事実・意見・知識に焦点を当てた観点に基づく複数文書要約

関　洋平 [1], 江口　浩二 [2,1], 神門　典子 [2,1]
総合研究大学院大学情報学専攻 [1]
国立情報学研究所 [2]
seki@grad.nii.ac.jp eguchi@nii.ac.jp kando@nii.ac.jp

本研究では，要約は，要約を受け取る利用者の情報要求・検索意図に応じて異なるタイプのものが求められると考える。ここでいう要約のタイプとは，(1) 事実をまとめたもの，(2) 意見性に焦点を当てたもの (3) 知識を重視したものの 3 つを設定する。本研究では，トピックと要約タイプの組み合わせを観点と呼び，複数文書を対象とした要約を観点に応じて自動作成することを目的としている。本稿では，その具体的な手法として，文書ジャンルを 4 つの特性の集合として定義を行い，特性の違いを反映して要約のタイプの出力を切り替えるシステムについて報告を行う。ジャンル特性は SVM を利用して学習した結果に基づいて与えられる。特定のジャンル特性の文書だけから複数文書要約を構成する場合に，すべての文書から複数文書要約を構成する場合との結果の違いについて実験を行った。実験データは，前回の NTCIR-3 TSC のテストコレクション 30 トピックと，自分で作成した解説記事や社説を多く含む 22 トピックのテストコレクションを対象として，同じ実験を行った。

Multi-Viewpoint Multi-Document Summarization Focused on Facts Opinion and Knowledge

Yohei Seki[1], Koji Eguchi[2,1], and Noriko Kando[2,1]
Department of Informatics, The Graduate University for Advanced Studies (Sokendai)[1]
National Institute of Informatics (NII)[2]
Tokyo 101-8430, Japan
seki@grad.nii.ac.jp eguchi@nii.ac.jp kando@nii.ac.jp

We assume that summaries vary according to the users and factors such as their information needs, situations and purposes in seeking information, so that different summaries can be produced according to the author's point-of-view. In our summarization system, a list of topics discussed in the given document set are presented, so that the user can select a topic of interest from the list as well as the summary type, such as fact-reporting, opinion-oriented or knowledge-centered. The purpose of this study is to build a multi-document summarizer to produce summaries according to viewpoints. We assume a relationship between summary type and genre because both features are orthogonal to topics. As an exploratory stage of investigation, we examined the effectiveness of genre for source documents to produce different types of summaries. We also analyzed the results of the multi-document summarization using automatic genre classification to reveal the association between genre dimensions and the summary types.

1 はじめに

テキスト自動要約の研究では，要約の結果が作成者により異なることが以前から指摘されている（[11, 13]）。これは作成者によって要約する際に注目する内容や観点が異なるからである。

Query-biased Summarization (SUMMAC[1], 1998) は，要約を作成する元となる記事で論じられているいくつかの話題（トピック）の中で Query（問い合せ）の内容に関連するトピックに焦点を当てる要約作成であるが、Query と要約に含める元記事の内容との関連の強さは、主としてトピックという側面について検討されてきた。しかしながら，観点の違いはトピックだけではなく，どういうタイプの情報を求めるかという側面にも関連している。

要約の評価会議である DUC2003 (Document Understanding Conference)[2] のタスク 3 では，観点に基づく要約 (viewpoint summary) が取り上げられている。この中で，Angheluta らの研究（[2, 1]）は，トピック分割技術と観点を組み合わせることを試みているが，そのことの有効性は明示的に示すことはできていない。また，文章論でも，嶋田 [14] は，各段落からまんべんなく要約の対象を取る大意型要約と，ある文章の特定の内容に注目してまとめる要旨的要約を区別している。観点に基づく要約や要旨型要約は，2.1 節で後述するように，トピック以外の観点の違いも包含しうる概念であるが，トピック以外の注目すべき観点については明確には提案されていない。

複数文書要約では，元文書集合に対して想定できる観点の種類が単一文書要約よりも多く，観点に応じた要約作成の研究がますます重要になっている。

本研究は，要約作成における「観点」をトピックと要約タイプの組み合わせと定義し，複数文書を対象とした観点に応じた要約を自動作成することを目的としている。トピックは，文書クラスタリングの手法を応用し，元記事集合中のパッセージを単位としたクラスタリングを行い，各クラスタの中心トピックを識別することにより，当該記事集合中の複数のトピックを自動的に認定する。要約タイプは，新聞記事の要約については，利用者が求める情報のタイプに応じて，「事実報告型」「意見重視型」「知識重視型」という分類を設定した。本提案に基づいて作成したシステムのインターフェースを図 1 に示す。

また，観点に応じた複数文書要約を生成するために，ジャンルの情報を利用することについて調査した。本研究では，‘ジャンル’ とは「日記」や「報告書」のような文書ジャンルの意味で使用し，主題分類の意味では用いない。ジャンルの定義の一例としては，“類似したスタイルで書かれた文書のグループで，主題分類とは直行する” [5] とする定義があるが，本研究でもこの解釈に従う。従来の要約研究では，特に事実の情報とトピックに焦点が当てられてきたが，ユーザは意見，評価，見通しのような主観的な情報を要求することもある。また，意見を重視した要約や多観点の質問応答 [4] の研究が注目されている。

本論文では，Biber の多次元分析 [3, chap. 6.4] を利用して ‘ジャンル’ を定義する。

本論文は 7 節から構成される。次節では，本研究で利用する要約データとジャンルの次元について説明する。3 節と 4 節では，自動ジャンル分類と複数文書要約への応用について説明する。5 節では，要約タイプを考慮して結果を分析し，6 節では要約タイプを区別してテストコレクションを作成し，評価を行う。最後に，7 節で結論を示す。

2 多観点に基づく複数文書要約で利用するジャンル特性

本節では，多観点に基づく要約を実現するために，複数文書要約において利用する “ジャンル特性” を中心に説明する。。

2.1 要約データ

本実験では，NTCIR-3 TSC2 [12, 10, 6][3] の日本語要約データを使用した。より新しい NTCIR-4 TSC3 データ を利用しなかった理由は，TSC3 に比べると TSC2 の方がさまざまな文書ジャンルの記事を含んでいるトピックが多いことによる (TSC3 は報道記事中心)。NTCIR-3 TSC2 フォー

[1] TIPSTER Text Summarization Evaluation Conference

[2] http://www-nlpir.nist.gov/projects/duc/pubs.html

[3] http://research.nii.ac.jp/ntcir/

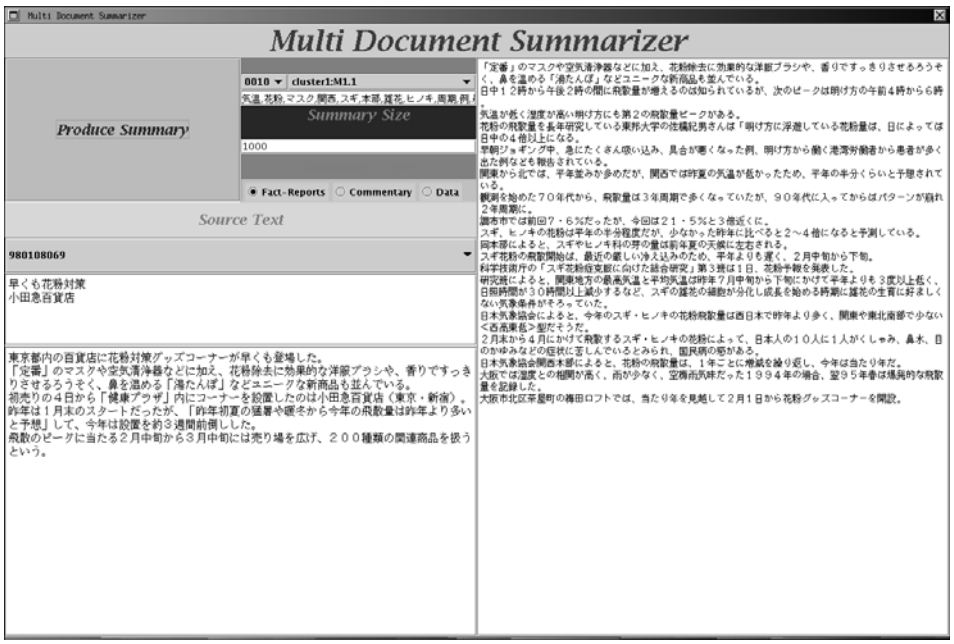

図 1: 観点を指定した複数文書要約システム

マルラン では 30 のトピックについて，日本語の新聞記事集合が指定されている。新聞記事は毎日新聞の 1998 年と 1999 年版を利用している。3 人の要約作成者によって人手作成要約が作成されている。TSC2 要約については，文書集合のほかにトピックが与えられているが，要約のタイプに関する指示は与えられていない。

2.2 文書集合のタイプ

NTCIR-3 TSC2 では，文書集合は，“報道記事のみ” か “報道記事以外を含む” の 2 つのタイプに分類されている。30 トピックに対して，21 トピックが前者に，9 トピックが後者に分類されている。

2.3 ジャンル特性

既存のニュース配信記事へのジャンル分類としては，IPTC (International Press Telecommunications Council) が設定したジャンル特性[4]がある。しかし，このジャンル特性は 40 を超える種類があり，意見 (Opinion) や背景 (Background) のようなものから，Music や Raw Sound のような取り扱う題材，Press Release のようなニュースソースに関する情報など，さまざまな情報を一次元の分類として取り扱っており，複雑で付与が困難であり，要約の効果を検証するには不適切である。

そこで，本研究では Douglas Biber の分類 [3, pp.135-155] に基づき，ジャンル特性を複数の基本的特性の組み合わせで表現することを提案する。利点を以下にあげる。

- 少数のクラスで記述できるため効率的
- 基本特性それぞれにおける要約への効果が明確になる点
- 新たなジャンルを追加する際にも全体の枠組みを変えることなく追加できる
- 付与の規則が単純化できる

Biber のジャンル特性の基本特性は 5 つである。

1. Elaborated/Situation-Dependent Reference
2. Overt Expression of Argumentation
3. Impersonal/Non-Impersonal Style.
4. Narrative/Non-Narrative Discourse
5. Involved/Information Production

Biber の次元のうち 5 つ目の次元は，新聞ではすべて情報産出 (Information Productoin) に当たると考えられるため，区別することができない。残りの 4 つを参考にして以下の 4 つの次元を定義した。

[4] http://www.iptc.org/metadata/mtdta_ts-table01.php

1. 状況即時性/詳細描写性
   状況に即して，見た状況を即時的にありのまま描写している記事は，状況即時性を持つ記事として判定する。後から振りかえり考察や分析など詳細な解説に記事が基づいている場合には，詳細描写性として判定する。
2. 議論あり/議論なし
   ある主張の根拠となる事実や引用が含まれており，説得や予測等の目標に向けて論旨の展開がなされている場合には議論性があると判定する。論旨に関係なく，事実に対する根拠があるだけの記事は議論なしと判定する。
3. 非人称スタイル/人称スタイル
   主語の不在，受動態などを手がかりとして，第3者から見た見地で書かれている場合に，非人称スタイルと判定する。著者が自分の意見を述べている場合は，受動態表現があっても，人称スタイルと判定する。
4. 事実報告性/意見性
   記事の文章構成を考慮したうえで，もっとも主張したい内容が事実報告か意見かに基づいて判定する。

本研究では，それぞれの文書に対する“ジャンル特性”を以上4つの次元で定義する。

3 ジャンル分類

本節では，ジャンル特性について，人間が付与する場合と自動分類で使用する際の一貫性について説明する。

3.1 ジャンル特性の人手付与

ジャンル特性の4つの次元を評価するために，以下の手順で実験を行った。

i) 本論文の第一著者がTSC2の19トピックにおける156の文書に対して4つのジャンル特性の次元について(1/0)の二値を割り振った。次に，付与のためのインストラクションを作成した。

ii) 3人の被験者(大学院生)にいくつかの文書にインストラクションに基づいて値を付与してもらい，インストラクションについてインタビューした。インタビューに基づき，付与のための曖昧さをなくすようにインストラクションを修正した。

iii) 別の被験者二名に156の文書を修正したインストラクションに基づいて値を付与してもらった。

表1: カッパ係数: 被験者間の付与の一致度

ジャンル特性の次元	被験者の組			
	(a1,a2)	(a1,a3)	(a2,a3)	平均
状況即時性	0.597	0.607	0.587	0.627
議論性	0.685	0.440	0.394	0.506
非人称性	0.729	0.585	0.447	0.587
事実報告性	0.606	0.724	0.484	0.605

iv) 修正したインストラクションに基づいて，第一著者の付与の一貫性について調べて修正した。

これらの手続きを行った後，被験者2名と第一著者の合計3名による156文書の付与の一致について，κ 係数を調べた。結果を表1に示す。結果は，人間の被験者の付与がやや信頼できることを示している。

3.2 自動ジャンル分類

TSC2の複数文書要約テストコレクションは，30トピックから構成される。ジャンル特性の次元について，要約タイプごとに要約に含まれる文抽出の精度が向上するかどうかを理解するために，“ジャンル特性”を文抽出パラメタとして利用することによる重回帰分析の決定係数の向上率が有意であるかどうかについて調査した[15]。この実験では4つ以下の文書数から構成されるトピックの複数文書要約については向上の効果を検証できなかった。また，ジャンル特性の次元を文書間で区別できないトピックも存在した。このようなトピックを除くことで，30トピックを二つのグループに分けた。すなわち，向上効果が検証できる17トピック(148記事)と検証できない13トピック(69記事)に分けた。

この文書集合に対して，テキスト分類で利用される機械学習技術であるサポートベクトルマシン(SVM)[9]を利用してジャンル特性の自動付与を実現した。ここでは，13トピックに含まれる69記事を訓練集合として，17トピックに含まれる148記事をテスト集合として使用した。ただし，これでは訓練集合の数が少ないことから，107の別の記事集合について，ジャンル特性の4つの次元を手動で付与し，訓練集合に追加した。また，ジャンル分類を実現するために，以

下の特徴素を定義した。本研究では，この特徴素のことを “機能的特徴素” と呼ぶことにする。

- 5 つの記事特徴素: 署名，紙面，写真，図，ニュースソース。
- 9 つの統計的特徴素 : 文字数，タイプ/トークン比，文の数，意見文の数，見通し文の数，背景文の数，接続詞の数，引用括弧の数，平均の文の長さ。
- 60 の機能語 (意見，見通し，背景の情報に関係している)。
- 93 の記号 (いくつかのパンクチュエーションを含む)。

これらの選択された特徴素の合計は， 167 個 である。この特徴素はジャンル分類に有効であった。ジャンル分類の結果を以下の表 2 に示す。この特徴素を，通常のテキスト分類で使用される内容語を特徴素とした場合と比較した。内容語の特徴素は 13,825 の内容語を含んでおり，本研究では，この特徴素のことを “内容語特徴素” と呼ぶことにする。

表 2: ジャンル分類の正確さ

特徴素	次元			
	状況	議論性	非人称性	事実性
機能的	79.1	86.5	90.5	88.5
内容語	63.5	85.8	93.2	89.2

状況 = 状況即時性

一見，内容語の特徴素が非人称性や事実報告性のジャンルの次元に効果的であるようにも思われるが，これは，表 3 に示すように，テストデータの偏りによるもので，内容語の特徴素は次元の 1/0 を区別できていないことがわかる。

表 3: ジャンル分類の結果

ジャンル Dimension	特徴素	システム/人間			
		p/p	p/n	n/p	n/n
状況	機能的	105	28	3	12
状況	内容語	65	11	43	29
議論性	機能的	6	5	15	122
議論性	内容語	0	0	21	127
非人称性	機能的	129	5	9	5
非人称性	内容語	138	10	0	0
事実性	機能的	127	12	5	4
事実性	内容語	132	16	0	0

p/p : システムと人間が 1 と分類。
p/n : システムは 1 と分類したが，人間は 0 と分類。
n/p : システムは 0 と分類したが，人間は 1 と分類。
n/p : システムと人間が 0 と分類。

表 4: 複数文書要約システムに対するカバレッジと精度

	参照要約					
	S1		S2		S3	
	cov.	pre.	cov.	pre.	cov.	pre.
全トピック	0.346	0.388	0.228	0.283	0.246	0.336
複ジャンル	0.257	0.317	0.209	0.265	0.178	0.257
単ジャンル	0.463	0.481	0.253	0.307	0.335	0.440

cov. = カバレッジ
pre. = 精度
複ジャンル = 17 のジャンルの区別が可能なトピック
単ジャンル = 13 のジャンルが区別されないトピック

4 ジャンル分類を利用した複数文書要約

本研究では，すでに開発していた複数文書要約システムを，ジャンル分類を利用する事で拡張した。評価は，TSC2 複数文書要約の集合に対して，“カバレッジ” と “精度” を NTCIR-4 TSC で配布された評価ツールを利用して計算することで行った。 (“カバレッジ” と “精度” の定義は [7] 参照; 正解文のデータは [8] で使用されたものと同じデータを使用した; 自動評価ツールは NTCIR-4 TSC3 [5] のオーガナイザに提供していただいた)。ジャンル分類なしで要約を生成すると，要約に対するカバレッジと精度は表 4 に示した結果となった。TSC2 のテスト集合は 3 つの人手作成要約がある; したがって，それぞれの要約に対して 3 つの結果が得られる。30 のトピックのうち，17 のトピックについてはそれぞれ 5 つ以上の文書が含まれており，区別できる値が割り振られているジャンル特性の次元が存在する。残りの 13 トピックについては， 10 トピックは 4 つ以下の文書から構成されており，3 トピックについては区別できるジャンル特性の次元を持たない。

これらの結果はジャンル分類なしの複数文書要約は，13 トピックに対しては効果的であるが，17 トピックに対しては効果的でないことを示している。17 トピックに対して，ジャンル分類を行い，4 つのジャンル次元の 1/0 の値の文書集合だけから要約を構成して，同じ評価を行った。結果を表 5 に示す。

表 5 は，ジャンルの次元がカバレッジと精度を向上させることを示している。しかし，この

[5] http://www.lr.pi.titech.ac.jp/tsc/tsc3-en.html

表 5: ジャンル分類を伴うカバレッジと精度の向上

ジャンル	参照要約					
	S1		S2		S3	
	cov.	pre.	cov.	pre.	cov.	pre.
base	0.257	0.317	0.209	0.265	0.178	0.257
G1-p	0.236	0.292	0.195	0.248	**0.184**	0.269
G1-n	0.066	0.205	0.055	0.185	0.049	0.173
G2-p	0.063	0.197	0.067	0.181	0.064	0.211
G2-n	0.243	0.311	0.207	0.262	**0.190**	0.277
G3-p	**0.261**	0.323	**0.212**	0.269	**0.186**	0.254
G3-n	0.016	0.148	0.037	0.153	0.049	0.194
G4-p	0.243	0.307	**0.216**	0.270	0.173	0.246
G4-n	0.024	0.062	0.019	0.044	0.017	0.052

G1 = 状況即時性　G2 = 議論性
G3 = 非人称性　G4 = 事実性

向上に有意差はない。そこで，次節では，要約タイプを区別して，この結果をより詳しく調査する。

5 要約タイプ

前節では，全体の要約の向上について議論した。本研究では，主題分類に対応した要求と直交する情報要求の軸として，以下の 3 つの要約タイプを定義する。

1. 事実報告型要約：主に，発生したイベントなど事実を中心に作成した要約。トピックに関連した事実ついてどんな報道がされているか知りたいユーザが望むような要約。
2. 意見重視型要約：著者の主張・識者の意見などの情報を重視し，そのトピックに関わるアドバイスや見通し・評価を得たいユーザが望むような要約。
3. 知識重視型要約：イベントよりも，すでに知識としてあるような定常的な事実を重視し，ユーザが要約を読むだけで特定の知識が得られるような要約。

TSC2 では，要約作成者は要約タイプに対するインストラクションを与えられていなかった。したがって，17 トピックに対する TSC2 の短い要約[6]について，人手で要約タイプを 3 つのタイプに分類した。知識重視型要約については，この分類の際には数値データを重視して分類した。

[6]TSC2 では長い要約と短い要約の二種類の要約を生成している。

表 6: TSC2 の短い要約に対する要約タイプの数

要約タイプ	要約作成者		
	C1	C2	C3
事実報告	6	9	11
意見重視	1	3	5
知識重視	10	5	1

C1-C3: 要約作成者は参照要約 S1-S3 を書いた人。

表 7: ジャンル特性の次元の利用によるカバレッジの向上

要約タイプ	ジャンル	Captionists		
		C1	C2	C3
事実報告	ベース	0.246	0.202	0.074
	事実性	0.199	**0.225**	0.074
意見重視	ベース	0.066	0.147	0.073
	議論性	0	0.037	**0.089**
知識重視	ベース	0.209	0.135	0.036
	状況	0.190	0.126	0.036
	事実性	0.202	**0.138**	0.036

5.1 要約作成者による傾向

参照要約は 3 人の要約作成者によって作成された。最初の作成者は “知識重視型” 要約を作成する傾向があるのに対して，3 番目の作成者は “意見重視型” 要約を作成する傾向がある。要約タイプの数を表 6 に示す。

5.2 要約タイプについての分析

ジャンル特性の各次元を利用した場合の要約タイプごとのカバレッジを表 7 に示す。

要約タイプの効果に対する結果から以下のことが言える。

- 事実報告型要約については，事実性のジャンルは要約作成者 2 によって作成された要約に対して効果的である。
- 意見重視型要約については，議論性のジャンルは要約作成者 3 によって作成された要約に対して効果的である。
- 知識重視型要約については，事実性のジャンルは要約作成者 2 によって作成された要約に対して効果的である。

これらの結果は，要約作成者 2 と 3 は，元文書のジャンルの次元を手がかりに，要約を作成する文書を選択していることを示唆している。

6 要約タイプを区別した複数文書要約の作成

NTCIR-3 TSC データの分析に基づき，3 種類の要約タイプを設定した。この要約タイプに基づき，同一の要約作成者に対して同じ文書集合から「事実報告型」「意見重視型」「知識重視型」の 3 つの観点から要約を作成してもらうかたちで，多観点に基づく複数文書要約のためのテストコレクションを作成した。

6.1 テストコレクションの作成

テストコレクションの作成は，以下の手続きで行った。

1. 1998 年〜1999 年に話題になったトピックに関連するキーワードを 22 個選ぶ。
2. 上記二年分の毎日新聞・読売新聞の記事集合を対象としてキーワードによる検索を行い，上位 60 件の記事を選択。
3. 60 件の記事から，文書ジャンル (報道記事・インタビュー記事・社説・解説記事・特集記事など) が分かれるように，平均 9〜10 記事程度を選択。
4. 22 トピックの文書集合から事実報告型・意見重視型・知識重視型要約を作成してもらった。トピックごとの 3 つの要約作成は同一人物が行った。トピック全体では 3 名の作成者により要約を作成した。
5. 要約文字数は 800 文字で統一した。

作成したテストコレクションの内容を表 8 に示す。NTCIR-TSC2, TSC3 の複数文書要約テストコレクションと比べると，以下の特徴がある。

1. 要約の元記事数は TSC2 より TSC3 に近く，一記事平均 9〜10 件程度。
2. 要約の圧縮率は TSC3 がおよそ 5 % なのに比べると少し小さめ。
3. 記事の文書ジャンルは TSC2 よりもさらに多くのジャンルを含む。

表 8: 要約タイプを区別したテストコレクション

課題		記事集合			
ID	トピック	文字合計	記事数	毎日	読売
S010	欧州通貨統合	20530	10	5	5
S020	年金支払い抑制	21704	10	6	4
S030	粉飾決算	21207	9	4	5
S040	イトマン事件	20647	10	5	5
S050	ペイオフ解禁	19251	11	6	5
S060	次世代デジタル携帯	20353	11	5	6
S070	ガイドライン関連法	20687	9	2	7
S080	コソボ	20583	11	5	6
S090	戦略兵器削減	15499	8	4	4
S100	脳死判定	21052	7	3	4
S110	少年審判	20967	11	6	5
S120	情報公開法	16953	8	3	5
S130	ドナーカード	15902	10	6	4
S140	確定拠出型年金	19131	12	4	8
S150	遺伝子組み換え食品	20225	12	7	5
S160	組織犯罪対策法	21425	8	2	6
S170	臨界事故	16935	7	5	2
S180	金融ビッグバン	19411	8	3	5
S190	プルサーマル	19092	9	3	6
S200	戦域ミサイル防衛	17323	8	3	5
S210	中国国有企業	13529	6	4	2
S220	北アイルランド紛争	14241	10	4	6

6.2 テストコレクションによる評価

上記作成したテストコレクションに対して，抽出文の数を [7] で定義されている通り，可能な解答の組み合わせのうちの最小文数として定義を行い，抽出文を単位として自動評価ツールによりカバレッジと精度を評価した。評価結果を表 9 に示す。

表 9: 22 トピックに対する要約タイプ別カバレッジと精度

要約タイプ	カバレッジ	精度
事実報告	0.099	0.130
意見重視	0.072	0.087
知識重視	0.085	0.102

圧縮率が高いことと複数ジャンルの入力文書を含んでいるためか，カバレッジの値は NTCIR3-TSC または NTCIR4-TSC のテストコレクションを使用した場合に比べてかなり悪い。全体的な傾向として，事実報告型要約がカバレッジが最も高く，知識重視型要約が続き，意見重視型要約がカバレッジが最も低いことは観察される。

要約タイプ毎のジャンル分類を利用することによるカバレッジの向上の効果については，実験を進めた上で報告したいと考えている。

7 おわりに

本論文では，特定の観点を与える手がかりとして，文書ジャンルが持つ効果について実験を行った。 3 人の作成者によって付与されたジャンル特性の次元に対して，κ 係数はこれらの次元に対する一定の信頼性を示した。ジャンルの次元ごとの特性値は，機能語を利用した手がかり句を特徴素として自動的に分類され，分類の正確さは 80 % 前後の評価値を得た。

また，ジャンル分類の結果を複数文書要約に応用した。アプローチは単純であり，特定の次元の特性値が 1/0 を示す文書だけから要約を構成している。この場合，該当しない文書中の文は要約に決して含まれないため，カバレッジの全体的な向上はむずかしい。それにもかかわらず，要約タイプと作成者によって向上するいくつかの事例を確認した。

謝辞

この研究の一部は文部科学省科学研究費補助金特定領域研究「情報学」 (課題番号 13224087) の助成を受けて遂行された。また，研究の一部は栢森情報科学振興財団の助成を受けて遂行された。

NTCIR-3 TSC2 と NTCIR-4 TSC3 の開催ならびにデータの利用に当たってご尽力された方々に対して感謝します。

参考文献

[1] R. Angheluta, R. De Busser, and M-F. Moens. The use of topic segmentation for automatic summarization. In *Workshop on Text Summarization (DUC 2002) in conj. with the Association for Computational Linguistics 40th Ann. Meeting (ACL 2002)*, Philadelphia, Pennsylvania, July 2002.

[2] R. Angheluta, M. F. Moens, and R. De Busser. K. u. leuven summarization system - duc 2003. In *Workshop on Text Summarization (DUC 2003) in conj. with the 2003 Human Language Technology Conference (HLT/NAACL 2003)*, Edmonton, Canada, June 2003.

[3] D. Biber, S. Conrad, and R. Reppen. *Corpus Linguistics - Investigating Language Structure and Use -*. Cambridge Approaches to Linguistics. Cambridge University Press, 1998.

[4] C. Cardie, J. Wiebe, T. Wilson, and D. Litman. Combining low-level and summary representations of opinions for multi-perspective question answering. In *AAAI Spring Symposium on New Directions in Question Answering*, pages 20–27, 2003.

[5] A. Finn, N. Kushmerick, and B. Smyth. Genre classification and domain transfer for information filtering. In F. Crestani, M. Girolami, and C. J. van Rijsbergen, editors, *Advances in Information Retrieval, 24th BCS-IRSG European Colloquium on IR Research, Proceedings (Lecture Notes in Computer Science)*, volume 2291, pages 353–362. Springer-Verlag, Glasgow, UK, March 2002.

[6] T. Fukusima, M. Okumura, and H. Nanba. Text summarization Challenge 2: Text summarization evaluation at NTCIR Workshop3. In *(Oyama, Ishida, & Kando 2003)*, 2003.

[7] T. Hirao, M. Okumura, T. Fukusima, and H. Nanba. Text summarization Challenge 3: Text summarization evaluation at NTCIR Workshop4. In *Proceedings of the Fourth NTCIR Workshop on Research in Information Access Technologies: Information Retrieval, Question Answering, and Summarization*, 2004.

[8] T. Hirao, K. Takeuchi, H. Isozaki, Y. Sasaki, and E. Maeda. SVM-Based Multi-Document Summarization Integrating Sentence Extraction with Bunsetsu Elimination. *IEICE Transactions on Information and Systems*, E86-D(9):1702–1709, 2003.

[9] T. Joachims. *Learning to Classify Text Using Support Vector Machines : Methods, Theory, and Algorithms*. Kluwer Academic Publishers, 2002.

[10] N. Kando. Overview of the Third NTCIR Workshop. In *(Oyama, Ishida, & Kando 2003)*, 2003.

[11] J. Kupiec, J. O. Pedersen, and F. Chen. A trainable document summarizer. In E. A. Fox, P. Ingwersen, and R. Fidel, editors, *SIGIR'95, Proceedings of the 18th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval. Seattle, Washington, USA, July 9-13, 1995*, pages 68–73. ACM Press, 1995.

[12] K. Oyama, E. Ishida, and N. Kando, editors. *Proceedings of the Third NTCIR Workshop on Research in Information Retrieval, Automatic Text Summarization and Question Answering*. National Institute of Informatics, 2003.

[13] K. Sparck Jones. Automatic summarizaing: Factors and directions. In I. Mani and M. T. Maybury, editors, *Advances in Automatic Text Summarization*, pages 1–14. MIT Press, 1999.

[14] 嶋田 保. 要約文・要旨. In 森岡健二・永野賢・宮地裕・市川孝, editor, **文章作法**, volume 2 of **作文講座**, pages 273–286. 明治書院, 1967.

[15] 関 洋平, 江口 浩二, and 神門 典子. ファセット化ジャンルに基づく複数文書要約の分析. In **電子情報通信学会言語理解とコミュニケーション研究会**, volume NLC2003-23, pages 49–54, August 2003.