

ブログ記事からのイベント文抽出によるシーンの生成

Scene Generation System using Event Describing Sentences from Weblogs

佐藤 圭太*1 西原 陽子*2 砂山 渡*1
Keita SATO Yoko NISHIHARA Wataru SUNAYAMA

*1 広島市立大学大学院情報科学研究科

Graduate School of Information Sciences, Hiroshima City University

*2 大阪大学大学院基礎工学研究科

Graduate School of Engineering Science, Osaka University

As many people write weblogs about events of a day and their opinions, they have dispatched much information positively. It is rare to read weblogs that have a lot of sentences and boring contents. However, we can attract attention of many people by adding images to the weblogs, and information sharing is also advanced. This paper proposes a scene generation system using by pictures that describe events in weblogs. We confirmed that the proposed system generates scenes where events are described from experimental results.

1. 序論

近年、個人の情報発信の手段として、ブログが注目されている。ブログの作成機能を提供するサービスやソフトウェアがあり、携帯電話、PDA を使って更新できるため、PCに詳しくない人でも簡単に作成、更新が行える。多くの人がブログに個人の体験や日記、意見を書くようになり、積極的に情報発信するようになった。しかし、テキストばかりの文章量の多いブログ記事は内容が興味深くないと、どんなに有益な情報が含まれていても人に読まれる機会は少ない。

本研究では、「喫茶店でケーキを食べた」のような出来事が書かれてあるブログ記事を対象に、イベント文を抽出し、イベント文からシーンを作る単語を抽出しシーンを画像として生成するシステムを提案する。テキストのみのブログ記事にシーンを加えることによって Web 上でのコミュニケーションの推進を図る。

2. 関連研究

小作らは新聞記事コーパスからイベント情報を抽出するシステムを提案している[小作 04]。周期的に出現する単語を抽出することでイベント情報を抽出し、抽出した単語からユーザーに適切なものを提示し、観光地に不慣れたユーザーの Web 検索を支援している。本研究との違いは小作らは「催し物」、本研究では「出来事」をイベントとする点になる。

野呂らは文章からイベント文を抽出し、イベントが起こった時間帯を判定するシステムを構築した[野呂 05]。このシステムは文章から動詞の過去形で終了する文をイベント文として抽出し、抽出した文から「朝」「昼」「夜」の時間帯を連想させる単語を抽出し、イベントの起こった時間帯を判定する。本研究では抽出したイベント文からシーンの生成を行う点で異なっている。

角らは文章を絵に変換し理解支援を行うシステムを提案した[角 05]。角らのシステムでは、文章中の登場人物や背景画像をあらかじめ用意するため、他の文章を即座に変換はできない。本研究では、画像検索で取得した画像からシーンを生成し、イベントを表すより多くの文章を変換する。

携帯端末を用いて、ユーザ同士のコミュニケーションを支援する研究[伊藤 06]では、各ユーザが撮影した写真とその上への書き込みをネットワーク上でリアルタイムに共有することで、知人同士のインフォーマルコミュニケーションや講習会などでの時

空間の共有を支援している。本研究では、時間によらず多くの人とのコミュニケーションを支援する点で異なっている。

また、携帯端末を用いて、母国語が違うユーザ同士のコミュニケーションを支援する研究があり[山田 02]、携帯端末を用いて伝えたい言葉をアイコンと母国語で書かれた文字を見せ伝えるシステムがある。本研究では、シーンの生成によって、文字を使わずコミュニケーションする点で異なる。

3. ブログ記事からのシーン生成システム

図 3.1 に提案システムの構成を示す。提案システムではブログ記事を入力とし、出来事を表す文であるイベント文を抽出する。そして、イベント文に含まれる単語の役割を特定し、「ブログ作者のいる場所を表す単語」、「ブログ作者の動作を表す単語」、「ブログ作者の動作の対象となっている対象物を表す単語」を抽出する。一つの場所について記述されたイベント文の集合であるシーンに切り分け、各シーンを画像として生成し出力する。

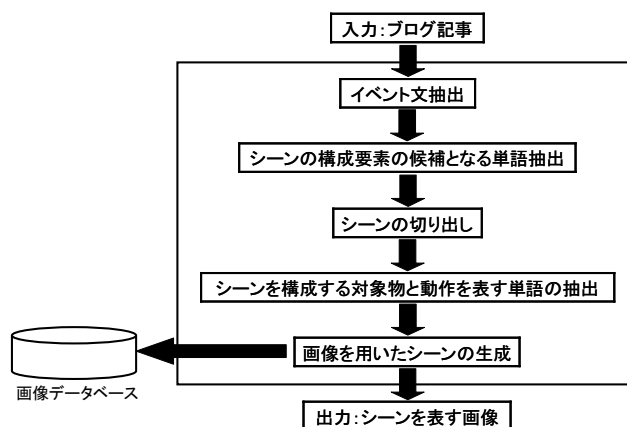


図 3.1: 提案システム構成

3.1 入力

提案システムへの入力、店に関する記述を含むブログ記事の URL とする。動作を表す画像を用意するときに、全ての画像を網羅できないため、店に関する記述を含むブログとする。

3.2 イベント文抽出

ブログの記事から、動詞の過去形で終了する文をイベント文として抽出する。入力されたブログ記事の一文ごとに「南瓜」

[Kudo 03]を用いて形態素解析を行い、文末に助動詞「た」があるか判別する。イベント文を抽出した後、各イベント文で、動詞の連用形があればその直前で、接続助詞があればその直前でイベント文を分割し、複文であるイベント文を単文にする。

3.3 シーンの構成要素となる単語の抽出

イベント文を入力とし、各文から「ブログ作者のいる場所を表す単語」、「ブログ作者の動作を表す単語」、「ブログ作者の動作の対象となっている対象物を表す単語」を抽出する。抽出されたイベント文に含まれる動詞を「動作を表す単語」とする。

抽出した動詞にかかる名詞で、名詞+格助詞となっているものを抽出し、取得した名詞の役割を後続の格助詞から特定する[渡邊 04]。役割を特定した名詞のうち、場所の役割を持つ単語、対象物の役割を持つ単語を抽出する。提案システムでは、7つの名詞の役割を扱う。役割が一つしかないものはその役割を与え、役割が複数ある場合、表 3.1に示す役割に対応するフレーズを名詞の後に付け、式(1)の値が最も高くなる役割を採用する。例えば、「家に」の名詞にフレーズをつけたものは「家に行って」「家にて」「家に対して」「家が原因で」「家に向かって」となる。

$$P(A) = \frac{\text{「名詞+フレーズ」のヒット件数}}{\text{「フレーズ」のヒット件数}} \quad (1)$$

そして、役割が1文の中で重複することがないように、かつ1文の中で各名詞の式(1)の合計が最大となるように、各名詞の役割を特定する。最後に、1文の中から「場所格」「目標格」「源泉各」の役割を持つと特定された名詞のうち、式(1)の値が最大のもを場所を表す単語として、「対象格」「道具格」「目標格」の役割を持つと特定された名詞のうち、文中の動詞のある箇所にもっとも近い名詞を対象物を表す単語として抽出する。

表 3.1: 名詞の役割を特定するためのフレーズ

格助詞	役割	フレーズ
に, で	場所格	に行って, にて
で	道具格	を使って
に, で, により	原因	が原因で
が, は, で	主格	の集まり
と, で, より	修飾節	のまま
を, に, と	対象格	に対して
に, へ, まで	目標格	に向かって
から, より	源泉格	から

3.4 シーンの切り出し

シーンの切り出しでは、背景が同じ文をまとめることによって、ブログを切り分ける。場所を表す単語から次の場所を表す単語の直前の単語までを1シーンとして切り出す。

3.5 シーンを構成する対象物と動作を表す単語の抽出

1シーンの構成要素の候補の単語集合から、対象物、動作を表す単語を抽出する。次に、シーン内の全ての対象物を表す単語のうち式(2)の値が最も高いものをシーンを作る対象物を表す単語として抽出する。シーンを構成する動作を表す単語は抽出された対象物を表す単語を含む文の動詞とする。p はシーン内の場所, t はシーン内の対象物を表す単語である。

$$relation(p, t) = \frac{hit(p \wedge t)^2}{hit(p)hit(t)} \quad (2)$$

3.6 画像を用いたシーンの生成

シーンの生成では、特定された場所と対象物と動作を表す単語を表現する画像からシーンを生成する。動作を表す画像は用

意したデータベースから取得する。場所と対象物を表す画像は単語を Yahoo!のイメージ検索[Yahoo!]にかけ、検索結果1位の画像を取得する。場所を表す画像を最奥、対象物、動作を表す画像は手前に配置する。このとき、データベースにある動作を表す画像の人物の向きに合わせるため、対象物、動作を表す画像の表示位置を「右上」「右下」「左上」「左下」のいずれかを動詞に合わせて指定し、シーンの画像を生成、出力する。データベース内の動作を表す画像の一部を図 3.2 に示す。データベースには30種類の画像がある。



図 3.2: 動作を表す画像の例

3.7 出力

図 3.3 にインターフェースの概観を示す。インターフェース左側はブログ記事を表示する。インターフェースの左上はURL入力部分と読み込みボタンがある。インターフェース右上は読み込んだブログ記事のシーンの一つが出力される。下にある矢印のボタン、サムネイルでシーンを切りかえられる。



図 3.3: インターフェースの概観

4. 評価実験

提案システムの有効性を確認するために3つの実験を行った。

4.1 シーンの切り出し精度評価

(1) 実験内容

システムがブログ作者のいる場所が変わった箇所でもブログを切り出せることを確認する。システムにブログ記事を入力し、切り出されたシーンと、被験者の回答を比較した。実験の手順を以下に示す。

手順 1: システムにブログの記事を入力し、システムによってシーンを切り出す。

手順 2: 被験者にブログ記事を、ブログ作者のいる場所が変わったと思う文と文の間に線を引いてもらう。

手順 3: システムの出力と被験者の回答を比較する。

Google ブログ検索[Google]で「で買った」「を食べた」「に行った」を検索し上位から店に関する記述を含むブログ記事 60 件取得したブログ記事のうち複数のシーンが含まれる 30 件のブログ記事を被験者に読んでもらった。被験者は 15 名で、被験者 1 人当たりブログ記事 10 件を読んだ。

(2) 結果と考察

切り出した文間のずれが前後2文までを一致するとみなしたときのシーンの切り出しの適合率と再現率を表 4.1 に示す。表 4.1 では 5 人中1人が正解としたシーンから 5 人中 5 人が正解としたシーンの適合率と再現率を示している。適合率と再現率は式 (3), (4)から求めた。

$$\text{適合率} = \frac{\text{正解と一致した数}}{\text{システムの出力数}} \quad (3)$$

$$\text{再現率} = \frac{\text{正解と一致した数}}{\text{正解の数}} \quad (4)$$

被験者切り出したシーンにはずれがあったため、前後2文までのずれをよしとした。過半数の被験者が切り出したシーンを正解としたときの適合率は 0.75 であった。これは場所が変わった文から文までをシーンとみなす提案システムの手法が適切だったためと考えられる。したがって、提案システムはシーンを切り出せると分かった。

表 4.1 の再現率は過半数を正解としたときに 0.496 と低い。この理由はシステムよりも被験者のほうがブログを細かく切り出したためと考えられる。システムがシーンを切り出せなかった理由を表 4.2 に示す。被験者が切り出したシーンには、回想などでシーンが切り替わったと判断し、イベント文でない文で切り出されたものがあつた。イベント文以外からも場所を表す単語を抽出する必要があると考えられる。しかし、イベント文以外では、感想や意見など、場所を含まない文が多く、全ての文から場所を表す単語を抽出すると、ノイズの単語が多く得られる。したがって、これらの文は提案システムの対象外の文とみなす。また、感想や意見などで文中に場所を表す単語がないものがあつた。文中に場所を表す単語がないとシーンを生成することはできないため、システムはシーンとして切り出す必要は無いと考えられる。以上より、提案システムによって、シーンの切り出しが行えたとわかつた。

表 4.1: シーン切り出しの適合率と再現率

何人以上が切り出したシーンを正解としたか	1/5 以上	2/5 以上	3/5 以上	4/5 以上	5/5 以上
適合率	0.938	0.85	0.75	0.613	0.45
再現率	0.28	0.389	0.496	0.57	0.621

表 4.2: システムが正解のシーンを切り出せなかった主な原因

切り出さなかった原因	シーンの数
文がイベント文ではない	26
文中に場所を表す単語が存在しない	24
場所を表す単語の後に格助詞がない	3
形態素解析のミス	4

4.2 シーンを構成する単語の抽出精度評価

(1) 実験内容

システムがシーンを構成する単語を抽出できることを確認した。システムによって切り出されたシーンをシステムに入力し、得られたシーンを構成する単語と、正解の単語を比較した。実験の手順を以下に示す。

手順 1: システムによって切り出されたシーンをシステムに入力し、システムによってシーンを構成する単語を抽出する。

手順 2: 被験者にシーンからブログ作者が「どこで(名詞)」「何を(何が, 何に, 何と)(名詞)」「どうした(動詞)」かを回答してもらい、正解の単語集合とする。

手順 3: システムの出力と正解を比較する。

先の実験で使用したブログ記事 60 件をシステムにかけ、切り出したシーン 251 個のうち、そのシーンだけを読んで意味が分かり、1 文ではない 31 個のシーンを被験者に読んでもらった。被験者は 20 名で、31 個の切り出されたシーンをそれぞれ 10 人の被験者に読んでもらった。

(2) 結果と考察

被験者に答えてもらった単語全てを正解としたときの場所、対象物、動詞を表す単語それぞれの適合率と再現率を表 4.3 に示す。表 4.3 で場所を表す単語は適合率、再現率共に高いが、対象物と動作を表す単語は再現率が低い。対象物を表す単語の再現率が低かった原因は、表 4.5 に示した、「対象物を表す単語の後に格助詞がない」文が多かったことが原因である。システムは、システムは格助詞から名詞の役割を特定するので、格助詞が後にない名詞は抽出されない。このため、対象物が何も出力されないことがあり再現率が低かった。「シーン中に場所がないとシステムが判断した」ことも対象物を表す単語の再現率が低かった原因である。システムでは、対象物を表す単語を抽出するとき、式(2)を計算し値の最も大きい名詞を対象物を表す単語として抽出する。しかし、システムが場所を表す単語が抽出できないと式(2)を計算できないため、同時に対象物を表す単語も抽出されなかつた。

場所、対象物、動詞を表す単語の3つが意味を成して一致しているときの適合率と再現率を表 4.4 に示す。適合率が 0.714 であつたことから、提案システムは 3 つの単語を抽出できると分かつた。しかし、3 つの単語が意味を成して一致しているときの適合率が、場所、対象物、動作を表す単語のそれぞれの適合率よりも低かつた。これは、場所、対象物、動作を表す単語一つ一つは抽出できているが、3 つの単語が意味を成して一致して抽出できていないことがあつたためである。システムは、対象物を表す単語を抽出するとき式(2)が最大となる対象物を表す単語を抽出する。このため、場所、対象物、動作を表す単語それぞれは取れていても、場所を表す単語と対象物、動作を表す単語が一致しなかつた。

表 4.3: 被験者が回答した全ての単語を正解としたときの場所、対象物、

動作を表す単語の適合率と再現率

	場所	対象物	動詞
適合率	0.857	0.857	0.81
再現率	0.774	0.581	0.548

表 4.4: 3 つの単語が意味を成して一致しているときの適合率と再現率

適合率	再現率
0.714	0.484

表 4.5: システムが正解の単語を抽出できなかった原因

場所	単語数	対象物	単語数
場所を表す単語の後に格助詞がない	3	対象物を表す単語の後に格助詞がない	6
場所を表す単語の後に付く格助詞が場所の役割を持たない	2	シーン中に場所がないとシステムが判断した	2

4.3 単語を表す画像取得精度の評価

(1) 実験内容

Yahoo!のイメージ検索から場所と対象物を表す画像が取り出せることを確認した。システムにブログ記事を入力し、得られたシーンを構成する単語を Yahoo!のイメージ検索にかけ、画像を取得した。単語との一致を被験者に回答してもらった。実験の手順は以下の通りである。

手順 1: システムによってブログ記事からシーンを構成する場所と対象物を表す単語を取り出す。

手順 2: 場所と対象物を表す単語を Yahoo!のイメージ検索にかけ 1 位から 5 位までの画像を取得する。

手順 3: 検索した単語と画像の内容との「一致」「不一致」「どちらでもない」を答えてもらう。

手順 4: 作業被験者が「一致」と答えた画像の割合をもとめる。

4.1 で用いた同じ、60 件のブログ記事をシステムにかけ取得した場所と対象物を表す単語 306 個のうち、表 4.6 に示す、目に見えない単語 30 個を除き、画像へのリンクが切れているものも除いた単語 168 個の画像合計 840 枚を見て回答してもらった。被験者は 20 人で画像 1 枚あたり 10 人の被験者に見てもらった。

表 4.6: 除いた単語

用事,性格,間際,種類,敷タイプ,前々,全盛,事,風,感想,途端,事情,連休,様子,他,気分,クリスマスセール,おかげ,久しぶり,最後,次,勇気,少なめ,原価,時点,文句,仕上げ,買い,社会,ランチ事情,
--

(2) 実験結果

被験者の過半数が「一致」と答えた画像を正解としたときの全画像における適合率は 0.421 であった。これは、「お昼、スーパー、コロッケ、トロ、バー」などの多義語では、さまざまな画像が検索されてしまったこと、表 4.7 に示す画像として表しにくい単語が多く適切な画像が検索されにくかったためである。

被験者の過半数が「一致」と答えた画像を正解としたときの各単語における正解の画像がはじめて出てきた順位が 1~5 位である単語の数と全ての単語における割合を表 4.8 に示す。表 4.8 では初めて正解の出てくる順位が 1 位である単語が一番多く、検索結果の順位と適切な画像には関連があると分かる。すなわち、検索結果の 1 位から 3 枚以上の画像を表示することでユーザはシーンを理解しやすくなると考えられる。

被験者が「一致」と答えた画像を正解としたとき、画像 5 枚全てが正解であった単語と全てが誤りであった単語を表 4.9 に示す。表 4.9 で、下位の単語は、「店」であれば「カレーの店」のように、他の単語と合わせて使う単語が多く、上位の単語は、それ以外の意味では使われない単語が多かった。多義語の対策として Web 上の画像集合にラベルをつけクラスタリングを行い、ブログ記事の内容と関連の強い画像を表示することで精度を上げることができる[砂山 04]と考えられる。

表 4.7: 画像として表しにくい単語の例

うしろ,先輩,お客さん,中,お昼,店,まわり,東ホール,もの,湯,インターネット,隣町,スタッフ,フック,学級委員,印鑑証明,

表 4.8: 各単語における初めて正解の出た順位と単語の数

	1 位	2 位	3 位	4 位	5 位
単語の数	78	19	22	9	3
割合	0.419	0.102	0.118	0.048	0.016

表 4.9: 全て正解だった単語と不正解だった単語

全て正解	露天風呂,旅館,夕飯,部活,ソフトクリーム,バイク,パン,ラジカセ,ロードバイク,椅子,花,絵,玄関,山,車,食堂,神社
全て誤り	うしろ,お客さん,お昼,お年玉,まわり,もの,インターネット,クルーザー,サンマルコ,スーパー,スタッフ,テレビ,トイレトペーパー,ブルブルマシーン,ラーメン屋さん,学級委員,顔,形,次女,実家,主人,場,場所,職場,親,身,先輩,中,店,東ホール

5. 結論

本論文ではブログ記事からイベントを表すシーンを画像として生成するシステムを提案した。本システムは、ブログ記事からイベント文を抽出し、場所を表す単語をもとにシーンを切り出し、シーンを構成する単語を抽出、画像検索と用意した画像のデータベースを用いてシーンを生成する。評価実験を行いシステムがシーンを切り出せること、シーンを構成する単語を抽出できること、単語を表す画像を取得できることを確認した。

今後の課題には、単語を表す適切な画像を取得するために、取得した画像の選択方法を考え、異言語間でのコミュニケーションを支援するシステムを構築していく。

参考文献

[小作 02] 小作, 内山, 井佐, 河野, 木戸出: 新聞記事コーパスでの単語出現特徴を利用した観光イベント情報の検索支援, 人工知能学会論文誌, Vol.19, No.4, pp.225-233, 2004.

[野呂 05] 野呂, 乾, 高村, 奥村: イベントの生起時間帯判定, 情報処理学会研究報告, 2005-NL-170, pp.7-14, 2005.

[渡邊 04] 渡邊, 太田, 片山, 石川: 格文法を用いた複数文書融合手法, DBSJ Letters, Vol.3, No.2, pp.93-96, 2004.

[角 05] 角, 田中: 理解支援のための電子コンテンツからアニメーション絵本へのメディア変換, The 19th Annual Conference of the Japanese Society for Artificial Intelligence, 2005.

[伊藤 06] 伊藤, 角, 久保田, 西田: Photo Chat: 互いの視点画像に「書き込む」事によるコミュニケーション支援, 人工知能学会全国大会第 20 回論文集, pp.263-pp.267, 2006.

[山田 02] 山田, 関: 異言語間コミュニケーション支援システムの構築, 武蔵工業大学環境情報学部情報メディアセンタージャーナル, 第3号, pp.64-69, 2002.

[Kudo 03] T. Kudo, Y. Matsumoto: Fast Methods for Kernel-Based Text Analysis, Proceedings of the 41st Annual Meeting of the Association for Computational Linguistics, pp. 24-31, 2003.

[Yahoo!] Yahoo!Japan:(URL)<http://search.yahoo.co.jp/>

[Google] Google:(URL) <http://blogsearch.google.co.jp/>

[砂山 04] 砂山, 永田, 谷内田: 画像検索のための Web テキストによる画像クラスタリング, 人工知能学会論文誌, 19 巻, 6 号, pp.580-588, 2004.