った日本語文書要約に関して良好な結果は報告されていない. そこで, 本研究では日本語単一文書を対象にして, GAの有用性に関する予備検討を行う. 具体的には, 参照要約をフィットネス関数に利用したGAにより, どの程度の品質の要約が得られるのかを確認する.

2 関連研究

自動要約とGAを組み合わせた研究報告はいくつかある. 例えば, Sillaら[3]の場合, テキストの自動要約タスクを通じて, 属性を分類するクラスタリング手法の性能を向上させるため, GAに基づく属性選択の有効性を調べている. Yehら[4]は, 学習機能を持った要約システムを新たに提案した. 具体的には, 文書の特徴をいくつか種類に分けた上で, コーパスベースのテキスト要約を改善するために次の2つの手法を提案している. 1点目は, 文章の位置が異なる文位置の意味を強調するためにランク付けを行っていること, 2点目は, スコア関数を算出する際にGAを使うことで特徴に関する適切な重みの組み合わせを見つけ出すことである.

上に述べたように, 英語を対象としたGAの研究は多い. しかし, 日本語を対象とした自動要約の研究においては, 小倉ら[5]による複数文書要約を対象にした要約手法が提案されているが, 自動要約は最適化問題であるにも関わらず, GAにより良好な要約結果は得られていない.

3 単一文書要約におけるGAの有用性の調査

調査の対象にする要約手法は, 文書の前処理, 重要文抽出, 文短縮の3つから構成される. ここでは, それぞれに

ついて 3.1~3.3 節で説明する. その後, 3.4 節において, 要約における GA の有用性の調査方法について示す.

3.1 前処理

本研究の対象とする機械加工技術文書には, 冗長な文書が多く存在する. このような文書から, 的確な重要文を抽出するのは難しい. そこで本研究では, 要約の精度の向上を目的に文書に対する下処理を行う. 機械加工技術文書には「箇条書き」と呼ばれる項目を並べた表現が多く存在するが, 箇条書きの部分においては重要な内容を含む可能性が非常に低い. そこでSVMを用いて, 行毎に箇条書きか否かという判別を行う. 箇条書きと判別した文については, 3.2節において要約の対象外とする.

(1) 素性の割り当て

箇条書きの特徴を取り出すために, 文毎に素性を抽出する. 表1に用いた素性の一覧を示す.

表1 用いた素性

素性	次元	内容
文の長さ	80	文Sの文字数
単語ID	1900	文中に出現した単語のID
文の位置	30	文書からみた文の位置番号
句読点	1	文中に句読点が存在するか
漢字の数	50	文に含まれる漢字の数
文を占める漢字の割合	10	文に含まれる漢字の割合
カタカナの数	50	文に含まれるカタカナの数
文を占めるカタカナの割合	10	文に含まれるカタカナの割合
ひらがなの数	50	文に含まれる単語の数
文を占めるひらがなの割合	10	文に含まれる単語の割合
半角文字数	50	文に含まれる英語や数字の数
文を占める半角英数字の割合	10	文に含まれる英語や数字の割合
文頭の文字	5	文頭の文字の種類

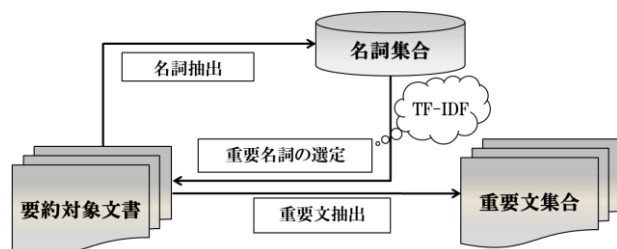


図1 重要文抽出フェーズの概要

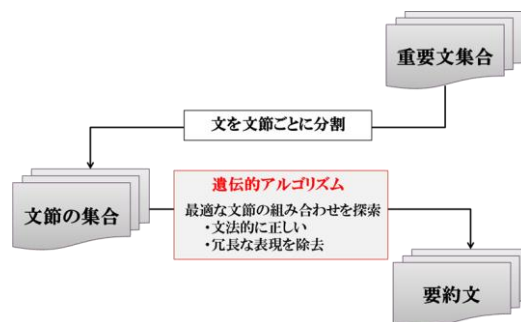


図2 文短縮フェーズの概要

(2) 機械学習による判別

SVMを用いて、文集合に対して一文毎に「箇条書きに類する表現なのか」についての判定を行う。

(3) 箇条書きの削除

前項 (2) による判別において、「箇条書きに類する表現」と判別した部分について、重要文抽出の対象外とするため、除去を行う。

3.2 重要文抽出

重要文抽出とは、入力された原文に対して、ある基準を用いて評価し、より重要だと判断した文を抽出して要約を出力する方法である。重要文を選択する評価尺度には、文の出現位置や節見出しなどといった様々なものが利用されているが、本研究では、文書の関心の中心となる事柄について言及する単語の出現頻度や重要度などを利用して、文の重要度を算出する手法を採用する。単語の重要度の算出には、TF-IDF法を用いる。

図1に重要文抽出フェーズの概要を示し、以下に対応する処理の流れについて示す。

(1) 名詞の抽出

TF-IDFスコアの算出のため、対象である機械加工技術文書に対し、形態素解析を行った後、品詞の同定を行う。次に、品詞の同定結果から名詞であった単語のみを抽出する。なお、形態素解析を行う際、機械加工の専門用語を追加したユーザ辞書を用いる。

(2) TF-IDFスコア算出

TF-IDF法を用いて、対象の文書毎に、文書に含まれる全ての名詞群に対して、TF-IDFスコアを算出する。

(3) 重要文判定及び抽出

各文書の任意の文中に出現する単語のTF-IDFスコアの総和を文の重要度として計算し、文書それぞれについて、文の重要度が高かった上位2文を重要文集合として抽出する。この際、ユーザ辞書に含まれている「専門用語」については、通常のTF-IDFスコアに重みを与えて計算する。

3.3 文短縮

文短縮は、一文毎に重要でない箇所を削り、主要な情報を減らすことなく、テキストを短く表現し直す要約手法である。この手法は、段落や文を対象とする重要文抽出とは違い、句や文字列を単位とした重要箇所抽出である。本研究では、GAを用いて文短縮を行う。図2に文短縮の概要を示し、以下に処理の流れについて示す。

(1) 前処理

3.2節の文書の重要文抽出にて生成した重要文集合の全ての文に対して、形態素解析、係り受け解析を行った。係り受け解析の結果から、文節毎に分割する。

(2) 最適解の探索

各文節を遺伝子として、GAを用いて冗長な修飾表現の除去を行う。

(3) 要約文生成

前項の探索で得られた結果を用いて、要約文を生成する。

3.4 参照要約を用いた要約の有用性の調査

本研究では、参照要約を利用しながら、3.1節から3.3節までの手順を行い、GAの有用性を調査する。GAのフィットネス関数については、参照要約に対する「ROUGE-1」「ROUGE-2」「文中に含まれる名詞のTF-IDFの総和(TF-IDF)」の3通りを用いる。

4 評価実験

4.1 対象データ

本研究において、原文書及び参照要約については、機械加工の現場で作成された報告書とその要約文（ヘッドライン）を用いた。データセットの概要を表2に示す。

4.2 前処理評価

箇条書きの分類のため、原文書全てに対して人手によるアノテーションを行った。そのうち、箇条書きの数は769であった。学習には交差推定を行い、 $k=8$ とした。また表3に、「箇条書き」の半別で算出したPrecision, Recall, F-scoreの値を示した。表3に示した通り、高い精度が得られた。

4.3 要約の質に関するスコアリング評価

表4に、前処理を行った96文書（要約対象文）について、ROUGE-1, ROUGE-2, TF-IDFをフィットネス関数と設定したGAによって生成された要約について、ROUGE-Nのスコアによる評価結果を示した。比較のため、同条件でSVMを用いた要約手法（SVM）[6]の評価結果についても示す。TF-IDFの重みは5倍と設定し、GAのパラメータは表5の通りに設定した。

表2 原文書と参照要約の概要

データ種類	文字数	行数
原文書	30,938	1,032
参照要約	4,570	191

表3 箇条書きに関する分類の精度

評価指標	recall	precision	F-score
評価値	0.926	0.939	0.932

表4 4つの手法によるROUGE-Nの評価の比較

手法	ROUGE-1	ROUGE-2	ROUGE-3	ROUGE-4
ROUGE-1	0.571	0.374	0.286	0.222
ROUGE-2	0.552	0.376	0.276	0.212
TF-IDF	0.442	0.274	0.200	0.147
SVM	0.486	0.300	0.211	0.160

表5 GAのパラメータ

	集団数	交叉	突然変異	要約率
実験条件	60個	80%	2%	70%

重要文抽出

きさげ加工とは、金属面の目には見えないほどの凹凸を平坦にならし、より平らな面を作り上げる技術です。機械では実現できない平面度を實現し、高精度の加工を可能にします。

生成要約文 (ROUGE-1)

きさげ加工とは、金属面のほどの平坦にならし、より作り上げる技術です。機械では実現できない平面度を實現し、高精度の

生成要約文 (ROUGE-2)

金属面の目には見えない凹凸をならし、より平らな面を作り上げる機械では平面度を高精度の加工を可能にします。

生成要約文 (SVM)

きさげ加工とは、金属面の見えない凹凸を平坦にならし、面を作り上げる技術です。実現できない平面度を高精度の可能にします。

生成要約文 (TF-IDF)

きさげ加工とは、目には見えないほどの凹凸を平坦にならし、平らな面を作り上げる技術です。機械では実現できない平面度を加工を可能にします。

正解要約文

きさげ加工とは、機械を超えた面粗度を實現する加工方法である。

図3 本要約システムによる処理結果の例

また、本要約システムを用いて生成された要約の比較例を図3に示す。表5より、(1)他の手法に比べてROUGEをフィットネス関数としたGAで要約を生成した時が高いROUGEスコアを得られていること、(2)TF-IDFをフィットネス関数としたGAの手法はSVMを用いて文短縮をした手法より評価結果が低いこと、が分かった。

4.4 要約の質に関する人手評価

4.3節の結果を踏まえ、提案手法及びSVMにおいて、生成した4つの要約文について可読性と適切性の二つの指標について評価値を割り振る主観評価を行った。評価者に与えた教示は表6, 表7の通りであり、評価指標を用いて、評価した結果を表8に示した。

表6~8から、(1)人手評価と自動評価による結果の優劣が一致しないこと、(2)内容の可読性についてはTF-IDFの和をフィットネス関数としたGAの手法が最も高評価であること、(3)内容の適切性についてはSVMを用いた手法が最も高評価であること、が分かった。

5 考察

評価結果から、自動評価ではROUGEをフィットネス関数としたGAによる手法が高評価だったが、他の手法と比

較して要約として優れているとはいえない結果となった。これは、日本語において、ROUGEと人間の評価結果との相関は0.9程度であるとする平尾ら[7]の結果とは異なるものとなった。また可読性においてはSVMによる文短縮より、本論文の手法の方が高い結果を得られたが、適切性においては従来手法を下回る結果となった。その原因として、SVMの素性には文の構造情報が含まれていたのに対し、GAによる手法では、その情報が一切利用されなかったからだと考えられる。

6 結論

本研究ではGAの有用性を検討することを目的として、参照要約を用いるフィットネス関数を設定し、GAを使った要約手法を提案した。その結果、分かった事は次の2点である。

表6 可読性の評価指標

評価値	評価尺度
評価 4	要約文に含まれる文節のうち、0%-25%の割合で文節を挿入、または別の単語に置換することで、流暢に読むことができる。
評価 3	要約文に含まれる文節のうち、25%-50%の割合で文節を挿入、または別の用語に置換することで、流暢に読むことができる。
評価 2	要約文に含まれる文節のうち、50%-75%の割合で文節を挿入、または別の用語に置換することで、流暢に読むことができる。
評価 1	要約文に含まれる文節のうち、75%-100%の割合で文節を挿入、または置換しなければ流暢に読むことができない。

表7 適切性の評価指標

評価値	評価尺度
評価 4	要約文に必要な内容が、75%以上含まれている。
評価 3	要約文に必要な内容が、50%-75%の割合で含まれている。
評価 2	要約文に必要な内容が、25%-50%の割合でしか含まれない。
評価 1	要約文に必要な内容が、25%未満しか含まれていない。

表8 GAによる人手評価の結果の比較

評価指標	GA			SVM
	ROUGE-1	ROUGE-2	TF-IDF	
可読性評価	2.70	2.94	2.57	3.12
適切性評価	2.74	2.59	2.83	2.81

まず、表5においてTF-IDFをフィットネス関数とした結果よりもSVMによる手法が上回ったように、日本語で書かれた機械加工文書において参照要約を使ったとしてもGAを用いた要約手法はSVMによる手法より劣るということ、次に、表5と表8の結果を比較すると高評価だった手法がそれぞれ異なったようにROUGEによるスコア評価と人手評価の結果は必ずしも一致しないこと、である。

今後の課題としては、構文情報を考慮したフィットネス関数の提案が挙げられる。SVMでは文の構造情報を含んでいたのに対し、本研究のGAによる手法では、単語のスコアリングしか行っていない。GAの有用性をより詳細に確認するため、構文に関する特徴もフィットネス関数に含めていく必要がある。

謝辞

本論文の執筆、及び、実験に貴重な助言を頂きました豊田工大三輪誠准教授に心より感謝します。

参考文献

- [1] J. H. Holland, Genetic algorithms, Scientific American, vol. 267, pp. 66-72, 1992.
- [2] L. G. Caldas and L. K. Norford, A design optimization tool based on a genetic algorithm, Automation in Construction, Vol. 11, No. 2, pp. 173-184, 2002.
- [3] C. N. Silla Jr. *et al.*, Automatic text summarization with genetic algorithm-based attribute selection, in Proceedings of the 9th Ibero-American Conference on Advances in Artificial Intelligence, pp. 305-314, Springer, 2004.
- [4] J.-Y. Yeh *et al.*, Text summarization using a trainable summarizer and latent semantic analysis, Information Processing & Management, Vol. 41, pp. 75-95, 2005.
- [5] 小倉由佳里, 小林一郎. 多目的GAを用いた複数文書要約への取り組み. 人工知能学会全国大会論文集, 28, pp.1-3. 2014.
- [6] 立林裕太郎, 藤田充洋, 古谷克司, 佐々木裕, 機械加工技術文書の自動要約, 言語処理学会 第20回年次大会 発表論文集, pp.638-641, 2014
- [7] 平尾努, 奥村学, 磯崎秀樹: 拡張ストリングカーネルを用いた要約システム自動評価法, 情報処理学会論文誌, Vol. 47, No. 6, pp. 1753-1766, 2006.