

地域 Blog 情報の可聴化インタフェースの提案と 音声化に適したテキストコンテンツの抽出手法

郡 宏志[†] 手塚 太郎[†] 田中 克己[†]

[†] 京都大学大学院情報学研究科 社会情報学専攻

〒 606-8501 京都市左京区吉田本町

E-mail: †{kori,tezuka,tanaka}@dl.kuis.kyoto-u.ac.jp

あらまし ユーザの置かれた環境によって利用可能なコンテンツのメディアは異なる。自動車の運転者による視覚的なコンテンツの利用は、安全性に支障を与えるため推奨されない。よって、自動車の中で通常利用されるコンテンツは音声コンテンツである。しかし、既存のラジオコンテンツや音楽は地域性が乏しく、旅行者にとって最適とは言えない。一方、増加を続ける Blog コンテンツには個人の特定の場所における体験や感想、最新の話題といった地域性の高い情報が多く含まれている。そこで、我々は自動車内のユーザに対し、地域性の高い Blog コンテンツを音声合成により可聴化して提示する BlogCarRadio システムを提案する。このシステムによって、ガイドブックに記載されていない情報をユーザに配信し、地域に対する理解を促進することが可能になる。Blog エントリは Blog 検索エンジンによって取得され、可聴化に対する適性や自動車の現在地からの距離、そして、カテゴリに対する適合度に基づいて選択される。本論文ではシステムに適した Blog コンテンツの取得手法と、特にそれを実現するのに必要な機能である、音声コンテンツへ変換するのに適したテキストコンテンツを抽出する手法について議論した。

キーワード Blog, マルチメディア, 最小注意負荷, 空間情報検索

Text Extraction and A User Interface for Regional Blog Sonification

Hiroshi KORI[†], Taro TEZUKA[†], and Katsumi TANAKA[†]

[†] Graduate School of Infomatics, Kyoto University

Yoshida-honmachi, Sakyo, Kyoto, 606-8501 Japan

E-mail: †{kori,tezuka,tanaka}@dl.kuis.kyoto-u.ac.jp

Abstract The media content presented to a vehicle driver is mainly auditory, since visual media is distracting and viewing it increases the chance of an accident. Music and talk are thus commonly listened to while driving. However, these types of content rarely reflect regional characteristics and are therefore not well suited for tourists who want to get information about the region they are visiting. We have developed the Blog Car Radio system that present blog entries in auditory style using sonification (speech synthesis). Blog entries are obtained from blog search engines, and selected based on their appropriateness for sonification, distances from the vehicle's current location, and relevance to the user-specified category. By using Blog Car Radio, a driver can obtain local information with only a small amount of distraction. In this paper, we particularly discussed the method to extract text contents which is suitable for sonification, and verified the effectiveness of the method.

Key words Blog, Multimedia, Minimal Attention, Spatial Information Retrieval

1. はじめに

情報システムは、様々な状況において利用されるようになりつつある。それに伴い、それぞれの状況に対して最適なコンテンツ提示手法が必要とされている。特に重要な例が自動車を運転するという状況である。というのも、運転者は特に自動車の

周辺に対して注意を払わねばならず、視覚による提示手法は適切ではない。よって、運転時におけるユーザインタフェースは、運転者にとっては最低限の注意しか要求されないものでなければならぬ [1]。この為、音楽やラジオは運転者が視聴するコンテンツとしては最も一般的なものである。しかし、このようなコンテンツは一般的に運転者の空間的状况に対してその関連性

は大きくない。

そこで、我々は地域 Blog(Weblog) 情報を音声コンテンツとして提示する BlogCarRadio システムを提案する。本システムは、地理的に関連性の高い Blog エントリを選択し、それを可聴化に対する適性とそれが含むカテゴリ情報に基づいてランキングする。そして、音声合成を利用してそれを可聴化し、ユーザに提示する。多くの Blog エントリが地域性の高い個人的な経験や感想を含んでおり、それは旅行者にとって有益な情報である。Blog エントリはしばしば口語的な表現が用いられるが、それは実際のラジオ番組で用いられるスタイルとも共通な点である。その土地に不慣れた旅行者は、このシステムから有益な情報を得ることが可能となる。また、普段同じ経路を利用するユーザに対しても、システムは異なる情報を提供することが可能となる。というのも、新しい Blog エントリは定期的に投稿されることが期待されるからである。

システムは、以下の 4 つの機能から構成される。

- Blog 検索のためにクエリを生成し、検索する
 - 取得した Blog を地理的特徴に基づき、フィルタリングする
 - フィルタリングした Blog からそのカテゴリを抽出する
 - 幾つかの特徴量による統合スコアによりランキングする
- 本論文においては、以上の 4 つの機能について説明する。まず、2. 節において関連研究について述べる。そして、3. 節においてシステムの概要を説明する。4. 節では可聴化に適したコンテンツの特性に関して議論し、5. 節では Blog エントリのランキング手法について説明する。そして、6. 節において可聴化に適した Blog エントリの抽出手法についての評価を行い、7. 節で今後の課題を、そして、8. 節で結論を述べる。

2. 関連研究

筆者ら [2] は、これまでにユーザが Web により客観的情報を閲覧する状況を想定し、その際にシステムが閲覧ページに関連する Blog エントリを主観的な情報として副次的に音声で提示するシステムを提案した。しかし、その際に提示される Blog エントリに対して、そのエントリが音声化に適したコンテンツかどうかという評価は行っていなかった。そこで、我々は本論文で提案するシステムの機能の一部として、音声化に適した Blog エントリを評価する手法を提案する。また、ラジオの利用形態として一般的なのは自動車内での利用であるため、本論文ではユーザが自動車内で Blog を聞く状況を想定し、空間的な要素についても言及する。

2.1 Web の可聴化に関する研究

Web の可聴化に関する研究としては、主に視覚的に障害のあるユーザの Web に対するアクセシビリティの向上を目的とした研究がある [3] ~ [5]。これらは、3D オーディオを使用する等と音声出力を工夫したり、幾つかの閲覧モードを用意する等とユーザの負担を減らしたりすることにより、Web コンテンツを聴覚により閲覧可能としている。本論文では、Web コンテンツの中では最も一般的なテキストコンテンツを対象としている。さらに、それを従来の Web の可聴化に関する研究の様に聴覚

により閲覧可能な形に変換するというアプローチではなく、聴覚による閲覧に適したコンテンツを抽出するというアプローチを採用している。このようにして、視覚的に閲覧するのに有益なテキストコンテンツと聴覚で閲覧するのに適したテキストコンテンツを区別することを試みる。

2.2 Blog と地理情報の統合

Blog と地理情報の統合に関する研究として、上松ら [6] は「場 log」というシステムを提案している。これは、カメラや GPS といった機能が付加された携帯電話端末を用いて、各人が個々に発信した情報を位置情報に基づいて整理し、新たな情報閲覧を可能にするシステムである。また、サービスとしても Blog を空間的にマッピングするものとして様々なものが公開されている [7], [8]。しかし、これらはユーザが自ら登録するという作業が必要となる。我々のシステムでは、このような地域性を持った Blog を、地名をクエリとして検索するため、地名とその地理的位置情報が存在する限り、あらゆる地域の Blog を対象とすることが可能である。

一方で、Hurst [9] らは Blog エントリを地名により地図に対してマッピングしており、これは各ユーザが自ら登録するという作業は必要ない。Hurst らはこうして複数の Blog ホスティングサービス間の相違点の発見を試みている。また、検索サービス maplog [10] では、各 blog サービスの更新情報等によって公開されている Blog に対してクローラを巡回させることによりそれらを取得する。そして、その中から位置キーワードという位置情報を有する単語を取得しその Blog を地図上に自動的にマッピングする。サービスの提供する機能として、地図で表示している範囲内にマッピングされている Blog をキーワードで検索し、エントリが投稿された時間情報に基づき、最新のものが上位にランクされるよう表示する。本システムは、各 Blog エントリをラジオメタファによる提示に適当なコンテンツが上位にランクされるようランキングする。また、倉島ら [11] は、Blog から人々の体験を、時間・空間・動作・対象属性間の相関ルールマイニングによって抽出することで、それらを検索・要約することのできるシステムを提案している。倉島らが体験の要約や検索を目的としているのに対し、我々はラジオメタファを用いて自動車内のユーザに人々の生の意見や体験を提示することを目的としている。

3. システム概要

本システムの目的は、音声による形式で地域 Blog エントリを自動車を運転しているユーザに提示することである。図 1 で示されるように、システムはクライアントサーバモデルにより構成され、情報を取得する各自動車がクライアントに対応する。クライアントは、自動車の位置、速度、進行方向といった情報を定期的にサーバに送信する。サーバは Blog を検索し、適切なものをクライアントへ配信する。システムの閲覧は、ユーザに要求する注意負荷を最小とする”Minimal Attention User Interface” [1] という構想に基づいている。その為、ユーザの動作は最小限のものとする。数少ないユーザの動作の一つとしては、カテゴリの選択が挙げられる。これは、実際のラジオにお

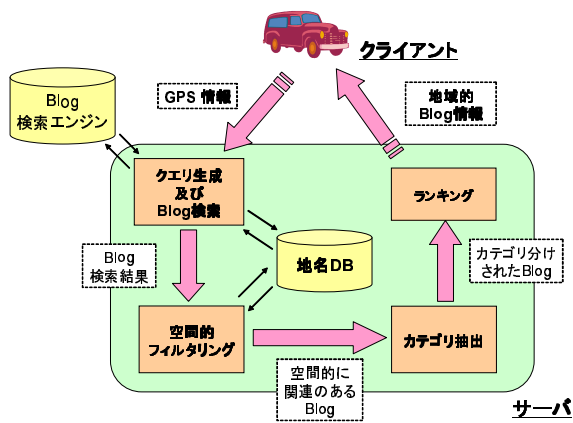


図1 BlogCarRadio システム概要

けるチャンネルの選択に類似するが、本システムにおいてはそれぞれのカテゴリがそれぞれあるトピックに該当する点が異なる点である。トピックとしては、食事やイベントなどが挙げられる。本システムの処理は、以下の4つのフェーズから成る。

Phase1: Blog 検索

Phase2: 空間的フィルタリング

Phase3: カテゴリ抽出

Phase4: ランキング

以下の各節においてこれらのフェーズについて詳しく述べる。

3.1 Blog 検索

システムにおける最初のフェーズとして、GPS(global positioning system)によって取得される自動車の座標データがBlogの検索に使用される。検索は、既存のBlog検索エンジンにより行う。座標データから、地名と座標を有するデータベースを用いて、自動車の現在地から近傍に存在する地名の集合を取得し、それらの地名の選言を検索クエリとして用いる。

3.2 空間的フィルタリング

Phase1により取得した検索結果から、自動車の現在地からなるべく近傍の地名について言及しているBlogを抽出するため、検索結果のフィルタリングを行う。6.2節の実験における事前処理において、1つのBlogエントリ中の平均地名数は、8.2であった。このように、地域性のあるBlogエントリの中では、複数の場所について言及されることが多い。図2におけるBlogエントリA、Bのように、どちらもユーザの近傍の同じ地名が含まれていても、そのエントリが含む他の地名はユーザの近傍に存在しない可能性がある。そのようなBlogエントリを取り除くため、各エントリをその中に含まれる地名と現在地との平均距離で評価し、上位*n*件のエントリを取得する。その評価式は以下となる。

$$G(E) = \text{avg}(\delta(l, p)) \quad (1)$$

*l*がエントリ*E*に含まれている地名の要素で、*avg()*が平均を、 $\delta(l, p)$ がユーザの現在地と地名間の距離を表す。ただし、*l*が地名を、*p*がユーザの現在地を表す。

3.3 カテゴリ抽出

カテゴリと関連のあるBlogを提示するために、Phase2にお

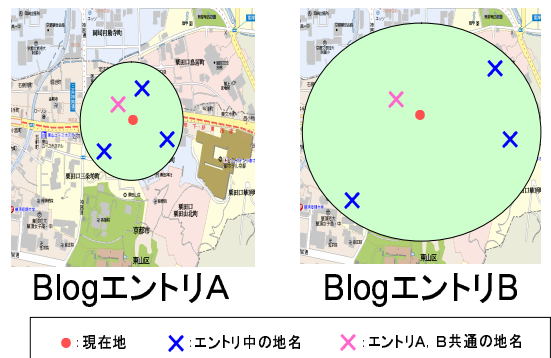


図2 2つのエントリ中に登場する地名の空間マッピング

けるフィルタリングで取得したBlogに対してカテゴリ抽出を行う。ここでは、あるカテゴリと関連のある単語をカテゴリ語と呼ぶ。カテゴリ語は、概念辞書を使用して取得し、こうしてそれぞれのカテゴリクラス*c*がカテゴリ語の集合*W_c*を有するコーパスを作成する。そして、各Blogエントリ*E*が含む単語の中からそれぞれのカテゴリクラスのカテゴリ語にマッチするものの延べ総数を求め、エントリ中の単語の延べ総数*m*で正規化する。よって、Blogエントリ*E*の各カテゴリクラス*c*に対するカテゴリ適合度*R(c, E)*は以下のように表される。

$$R(c, E) = \frac{1}{m} \sum_{w \in W_E \cap W_c} tf(w) \quad (2)$$

ただし、

$$m = \sum_{w \in W_E} tf(w) \quad (3)$$

*tf(w)*は、単語*w*のエントリ中の出現頻度を表す。このようにして表された*R(c, E)*の中で、最大値をとるカテゴリクラス*c*にBlogエントリを割り当てる。

3.4 ランキング

ランキングのフェーズでは、Blogエントリは以下の尺度に基づいてランキングされる。

- 可聴化に対する適性
- カテゴリ適合度

以下の4.節において、可聴化に適性のあるテキストコンテンツについて議論し、5.節においてランキングの仕組みを具体的に述べる。

4. 可聴化に適したテキスト情報の抽出手法

表1は、テキストコンテンツを視覚により閲覧した場合と聴覚により閲覧した場合の特性を、特に聴覚による閲覧の場合の制約に基づいて、比較したものである。このテーブルに関しては、一部黒橋ら[12]の研究に基づいて作成されたものである。このように、同じテキストでも、視覚により閲覧するのと同様に聴覚により理解するのも容易なコンテンツが存在する一方で、聴覚により閲覧する場合にはほぼ理解不可能なコンテンツも存在する。Blogの可聴化に際しては、このような聴覚による閲覧における制約を考慮する必要がある。

閲覧メディア	視覚	聴覚
重要な部分の見逃し	困難ではない	困難
記号の理解	容易	困難
1文が長い文章の理解	困難ではない	困難
頻出する指示語の理解	困難ではない	困難

表 1 閲覧メディアによる違い

4.1 テキスト形式の特徴

Blog には、聴覚により閲覧した場合に理解が困難となるテキストの形式上の特徴が含まれていることがある。例えば、顔文字等といった記号表現は Blog でしばしば用いられるが、そのような記号を可聴化すると、ユーザは理解できず混乱する。黒橋らは、書き言葉を話し言葉へ変換する場合に直面するこのような幾つかの特徴について議論した [12]。我々のシステムにおいてもこのような特徴をいくつか考慮する。本システム上で評価する特徴量は、以下の値である。

- 記号の出現率: 文書中の文字数に対する記号数の割合
- 平均文長: 文書中の文の平均長
- 指示語の出現率: 文書中の単語数に対する代名詞数の割合

すると、テキスト形式の特徴に基づく Blog エントリーの重要度 $T(E)$ は、以下のように定義される。

$$T(E) = t'_1(E) + t'_2(E) + t'_3(E) \quad (4)$$

$t'_i(E) (i = 1, 2, 3)$ は、評価した特徴量を正規化したものであり、平均 μ_t 、標準偏差 δ_t 及びデータの实测値 t_i を用いて以下のように表される。

$$t'_i = \frac{\mu_t - t_i}{\delta_t} \quad (5)$$

4.2 キーワードの出現位置

コンテンツを視覚により閲覧する場合と聴覚により閲覧する場合には、重要な相違点の一つ存在する。それは、聴覚による閲覧の場合は、重要で無い部分を省略して閲覧することが困難、あるいは不可能であるという点である。Blog エントリーの一部だけがユーザにとって重要であるという状況もしばしば存在する。そのような場合、テキストを視覚により閲覧する場合は、ユーザは興味のあるトピックが見つかるまでテキストを読み飛ばすことが出来る。一方で、コンテンツを聴覚により閲覧する場合は、ユーザは重要な部分が閲覧出来るまで待たなければならない。それは、ユーザにとっては負担となる。特に、重要でない部分が長いほどユーザの負担は大きくなる。キーワードの出現をユーザにとって重要な情報の出現と仮定すると、以上の議論から、キーワードの出現位置が重要となる。読み飛ばす負担を考慮すると、テキストを視覚により閲覧する場合もキーワードの出現位置は重要なのだが、それは聴覚により閲覧する場合ほどではない。以上のテキストの音声コンテンツへの可聴化の特性を図 3 に示す。

また、談話解析の手法としてセンタリング理論 [13] が存在する。これは、文中で話題の中心になっている事物、すなわちセンターが連続する場合にはセンターである事物を代名詞で表現

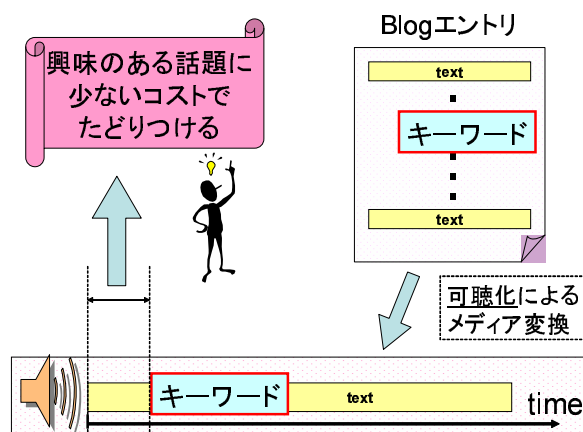


図 3 メディア変換における特性

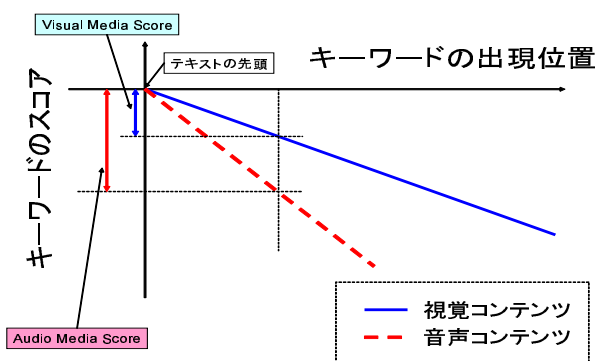


図 4 キーワード出現位置に対するスコアリングモデル

するはずである、という基本原則に基づいて談話の解析を行うものである。この原則に従うと、話題の中心は先に出現し、その後は省略されることが多いと考えられるため、テキストの最初の部分だけにキーワードが出現するとしても、テキストの最後の部分にキーワードが出現する場合と比較して、文書全体がユーザにとって重要な情報である確率は高い。これは、空間的特性から、特に地名について重要である。テキストにおいて一度地名が出現すると、それ以降のテキストはその場所でのイベントであることが明記されていなくても文脈からそうと判断可能なことが多く、地名が繰り返し出現することはあまりない。したがって、地名の場合、文書の最初の部分に出現すると、そのテキスト全体がその地名に関係があり、結果重要である可能性は大きい。

しかし、ここで議論されている違いは、メディアの種類に起因している。聴覚での閲覧のために Blog エントリーをランキングするため、我々のシステムは、重要な部分はコンテンツの最初に出現するという事実をより重視しなければならない。以上のモデルをグラフにより定式化すると、図 4 のようになる。このように、キーワードがテキストの最も先頭に出現する際は、コンテンツを視覚により閲覧する場合も聴覚により閲覧する場合も負担は等価であるため、キーワードに対するスコアも等価となる。けれども、キーワードの出現位置がテキストの後方になればなるほど、聴覚により閲覧する場合の負担は視覚の場合

と比較して大きくなり、そのキーワードに対するスコアは小さくなる。したがって、各々のメディアにおけるキーワードの評価式 $p(x)$ は、キーワードの出現する先頭からの距離（以下、出現位置とする）を x とすると、以下のようになる。

$$p(x) = \begin{cases} -\beta_v x & (\text{visual media}) \\ -\beta_a x & (\text{auditory media}) \end{cases} \quad (6)$$

$$0 < \beta_v < \beta_a \quad (7)$$

この場合の傾き β_v, β_a は、キーワードの出現する場所をどれだけ重視するかという係数とも考えることができる。また、本システムにおけるキーワードには、地名とカテゴリ語と 2 種類存在する。したがって、本システム上で評価する特徴量は、以下の値である。

- 地名の出現位置
- カテゴリ語の出現位置

ただし、評価するカテゴリ語は、そのエントリが分類されたカテゴリに属するカテゴリ語とする。それぞれの地名の出現位置を l_i 、カテゴリ語の出現位置を c_j とすると、キーワードの出現位置に基づく Blog エントリ E の重要度 $P(E)$ は、以下のよう

$$P(E) = \frac{1}{m+n} \left\{ \sum_{i=1}^m p(l_i) + \sum_{j=1}^n p(c_j) \right\} \quad (8)$$

ただし、 m, n はそれぞれエントリ中に出現する地名、カテゴリ語の延べ総数である。

5. Blog ランキング手法

可聴化に対する適性に加えて、Blog エントリはそのカテゴリ適合度によっても評価される。そして、ユーザに対して提示されるエントリは、最終的に統合されたスコアの内、最も高いスコアを持ったエントリである。

5.1 カテゴリ適合度

カテゴリ抽出フェーズに加えて、ランキングフェーズにおいてもカテゴリ適合度を利用する。文書に含まれるカテゴリ語数の割合はその文書のカテゴリに関する適合度としてはもっとも基礎的な尺度である。以下、エントリが分類されたカテゴリ c に対するカテゴリ適合度 $R(c, E)$ を $R(E)$ とする。

5.2 統合スコア

本システムでは、テキスト形式の特徴に基づいたスコア $T(E)$ 、キーワードの出現位置に基づいたスコア $P(E)$ 、及びカテゴリ適合度 $R(E)$ という 3 つのスコアを統合してエントリをランキングする。そして、最も大きいスコアを有するいくつかのエントリを実際にユーザに提示する。すると、エントリ E に対する統合スコア $S(E)$ は、以下のように定式化される。ただし、 α 、 β 及び γ は任意の係数である。

$$S(E) = \alpha T'(E) + \beta P'(E) + \gamma R'(E) \quad (9)$$

$$\alpha + \beta + \gamma = 1 \quad (10)$$

ここにおける各項のスコア $T'(E), P'(E), R'(E)$ は、それぞれ元のスコア $T(E), P(E), R(E)$ を正規化したものであるとする。正規化を行うのは、それぞれ異なる単位の値であるため、そのままの値を使用しては、それぞれの評価尺度を等価に扱うことが出来ないためである。 $T'(E)$ の場合は、式 4 の各項に対して適当な係数 $\lambda_i (i = 1, 2, 3)$ を適用し、以下のように表される。

$$T'(E) = \lambda_1 t'_1(E) + \lambda_2 t'_2(E) + \lambda_3 t'_3(E) \quad (11)$$

$$\lambda_1 + \lambda_2 + \lambda_3 = 1 \quad (12)$$

また、 $P'(E)$ に関しては、式 6 における β_a 及び β_v がそのまま式 9 における β に対応する。というのは、式 9 の統合スコアにおける β はキーワードの出現位置をどれだけ重視するかという係数を表し、それは式 6 の β も同様であるためである。

$$\begin{aligned} P(E) &= \frac{1}{m+n} \left\{ \sum_{i=1}^m p(l_i) + \sum_{j=1}^n p(c_j) \right\} \\ &= \frac{\beta}{m+n} \left\{ \sum_{i=1}^m (-l_i) + \sum_{j=1}^n (-c_j) \right\} \end{aligned} \quad (13)$$

上記の式が成り立つため、 $-l_i$ 及び $-c_i$ に対して正規化を行うと、それぞれの平均 μ_c, μ_l 及び標準偏差 δ_c, δ_l を用いて、 $P'(E)$ は以下のように表される。

$$P'(E) = \frac{1}{m+n} \left\{ \sum_{i=1}^m \frac{(-l_i) - \mu_l}{\delta_l} + \sum_{i=1}^n \frac{(-c_i) - \mu_c}{\delta_c} \right\} \quad (14)$$

また、 $R'(E)$ も同様に平均 μ_r 、標準偏差 δ_r を用いて以下のよう

$$R'(x) = \frac{R(E) - \mu_r}{\delta_r} \quad (15)$$

こうして得られた式 9 の統合スコアは、 α, β, γ の重み付けによりテキストコンテンツを視覚メディアとしても聴覚メディアとしても評価することが可能となる。まず、テキストコンテンツを視覚コンテンツとして評価する場合は、各重みを以下の条件を満たすように定める。

$$\alpha_v = 0, \quad 0 < \beta_v \ll \gamma_v \quad (16)$$

一方、聴覚コンテンツとして評価する場合は、以下の様に定める。

$$\beta_a \gg \beta_v > 0, \quad \alpha_a > 0, \quad \gamma_a > 0 \quad (17)$$

6. 実験

本節において、我々は 4. 節で議論したランキング手法により、可聴化に適した Blog が高くランキングされているかを検証する。その為、以下の実験を 4 回行い、それぞれの結果を統合し、適合率-再現率曲線を求めた。統合法に関しては、各質問に対する結果の再現率及び適合率を平均するマクロ平均 [14] を用いた。

(1) 特定の場所をユーザの現在地と仮定し、その場所から近傍の地名を空間データベースを用いて幾つか取得する

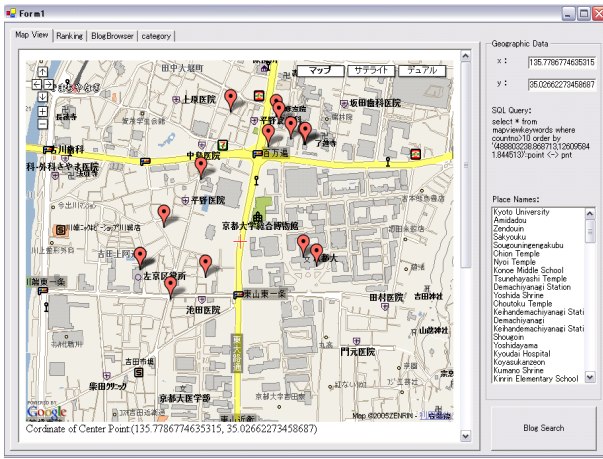


図 5 実験用システム・インタフェース

- (2) 取得した地名の選言をクエリとし、得られた Blog エントリが 200 件を超えるまで Blog を検索する
- (3) 得られた検索結果に対して空間的フィルタリングを行い、上位 100 件の Blog を取得する
- (4) 取得した Blog に対してカテゴリ抽出を行い、Blog をランキングする
- (5) 各 Blog エントリが正解かどうかを実際に音声で聞くことにより人手で分類し、適合率及び再現率を算出する
- (6) 上位 n 件に対する適合率及び再現率を算出し、適合率-再現率曲線を求める。($1 \leq n \leq 100$)

今回の実験では各エントリを正解かどうか判定するのに評価基準をなるべく明確にする為、以下の点に留意して正解データの作成を行った。

- 最後まで意味が理解出来る
- コンテンツの先頭近くに登場する話題がエントリが分類されたカテゴリを表している、もしくは地域的な話題である

6.1 実験環境

我々は、実験用のシステムを C# を用いて実装した。それを図 5 に示す。自動車の現在地を想定した Blog 検索時の入力座標は、システムの地図インタフェースから入力出来るものとした。地図インタフェースに関しては、Google Maps API [15] を用いて実装した。また、地図上のマーカーは、検索地点から最も近傍に存在する地名 20 件を空間上にマッピングしたものである。Blog 検索エンジンとしては、RSS 検索エンジンの Bulkfeeds [16] を用いた。本システムで抽出したカテゴリに関しては、日本観光協会による全国地域観光情報センター [17] における「旅行目的」の分類に依った。また、各カテゴリに対するカテゴリ語セットの抽出は、日本語語彙大系ソフト [18] から人手で分類を行い、作成した。カテゴリ及びカテゴリ語の一部を表 2 に示す。また、ゼンリンの住宅電子地図情報 [19] の京都に関するデータを空間データベースにおける地名及び座標データとして用いた。データベースとしては PostgreSQL [20] を使用し、空間検索手法として R-tree [21] を使用した。音声合成に関しては、ANIMO 社の Fine Speech [22] を使用した。

カテゴリ	カテゴリ語の一部
見る	神社、仏堂、城跡、博物館、溪谷、花見、紅葉
イベント	神事、火祭、ショー、バレンタインデー
遊ぶ	ハイキング、スポーツ、ビリヤード、百人一首
泊まる	ホテル、旅館、宿坊、ユースホステル、
乗る	ケーブルカー、ゴカート、人力車、遊覧バス
食べる	漬け物、湯葉、懐石料理、カレー、ラーメン
買う	古着屋、有名店、雑貨店、百貨店、デパート

表 2 カテゴリ及びカテゴリ語の一部

対象データ	平均	標準偏差
記号出現率	0.0583	0.0511
1 文の長さ	41.1	57.3
指示語出現率	0.0070	0.0084
地名出現位置	773	940
カテゴリ語出現位置	814	935

表 3 事前処理により取得した対象データの特性

6.2 事前処理

まず、テキストの形式的特徴、カテゴリ語及び地名の出現位置の平均及び標準偏差を求めた。その手法は、ランダムな空間上の位置において近傍の地名を取得し、その選言により Blog を検索する。得られた Blog のテキストの形式による各特徴量、地名及びカテゴリ語の出現位置を求める。以上の処理を何度も繰り返す。こうして、我々は 8719 件の Blog を解析した。このようにして求めたパラメータを表 3 に示す。

また、本実験の目的は、可聴化に適したコンテンツのランキング手法の有効性の検証のため、式 9 における各項の重みは以下のようにした。

$$\alpha = \beta = \frac{1}{2}, \quad \gamma = 0 \quad (18)$$

また、式 11 における各係数は以下のようにした。

$$\lambda_1 = \lambda_2 = \lambda_3 = \frac{1}{3} \quad (19)$$

6.3 評価

このようにして実験を行った結果得られたグラフが、図 6 である。図中の $S(E)$ が式 9 における $S(E)$ 、すなわち本論文で提案したスコアリング手法に対応する。各項の係数に関しては、6.2 節で定めた通りである。図から分かるように、再現率が 0.1 から 0.3 の間の場合には、高い適合率が得られていることが分かる。一方で、再現率が 0.0 から 0.1 の間で適合率が低下しているが、その原因は写真を含む Blog が不正解と判断されたからである。写真を含むような Blog の本文には、写真に対する説明が含まれており、本文だけを聞いても意味が理解できない場合が多い。また、このような Blog で特に「嵯峨大覚寺の心経宝塔」等と写真の対象である地名だけを記述した Blog もあり、その場合は本提案手法では高くランキングされる。以上から、Blog に写真が含まれるかを評価する必要があることが分かる。また、リンクが張られている Blog はリンクを参照しなければ意味が分からないものもあった。さらに、カテゴリ分けがうまくいっておらず、テキストの最初の部分に該当するカテゴリの内容が存在しない Blog も存在した。そのため、地域的

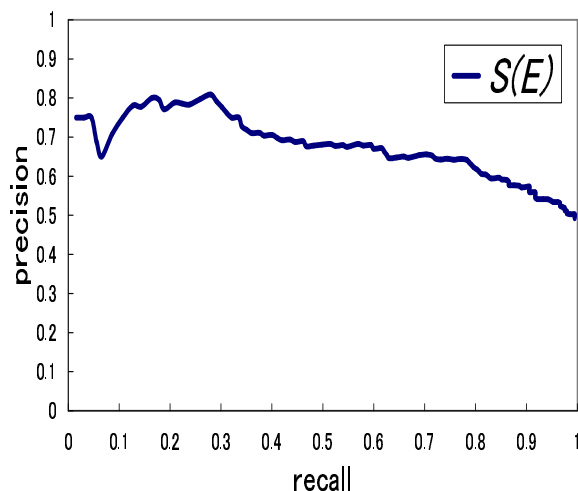


図 6 適合率-再現率曲線

な Blog のカテゴリ分けの精度を上げる工夫も必要である。表 4 に高い値にランクされた Blog エントリーの例を、表 5 に低い値にランクされた例を示す。表 5 の例においては、記号表現が視覚メディアには向くが、聴覚メディアには向かないことが分かる。視覚でこのコンテンツを閲覧する場合は、「」の数を比較することが出来るが、聴覚では聞きながら比べることが不可能である。

7. 今後の課題

今後の課題としては、Blog ランキング手法の更なる精度の向上のため、画像やリンクに関する解析、そして Blog のカテゴリ抽出精度向上が挙げられる。距離の評価手法に関しても、現在地と地名の位置という点同士の評価ではなく、自動車の経路と地名の位置という線と点の評価を行い、ユーザがある場所に最も近づく瞬間にその場所に関する Blog が聞ける等といったチャンネル編成を行い、ユーザが違和感なくシステムを利用できる工夫が必要となる。また、それと関連して、システムは常にユーザにコンテンツを提示するために、連続的に検索を行う。そのために、連続的にクエリを生成する必要がある。そのようなクエリを *Continuous Query* といい、こうしたクエリ生成手法についても、これから考えていく必要がある。

また、ユーザを真に満足させる為には、実際のラジオコンテンツのようなラジオ番組の生成も行わなければならないと考えられる。その為には音楽等といった Blog 以外のコンテンツを利用することも必要である。また、例えば、海沿いを走っている時や山を走っている時、もしくは自動車が渋滞に巻き込まれている時では聞きたい音楽は異なる様に、ユーザの嗜好は空間により変化する。したがって、最終的な目標は、空間的な要因からユーザを満足させるコンテンツを生成することである。

タイトル：京都ハリストス正教会（京都市）

カテゴリ：見る

スコア：0.662631128

本文：京都ハリストス正教会（京都市）。宿泊しているホテルの周辺を散策している途中に偶然見かけました。京都の街並のなかに突然現れた白い建物なのですが、不思議と街並と調和していました。やはり 100 年の歴史がそうさせるのかな？ロシア・ビザンチン様式の教会堂です。明治 36 年（1903）の建築です。設計は、松室重光という方だそうです。京都市指定文化財となっています。

URL：<http://kyotogoriyaku.blog24.fc2.com/blog-entry-106.html>

表 4 高くランキングされた記事の例

タイトル：繁栄をもたらす商売の神様！【京都・伏見稲荷大社】

カテゴリ：見る

スコア：-0.334775998

本文：【伏見稲荷大社】京都市伏見区深草藪之内町（地図はこちら） JR 稲荷駅下車徒歩 4 分・京阪伏見稲荷駅下車徒歩 6 分 ぼのこの的オススメ度（5 段階評価・星が多いほどオススメです） ・商売繁盛度 + ・家内安全度 ・五穀豊穡度 ・子授け度 ・開運招福度 伏見稲荷大社は、全国に約 4 万あるといわれる稲荷社（お稲荷さん）の総本社です「総本社」つつーくらいやから、全国の子社はここからビジネス運 UP のピームをもらってるわけですこちらの神社さん、お正月三日だけで 250 万人以上の人たちが初詣に来るんですじょ（すげー）…しかしこの人数、どーやって数えてるの？（…？）（いつも思うんやけどさ）

URL：<http://kyotogoriyaku.blog24.fc2.com/blog-entry-106.html>

表 5 低くランキングされた記事の例

8. まとめ

我々は、自動車内のユーザに対し、地域性の高い Blog コンテンツを音声合成により可聴化し、提示する BlogCarRadio システムを提案した。さらに、インタフェースの方針として、ユーザに対する注意負荷を最小とすることを目標とした。このシステムによって、ガイドブックに記載されていない情報をユーザに配信し、地域に対する理解を促進することが可能となる。また、システムに適した Blog コンテンツの取得手法を実現するために必要な機能として、Blog の空間的フィルタリング、そして、カテゴリ適合度とコンテンツの可聴化に対する適性を考慮したランキング手法を提案した。本論文では、その内の可聴化に対する適性の評価手法について特に議論した。そして、実験によりその手法の有効性を検証し、さらに今後の課題について述べた。

謝 辞

本研究の一部は、21 世紀 COE プログラム「知識社会基盤構築のための情報学拠点形成」、文部科学省科学技術振興費プロジェクト「異メディア・アーカイブの横断的検索・統合ソフトウェア開発」（代表：田中克己）、および、平成 17 年度科研費特定領域研究 (2)「Web の意味構造発見に基づく新しい Web

検索サービス方式に関する研究」(課題番号:16016247,代表:田中克己)によるものです。ここに記して謝意を表すものとします。

文 献

- [1] J. Pascoe, N. Ryan and D. Morse: “Issues in developing context-aware computing”, First International Symposium on Handheld and Ubiquitous Computing (HTC 99) (1999).
- [2] 郡宏志, 竹原幹人, 大島裕明, 小山聡, 田中克己: “Blogradio: Blog 情報の感情マイニングと可聴化に基づく web 閲覧補完”, 第 16 回データ工学ワークショップ (DEWS2005) (2005).
- [3] F. James: “Presenting html structure in audio: User satisfaction with audio hypertext”, ICAD 96 Proceedings, pp. 97–103 (1996).
- [4] M. Wynblatt, D. Benson and A. Hsu: “Browsing the world wide web in a non-visual environment”, ICAD 97 Proceedings (1997).
- [5] P. Roth, L. Petrucci, A. Assimacopoulos and T. Pun: “Ab-web: Active audio browser for visually impaired and blind users” (1998).
- [6] 上松大輝, 沼晃介, 徳永徹郎, 大向一輝, 武田英明: “場 log: weblog 環境における位置情報利用の提案”, 第 6 回人工知能学会セマンティック Web とオントロジー研究会 (2004).
- [7] DC Metro Blogmap
<http://www.reenhead.com/map/metroblogmap.html>.
- [8] nyc bloggers
<http://www.nycbloggers.com/>.
- [9] M. Hurst: “Gis and the blogosphere”, WWW2005, 2nd Annual Workshop on the Blogging Ecosystem: Aggregation, Analysis and Dynamics (2005).
- [10] maplog
<http://maplog.jp/>.
- [11] T. Kurashima, T. Tezuka and K. Tanaka: “Blog map of experience: Extracting and geographically mapping visitor experiences from urban blogs”, Proceedings 6th Web Information Systems Engineering(WISE2005), pp. 496–503 (2005).
- [12] 黒橋禎夫, 大泉敏貴, 柴田知秀, 鍛冶伸裕, 河原大輔, 岡本雅史, 西田豊明: “会話型知識プロセスのための言語情報のメディア変換”, 社会技術研究論文集, 2, pp. 173–180 (2003).
- [13] B. J. Grosz, A. K. Joshi and S. Weinstein: “Centering: A framework for modeling the local coherence of discourse”, Computational Linguistics, 21(2), pp. 203–225 (1996).
- [14] “Information Retrieval”, Butterworths, second edition.
- [15] Google Maps API
<http://www.google.com/apis/maps/>.
- [16] Bulkfeeds
<http://bulkfeeds.net/>.
- [17] 日本観光協会 全国地域観光情報センター
<http://www.nihon-kankou.or.jp/nkks/index.html>.
- [18] Japanese Vocabulary System
<http://www.ntt-tec.jp/technology/C404.html>.
- [19] Zenrin Co.,Ltd
<http://www.zenrin.co.jp/>.
- [20] PostgreSQL
<http://www.postgresql.org/>.
- [21] A. Guttman: “R-trees: A dynamic index structure for spatial searching”, In Proc. of the ACM SIGMOD Intl. Conf. on Management of Data, pp. 47–57 (1984).
- [22] ANIMO Fine Speech
<http://www.animo.co.jp/products/tts/fs/index.jsp>.