

ベイジアンネットワークに基づくライフログ検索モデル

牛尼 剛聡

九州大学 大学院芸術工学研究院

ushiana@design.kyushu-u.ac.jp

渡邊 豊英

名古屋大学 大学院情報科学研究科

watanabe@is.nagoya-u.ac.jp

概要 ライフログは同一人物の活動を記録した複数のメディア形式から構成されるデータ集合である。本論文では利用者が蓄積したライフログを検索するためのデータモデルを提案する。提案するモデルはベイジアンネットワークに基づく文書検索モデルである推論ネットワークを拡張したものである。本モデルはライフログ検索で特徴的な二種類の検索をサポートする。一つはコンテキストに基づいたメディアの検索の実現である。もう一つは不確実性を伴う検索である。これらの検索を実現する際に問題となるのは異種メディアの対応付けの方法である。提案手法ではメディアの時間空間的な特徴に基づいて依存関係の強さを推定する。

A Data Model for Searching Life-log Based on Bayesian Network

Taketoshi USHIAMA

Faculty of Design,

Kyushu University

ushiana@design.kyushu-u.ac.jp

Toyohide WATANABE

Graduate School of Information Science,

Naogoya University

watanabe@is.nagoya-u.ac.jp

Abstract A life-log set is a collection of media objects, which represent activities of a person. This paper introduces a data model for the searching of media objects in the life-log set of a user. The model is designed by giving some extensions to the inference network model, which is a data model for searching documents based on Bayesian network methodologies. Our life-log data model supports two types of searches: (1) the context-based search and (2) the search with uncertainty. One of the important problems on life-log composition is the specification of relationships between different media objects. We specifies the relationships based on temporal-spatial features of media objects.

1 はじめに

個人の経験を記録することにより人間の記憶を増強し、高度な知識活動を支援するシステムのコンセプトは、1945年に Bush により Memex として提案された [1]。Memex が提案された当時は、技術的な問題から実現は困難であった。しかし、技術の発達により Bush が示した Memex のコンセプトの実現が現実味を帯びてきたことから、近年に再び注目されている [2, 3]。

個人の経験を表す記録はライフログと呼ばれる。Memex のようにライフログを管理し必要に応じて検索可能とするシステムを実現するために必要な技術的な課題は二種類に大別できる。一つはライフログの取得方法であり、もう一つは、取得したライフログの管理方法である。

ライフログを取得するために、実世界で個人の活動をセンシングすることが考えられる。ユビキタスコンピューティングでは生活環境に各種のセンサを埋め込み、個人活動に関するデータを取得可能とする。また、ウェアラブルコンピューティングでは、各種センサを携帯し個人が経験した対象を記録可能とする。さらに、情報家電を利用することで日常生活で利用する家電製品を操作する記録をデー

タとして取得可能となる。

一方、従来実世界で行ってきた活動が仮想空間上での実行可能とすることにより、ライフログを取得することが考えられる。近年、インターネットが普及し Web 上で様々なサービスが提供されるようになった。オンラインショッピングや E ラーニング等、実世界で行ってきた日常生活や業務上の活動の一部分を仮想空間上で遂行できるようになってきた。仮想空間は電子的に実装されているため、仮想空間上での活動の記録データを取得可能である。

ライフログを管理するためには、大容量の二次記憶が必要である。近年、HDD の大容量化と低価格化は着実に進んでおり、個人の経験に関する記録を蓄積する HDD を個人で保有することは非現実的ではなくなっている。

このように、ライフログを取得するための環境は着実に整備されつつあり、取得したライフログを蓄積する記憶媒体を個人が保有することも現実的になりつつある。しかし、ライフログを効率的に管理し効果的に活用するための一般的な枠組みに関して十分な研究が行われているとはいえない。本研究では汎用性を有するライフログ検索モデルの開発を目標とする。

本論文ではライフログ検索の特徴として、コンテキスト

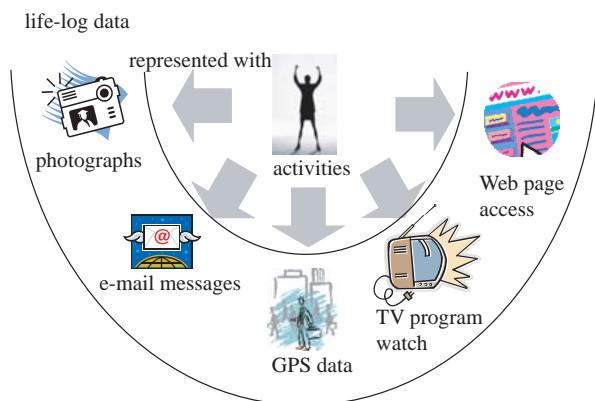


図1 ライフログの概念図

表1 ライフログを構成するメディアの分類と例

分類	例
記録	カメラ付き携帯電話で撮影した画像, ウェアラブルカメラで撮影したビデオ
コンテンツ利用ログ	TV 番組視聴履歴, Web アクセスログ, プレーヤ再生楽曲
コミュニケーションログ	電子メールメッセージ, 電話通話記録
センサデータ	GPS データ, 気温
生体情報	脳波, 血圧

に基づいた検索, および不確実性に基づいた検索が存在することを示し, それらを実現するための検索モデルの基本的な枠組みを提案する.

2 ライフログを利用した検索

ライフログは複数の形態のデータから構成される. ライフログを構成するデータをメディアと呼ぶ. 複数のメディアが同一の活動の異なる側面や特徴を表現している. ライフログと活動の関係を表す概念図を図1示し, ライフログを構成するメディアの例を表1に示す.

本研究で対象とする検索は, 利用者が指定した検索要求に基づいて, 特定の人物に関するライフログ集合の中から指定した種類のメディアを検索するものである. 本節では, 我々が提案するライフログ検索の特徴について述べる.

2.1 コンテキストに基づく検索

画像, ビデオ, テキストなどのメディアに対して, 付随するメタデータを利用せず, メディア自体の特徴を利用する検索は, 内容に基づく検索 (content-based retrieval) と呼ばれる. 内容に基づく検索ではメディアの種類によって表現可能な特徴と表現不可能な特徴が存在する.

利用者は検索対象となるメディアが直接表現していない特徴に基づいた検索を希望する場合がある. 例として, デジタルカメラで撮影した記録画像の検索を考える. 記

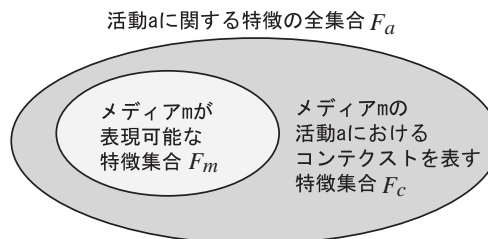


図2 コンテキストの概念図

録画像を検索する際には, 検索キーとして撮影した日の場所や天候を指定する要求が想定できる. しかし, 記録画像から記録場所や天候を画像的な特徴から判断することは困難であるため, 通常の内容検索ではそうした検索要求に対応できない.

ライフログでは同一人物の活動を記録している点から, 異なるメディアを意味的に関連付け可能である. この関連付けを活動に基づくメディア統合と呼ぶ. メディア統合の対象となるメディアは同一の種類に属することもあるし, 異なる種類に属することもある.

メディア統合により検索対象が直接表現することができない特徴を利用して検索を行うことを考える. いま, 活動 a が有する総ての特徴から構成される集合を F_a とする. 活動 a を表すメディア m が表現する特徴の集合を F_m とする. $F_c = F_a - F_m$ となる F_c を F_m の活動 a に対するコンテキストと呼ぶ. コンテキストの定義を表す概念図を図2に示す.

メディアによって直接表現されていない内容に基づく検索をコンテキストに基づく検索 (context-based retrieval) と呼ぶ. ライフログ上でメディア統合を行うことによりコンテキストに基づく検索を実現できる. コンテキストに基づいた検索の概要を表す概念図を図3に示す. 図ではメディア2を検索する為に, メディア1の特徴を利用する様子を表している. 記録画像に対する撮影場所や天候などはコンテキストであると考えられる.

2.2 不確実性に基づく検索

メディアを検索する際に不確実性が存在することがある. 例えば, 記録画像に対して画像処理によってオブジェクト抽出や分類を実行した場合には不確実性を含むことが多い. また, 文書をキーワードで検索する場合も, 一つのキーワードが含む文書が多数存在する場合があり, tf や idf 等の特徴量を利用して文書毎にキーワードの重みを変化させることが多い.

ライフログ検索では, 同一の活動が複数のメディアで表現され, 利用者が与えた質問に対応する概念が異なるメディアの特徴として表現されることがある. このような場合, 単一の概念で表された特徴は不確実性が高い場合でも, 複数のメディアで表される特徴を統合的に解釈することで確実性を向上させることができる. この検索の概念図を図4に示す. 図では利用者が与えた質問に対応する概念がメ

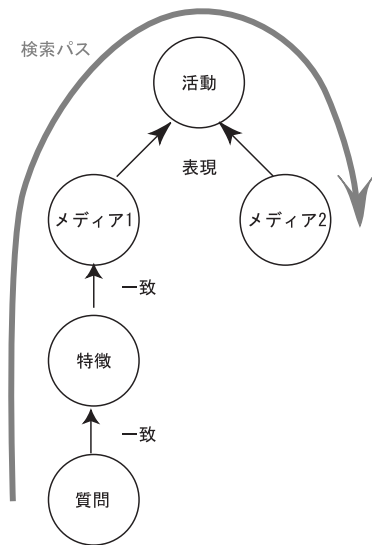


図3 コンテキストに基づく検索

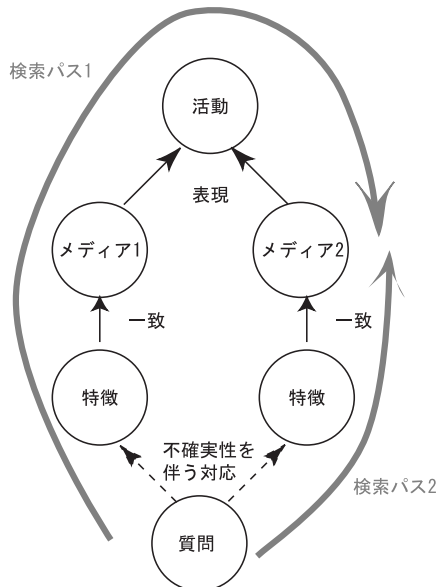


図4 不確実性を伴う特徴の統合的な解釈に基づく検索

メディア1とメディア2の異なる概念によって表現され、メディア2の適合度は検索パス1と検索パス2の2個の経路に基づく評価を統合的に利用することを示している。

2.3 ライフログ検索の応用例

ライフログに対して、コンテキストに基づく検索や不確実性を伴う検索を利用することで、以下に示すような応用例に適用することが考えられる。

- デジタルカメラで撮影した画像を人手によるアンテーションを付与せずにキーワードで検索する。
- 指定した文書を作成した際に参考にしたWebページを取得する。
- 指定した状況や目的において再生した楽曲を類似し

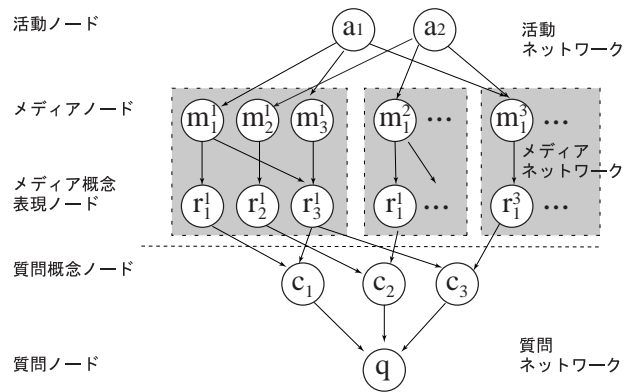


図5 ライフログネットワークの概念図

た状況で再生する。

3 ライフログ検索モデル

本論文で提案するライフログ検索モデルは、推論ネットワーク [4] を拡張したものである。推論ネットワークはベイジアンネットワーク [5] に基づく代表的な文書検索モデルの一つである。

ベイジアンネットワークは確率理論に基づいた不確実性を含む推論を行う枠組みである。ベイジアンネットワークでは命題がノードとして表され、ノード間の依存関係がエッジで表される。依存関係は条件付き確率で表現される。命題に対する確率は、命題に関する信念の強さを表す。親ノードを前提条件とする子ノードの条件付き確率の確率分布は、子ノードに与えられた条件付き確率表 (CPT) で管理される。親以外のノードとの間の依存関係もベイズの定理を利用することによって、いくつかのノードが管理するCPTを利用して導出可能である。

3.1 データ構造

ライフログを検索するために利用するライフログネットワークの概念図を図5に示す。ライフログネットワークはベイジアンネットワークであり、図中の円がノードを表し、矢印は依存関係を表している。

3.1.1 ノードの種類

ライフログ検索モデルは以下の5種類のノードから構成される。

活動ノード 利用者が実世界で行った活動に対応付けられる。「利用者が活動 a を (検索対象として) 興味を持つ」という命題を表す。

メディアノード 活動を記録したメディアに対応する。「利用者がメディア m に対して (検索対象として) 興味を持つ」という命題を表す。例えば、デジタルカメラで撮影された1枚の記録画像、1通の電子メールメッセージは1個のメディアノードとして表現される。

活動ノードとメディアノードの依存関係は条件付き確率 $P(m|a)$ によって表現される。 $P(m|a)$ は、「利用者が活動

a に注目している場合に、メディア m がその内容を記録している」という命題に対する信念の強さを表す。

メディアノードは型に分類される。メディアノードの親ノードは表現対象となる活動ノードである。複数のメディアノードが同一の活動ノードを親とすることが許されるため、同一の活動が複数のメディアで表現される構造を構成可能である。

メディア概念表現ノード メディアノードの特徴を表す概念に対応する。「利用者がメディアで表現される概念 r に注目する」という命題を表す。メディアノードに対する索引としての役割を有し、文書検索における索引語に相当する。索引付けの対象となるメディアノードが親ノードとなる。

メディアノードとメディア概念表現ノードの依存関係は条件付き確率 $P(r|m)$ によって表現される。 $P(r|m)$ は、「利用者がメディア m に注目している場合に、メディア概念 r に注目する」という命題に対する信念の強さを表す。

メディアノードと同様に、メディア概念表現ノードは型に分類される。同一のメディア概念表現型に属するノードの親ノードは総て同一のメディア型に属する。メディア概念表現型はメディア型に依存するため、同一の概念であっても、異なるメディア型に属する概念は異なるメディア概念ノードとして表現される。

質問概念ノード 利用者の検索要求を構成する概念に対応する。「利用者が質問を構成する概念として c を指定する」という命題を表す。文書検索における検索語に相当する。同一の質問概念ノードが、複数のメディア概念表現ノードを親ノードとすることがある。これは、同一の概念が複数のメディア型によって表現されることに相当する。

メディア概念表現ノードと質問概念ノードの依存関係は条件付き確率 $P(c|r)$ によって表現される。 $P(c|r)$ は、「利用者がメディア概念 r に注目している場合に、質問概念 c を指定する」という命題に対する信念の強さを表す。

質問ノード 利用者の検索要求に対応する。「利用者が検索要求を q とする」という命題を表す。質問ノードを構成する質問概念を親ノードとして有する。複数の親ノードを有するため、複数の質問概念を利用した複合的な検索要求を表現可能である。

質問概念ノードと質問ノードの依存関係は条件付き確率 $P(q|c)$ によって表現される。 $P(q|c)$ は、「利用者が質問概念 c を指定した場合に、検索要求 q を有している」という命題に対する信念の強さを表す。

3.1.2 部分ネットワーク

ライフログネットワークは活動ネットワークと質問ネットワークから構成される。

活動ネットワークは、活動ノード、メディアノード、メディア概念表現ノードから構成される。活動ネットワークは対象とする人物の活動が、どのメディアでどのように表現されているかを示している。

メディア概念表現ノードは同じ型のメディアノードのみ

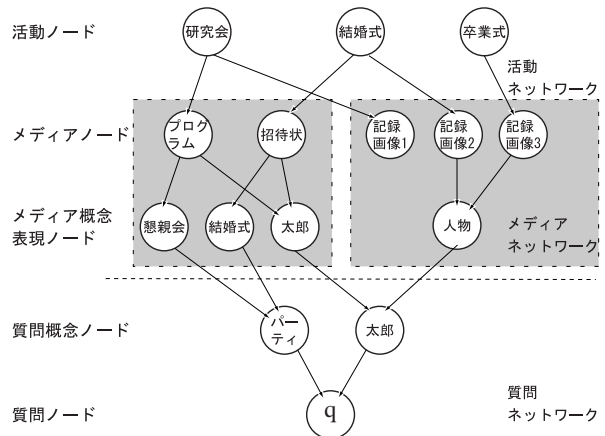


図6 ライフログネットワークの例

を親とする。同じ型に属するメディアノードとメディア概念表現ノードから構成されるネットワークをメディアネットワークと呼ぶ。活動ネットワークはメディア型と同数のメディアネットワークを含む。

質問ネットワークは質問概念ノード、質問ノードから構成される。質問ネットワークは、利用者の検索要求が表現する内容を示している。

ライフログネットワークの具体例を図6に示す。例は記録画像と電子メールの2種類のメディア型から構成されている。右側のメディアネットワークが記録画像であり、左側のメディアネットワークが電子メールである。

3.2 検索

検索時には利用者は質問ネットワークと検索対象となるメディア型を指定する。いま、検索対象となるメディア型が M^t であるとし、 M^t に所属するメディアノードの集合が $M^t = \{m_1^t, \dots, m_n^t\}$ であるとする。このとき、メディアノードと質問ノード q の間の依存関係によって指定されたメディア型に含まれる集合を検索結果としてランク付けする。 M^t に属する全ての要素に対して条件付き確率 $P(q|m_i^t = \text{ture})$ を計算することにより、条件付き確率の値が大きいほど質問に対する適合度が高いと考える。これは、利用者がメディア m_i^t を検索結果とした場合に質問が q であると考えた信念の強さを表し、 m_i^t の内容を q がどの程度表現しているかの度合いを表していると解釈できる。

4 活動ネットワークの構成

活動ネットワークはメディア間の相互関係とインデックス構造を表し、利用者の質問とは独立に存在する。外部からシステムに与えられるのはメディアとメディア概念表現ノードの一部分である。外部から与えられないノードと依存関係は与えられたデータに基づいてシステムが自動的に構成する必要がある。

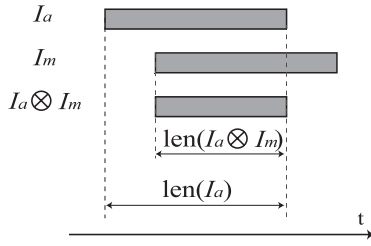


図7 時間区間の関係

4.1 活動ノードとメディアノードの依存関係

本研究で活動とは「特定の人物が時間区間及び空間領域を占有する一連の状態」であると考えられる。本研究ではメディアの時間的な特徴及び空間的な特徴を利用して依存関係を考える。活動に関する時間または空間に関する特性を表現可能なメディアのみを対象とする。

活動はメディアによって表現される対象である。実世界には様々な活動が存在する。全ての活動が一般的に有する特徴は時間区間と空間領域を占有するという点である。

活動ノード a とメディアノード m の依存関係は「利用者が活動 a に注目している場合に、メディア m がその内容を記録している」と考えられる信念の強さを表す。この依存関係を条件付き確率 $P(m|a)$ として表現する。

依存関係を導出する基本的なアプローチを説明するために、まず、時間的な特徴のみを考える。活動は時間的な区間を占有する。メディアは実世界における特定の時間区間での活動の様子を記録している。活動 a が存在した時間区間を I_a とし、メディア m が記録している時間区間を I_m とする。このとき、 I_a と I_m の共通区間を $I_a \otimes I_m$ と表記する。時間区間の例を図7に示す。条件付き確率の値を以下のように定義する。

$$P(m|a) = \frac{\text{len}(I_a \otimes I_m)}{\text{len}(I_a)} \quad (1)$$

ここで $\text{len}(I)$ は時間区間 I の長さを表す。 $P(m|a)$ は、利用者が活動が存在した時間区間に対して均一的に興味を持っている場合、メディアに記録されている時間区間の割合を表している。例えば、1時間の活動 a の一部分を撮影した30分の映像 m が存在したとき、 a に興味がある利用者が m で要求を満足すると考える信念の強さは0.5である。

次に、空間的な特徴を考えるため、時間区間を空間的な領域に拡張する。活動 a の時間と空間に関する存在を表す関数 $f_a(n, e, t)$ を導入し、これを存在関数と呼ぶ。 n は緯度(北緯)、 e は経度(東経)、 t は時刻を表す。 $f_a(n, e, t)$ は時刻 t において (n, e) において活動 a が行われている場合1を返し、それ以外の場合は0を返す。一方、メディア m の記録対象を表す関数を記録関数と呼び、 $g_m(n, e, t)$ と表記する。

いま、活動ノードとメディアノードが存し、それぞれに対して存在関数と記録関数が与えられているとする。活動ノード a とメディアノード m 間の依存関係を表す条件付

き確率 $P(m|a)$ を、存在関数と記録関数の共通部分の大きさとして、以下のように定義する。

$$P(m|a) = \frac{\int_0^\infty \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \int_{-\pi}^\pi f_a(n, e, t) g_m(n, e, t) dndedtdt}{\int_0^\infty \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \int_{-\pi}^\pi f_a(n, e, t) dndedtdt} \quad (2)$$

対象とするメディアが、時間的な特徴と空間的な特徴のどちらか一方の特徴しか有さない場合がある。時間的な特徴だけが抽出できる場合には、時間的な記録のみを表す関数 $g_m^t(t)$ を用いて

$$g_m(n, e, t) = g_m^t(t) \quad (3)$$

として計算を行う。同様に空間的な特徴のみを有するメディアに関しては空間上の記録を表す関数 $g_m^s(n, e)$ を利用して計算する。

活動ノードとメディアノードに存在関数が与えられていれば依存関係を推定できるため、活動ノードとメディアノード間の依存関係を決定する問題は、それぞれのノードに割り当てる存在関数と記録関数を推定する問題になる。

4.2 メディアノードに関する記録関数推定

記録関数の推定方法はメディアに依存する。本論文では代表的なメディア形式の例として、GPSデータ、カメラ付き携帯電話やデジタルカメラで撮影された記録画像、利用者が送受信した電子メールメッセージを用いる。

以下に、それぞれのメディアにおける存在関数の推定方法について述べる。

4.2.1 GPSデータ

GPSデータは緯度 n_i 、経度 e_i 、時刻 t_i から構成されるレコード (n_i, e_i, t_i) の系列である。GPSレコードはサンプリング周期を s 毎に取得される。GPS観測結果には測定誤差 d_e が含まれる。レコード (n_i, e_i, t_i) が表す記録関数を以下のように推定する。

$$g_i^{GPS}(n, e, t) = \begin{cases} 1 & t_i - \frac{s}{2} \leq t < t_i + \frac{s}{2} \text{かつ} \\ & (n, e) \text{ と } (n_i, e_i) \text{ 間の距離 } \leq d_e \\ 0 & \text{その他} \end{cases}$$

GPSメディアノードはGPSレコードに基づいて構成される。原理的には、個々のGPSレコードを個別のメディアノードとすることが考えられる。しかし、個々のGPSレコードを単位とするとメディアノードの数が膨大になり現実的ではない。例えば、地名、人名、出来事名などをキーワードに記録画像を検索するアプリケーションを想定する場合には、複数のGPSレコードを集約化しても問題がないと考えられる。単位時間に含まれるデータを単一のメディアノードに集約する。集約されたメディアノードの記録関数は、構成データの記録関数の和として導出する。

4.2.2 記録画像

記録画像から時間的な情報として撮影時刻を獲得することができる。しかし、空間的な情報を獲得することは困難であるため、記録画像に対しては時間に関する記録関数の

みを考える．撮影時刻が t_i である記録画像 m_i に関する時間概念を表す記録関数を以下のように推定する．

$$g_{m_i}^{Image}(t) = \begin{cases} 1 & t_i - \frac{d}{2} \leq t < t_i + \frac{d}{2} \\ 0 & \text{その他} \end{cases} \quad (5)$$

d は 1 個の記録画像が表現可能な時間的な幅を表す．厳密には， d は撮影機材で連続撮影可能な最小時間となるが，スナップ写真を検索するような場合には数十秒から数分として考えるのが適切であると予想される．

記録画像以外にも，記録時刻がタイムスタンプとして保持されるデータが多数存在する．例えば，心電図のデータ等である．それらに対して基本的には同様な手法で記録関数を推定可能である．

4.2.3 電子メールメッセージ

電子メールは一般に普及し，予定の調整や報告などに広く利用されている．電子メールメッセージには送信者や受信者の活動に関する有用な情報が多く含まれている可能性が高い．

簡単化のために，メッセージに含まれている時間概念と空間概念の間に依存関係が存在しないと仮定する．時間に関する記録関数 $g_m^{email}(t)$ と空間に関する記録関数 $g_m^{email}(n, e)$ は直交する．時間と空間を表す記録関数 $g_m^{email}(n, e, t)$ を以下のように定義できる．

$$g_m^{email}(n, e, t) = g_m^{email}(t) \times g_m^{email}(n, e) \quad (6)$$

以下に時間に関する記録関数と空間に関する記録関数の推定方法を述べる．

時間に関する記録関数 電子メールメッセージから抽出可能な時間表現には，送受信等のコンピュータ上の処理時刻を示す時間表現とメッセージの内容として本文に含まれる時間表現がある．送信時刻は送信者がコンピュータ等で電子メールを送信するという活動が終了した時刻を記録している．一方，本文には様々な種類の時間表現が含まれている．メッセージ中の時間表現は自然言語処理によって抽出する．

時間表現が表す概念は時間区間と時間点に分類できる．日常的な時間表現では時間軸上の点を指すように見える表現が区間を表していることが多い．例えば，「2004年10月1日」という表現は，「2004年10月1日」が始まってから終わるまでの時間区間として考えられる．このように，電子メールメッセージ本文から抽出された時間点表現は全て時間区間として扱う．

いま，メッセージ m に含まれる時間区間集合を $T_m = \{I_1, \dots, I_n\}$ とする．時間区間 I_i に関する存在関数を以下のように定義する．

$$g_m^{email}(t) = \begin{cases} 1 & \text{begin}(I_i) \leq t < \text{end}(I_i) \\ & \text{を満足する } I_i \text{ が } T_m \text{ 内に存在する} \\ 0 & \text{その他} \end{cases} \quad (7)$$

ここで， $\text{begin}(I_i)$ は I_i の開始時刻， $\text{end}(I_i)$ は I_i の終了時刻を表す．

空間に関する記録関数 空間を表す記録関数について考える．メッセージに含まれる地名等の地理概念の表現を自然言語処理で抽出する．地理概念とは，地名，ランドマーク名など実世界上の領域に対応付けられる概念である．単一のメッセージ内に複数の地理概念が含まれることがある．抽出された全ての地理概念は辞書を用いて対応する領域に変換する．得られた領域集合を $S_m = \{s_1, \dots, s_n\}$ とするとき，メッセージ全体の空間に関する記録関数 $g_m^{email}(n, e)$ を以下の式で定義する．

$$g_m^{email}(n, e) = \begin{cases} 1 & (n, e) \text{ が含まれる領域 } s_i \text{ が} \\ & S_m \text{ 内に存在する} \\ 0 & \text{その他} \end{cases} \quad (8)$$

4.3 活動ノードに関する存在関数推定

あらかじめシステム内に活動ノードは存在しない．システムに与えられるのはメディアノードのみである．そこで，利用者によって与えられたメディアノードに基づいて活動ノードを導出する．

メディアが活動の時間・空間的な特徴表現の正確さはメディアの種類によって異なる．GPS データと電子メールメッセージはともに時間的特徴と空間的特徴を表現している．GPS データから得られる特徴は活動を正確に反映している可能性が高い．電子メールメッセージに含まれる時間的特徴と空間的特徴は，利用者の活動を表していない対象が含まれていることが多い．

GPS データから得られる特徴が正確に活動を表現していると仮定する．これは GPS データを表すメディアノード m_i に対して以下の条件を満足するような活動ノード a_i を生成することである．

$$\begin{aligned} P(m_i|a_i) = 1 \text{ かつ } P(\neg m_i|a_i) = 0 \text{ かつ} \\ P(m_i|\neg a_i) = 0 \text{ かつ } P(\neg m_i|\neg a_i) = 1 \end{aligned} \quad (9)$$

a_i の存在関数として m_i の記録関数を割り当てることにより，上記の条件を満足する a_i を導出可能である．

GPS データは常に取得できるとは限らない．GPS データが存在しない時間区間では，単位時間で分割した時間区間を個別の活動ノードとして利用する．

4.4 メディア概念ノードの構成

メディア概念ノードはメディアノードの索引となる概念を表す．メディア概念ノードはメディアの型に依存している．以下に，それぞれのメディア型に関する概念表現ノードの構成方法を示す．

4.4.1 GPS データ型の概念表現ノード

GPS データ型のメディア概念ノードは，地理概念を表す．地理概念は，行政区域，公共施設，ランドマーク等である．ノードとなる地理概念は電子地図から抽出する．

地理概念は実世界上の地理的対象に関する表現であると考えられるので，地図上の領域に基づいて空間的な記録関数を考えることができる．GPS データ型のメディアノード

には時空間的な記録関数が与えられているので、活動ノードとメディアノード間の条件付き確率を推定した場合と同様に、メディア概念ノードとメディアノードに付与された記録関数の共通領域の大きさにより依存関係を決定することができる。

4.4.2 記録画像型の概念表現ノード

記録画像に対しては画像処理によって判断可能な概念がメディア概念ノードとなる。

一般的に画像に撮影されている対象を理解することは困難である。しかし、対象を限定することにより利用可能となる場合もある。例えば、画像中の肌色情報を利用して未学習の人物を抽出する研究がなされている [6]。このような技術を利用して画像中の肌色領域の割合を人物が撮影されている確率として人物が撮影されているかどうかを判断するための処理に利用することが考えられる。

4.4.3 電子メールメッセージ型の概念表現ノード

電子メールメッセージからテキスト解析によってメディア概念ノードを作成する。電子メールメッセージに関するメディア概念ノードはヘッダ情報から得られる情報とメッセージの本文から得られる情報がある。

ヘッダからはメッセージの投稿時刻、送信者のメールアドレス、表題等の情報を取得できる。これらは個別のメディア概念ノードとする。また、メッセージの本文からは字句解析を行い得られた名詞を抽出する。それらを個別のメディア概念ノードとする。

概念表現ノードの条件付き確率を求める場合には、 tf 値や idf 値といったテキスト検索で利用される特徴量を利用することが考えられる。概念表現ノードは 1 個の索引語を表すとする。メッセージ m_i に関するメディア概念ノード r の重みを表す条件付き確率を推定するために、推論ネットワークに基づく文書検索システム INQUERY [7] で索引語の重み付けに利用された以下の式を利用する。

$$p(r|m_i) = 0.4 + 0.6 \times \frac{tf_{r,m_i}}{tf_{r,m_i} + 0.5 + 1.5 \frac{\text{length}(m_i)}{\text{avg len}}} \times \frac{\log \frac{N+0.5}{n_r}}{\log N + 1} \quad (10)$$

ここで n_r は r を含むメッセージの数であり、 N はメッセージ数であり、 avg len はメッセージ中の単語数であり、 tf_{r,m_i} はメッセージ m_i 中で r が出現した回数である。

5 質問ネットワークの構成

質問ネットワークは利用者の検索要求の表現であるため、利用者の検索要求に基づいて動的に生成する必要がある。

質問概念はキーワード検索における検索語に相当する。質問概念では文字列から構成されるキーワードの他にも、例となる画像や地理上の領域などいくつかの対象が考えられる。質問概念ノードは利用者が指定した検索要求にもとづいて動的に生成される。

本研究では文字列から構成されるキーワードに対象を限

定する。個々の質問概念ノードが個別のキーワードを表す。

電子メールメッセージ型のメディア概念表現ノードと質問概念ノードとの依存関係は単純なキーワードマッチングによって決定する。一致した場合に依存関係が存在すると考え、それ以外の場合は存在しないと考える。シソーラスを利用することで「結婚式」を「パーティ」に対応づけるような対応付けも可能である。

GPS データ型のメディア概念表現ノードとの関係も同様に地理的な概念を表すキーワードに関するマッチングによって決定する。

記録画像型のメディア概念表現ノードとの対応付けは自動的に行うことは難しい。電子メールメッセージをはじめとするテキストメディアではメディア概念ノードを自動的に生成可能である。記録画像のメディア概念表現ノードは画像処理方法を管理者が個別に設定する必要がある。メディア概念表現ノードを付与した際に、対応付け可能なキーワードを指定しておく。

6 関連研究

ウェアラブルカメラで撮影された映像を対象として、異種メディアのデータを利用して自動的にインデキシングし、高度な検索やアクセスを実現することを目指していくつかの手法が提案されている。上田ら [8] は GPS で得られた空間情報とウェアラブルカメラで撮影した映像を統合化し、地理情報を表すキーワードで映像を検索するための方法を提案している。相澤ら [9] はウェアラブルカメラで記録した映像と脳波データを統合化することにより、映像のダイジェスト化を行う手法を提案している。これらは、時間に基づいて映像に対して異種メディアの対応付けを行っているが、不確実性を伴う対応付けについては考えていない。

個人データの管理のための機構として、Freeman [10] らはファイルをデスクトップメタファではなく時系列によって整理する手法を提案している。Stuff I've Seen [11] は個人が読んだ電子メールやウェブページなどをテキスト情報に基づいて統一的にインデックス付けして検索可能としている。タイムマシンコンピューティング [12] では異種メディアデータを時間軸上の関係を様々な視覚な形式で閲覧する機構を提案している。MyLifeBits [13] では、写真、電子メールなどを様々な情報を実際にデジタル化し異種メディアデータを関係づける数種類のリンクを提案している。これらの研究では、複数の不確実な情報を統合して対象を検索することは考えられていない。

7 おわりに

本論文では、ライフログ検索を目的として、ページネットワークに基づく検索モデルを提案した。本モデルはすべてのメディアをノードとして表現し、ノードの特徴は活動ノードとメディア表現ノード間との依存関係として表現されるため、異種メディアに対して統一的な構造化を実現できる。次に、GPS レコード、記録画像、電子メールメッ

ページを例として、異種メディアが記録している対象の時間的な特徴と空間的な特徴に基づいて検索に必要なノード間の依存関係を推定する手法を提案した。他のメディアに対してもメディア認識処理により抽出できる概念をメディア概念表現ノードとして実現することにより、基本的に同様な方針で検索に利用できると考えている。

現在、プロトタイプシステムの実装を行っている。実装が完了した後に、評価実験により本手法の有効性を評価する。

本論文で提案した手法では、メディアノードの表現対象に関する不確実性を考慮していない。電子メールメッセージに自然言語で記述された時間概念や空間概念は曖昧さを多く含むことが予想される。不確実性を表現可能となるように記録関数を拡張することは今後の課題である。

本研究の発展として、信頼できるコミュニティでの情報共有を目的として、分散して存在する他人のライフログのアクセスを実現する機能を検討する予定である。この場合、他人のコンピュータで管理されるメディアネットワークを自分のライフログネットワークに組み込む仕組みが必要であると考えている。ライフログはプライバシーデータであり、厳重かつ柔軟なアクセス管理が必要である。そのために利用者との人間関係のモデル化と、それに基づく活動内容に基づくメディアデータのアクセス制御機構を開発することが重要であると考えている。

謝辞

日頃からご指導頂いている九州大学藤村直美教授、並びに本研究に関して熱心に議論して下さった名古屋大学渡邊研究室の皆様へ感謝致します。本研究の一部は、科学研究費補助金（課題番号 15500073, 16700107）の助成をうけて行われた。ここに記して謝意を表します。

参考文献

- [1] Bush, V.: As We May Think, *Atlantic Monthly*, pp. 101–108 (1945).
- [2] Gray, J.: What Next? A Dozen Information-Technology Research Goals, *Journal of the ACM*, Vol. 50, No. 1, pp. 41–57 (2003).
- [3] Croft, W. B.: Information Retrieval and Computer Science: An Evolving Relationship, in *Proc. of Int'l Conf. of SIGIR'03*, pp. 2–3 (2003).
- [4] Turtle, H. and Croft, W. B.: Evaluation of an Inference Network-Based Retrieval Model, *ACM Transactions of Information Systems*, Vol. 9, No. 3, pp. 187–222 (1991).
- [5] Pearl, J.: *Probabilistic Reasoning in Intelligent Systems*, Morgan Kaufmann (1988).
- [6] 松村朱里, 岩井儀雄, 谷内田正彦: 肌色情報を用いた複数人物追跡, 情報処理学会 CVIM 研究会報告, No. 133, pp. 133–138 (2002).
- [7] Allan, J., Connell, M. E., Croft, W. B., Fang-Fang Feng, D. F. and Li, X.: INQUERY and TREC-9, in *Proc. of TREC-9* (2001).
- [8] 上田隆正, 天笠俊之, 吉川正俊, 植村俊亮: 位置情報と地理情報を用いた映像データの索引付け手法, 第12回データ工学ワークショップ (DEWS2001) (2001).
- [9] 相澤清晴, 石島健一郎, 椎名誠: ウェアラブル映像の構造化と要約: 個人の主観を考慮した要約生成の試み, 信学論, Vol. J86-D-II, No. 6, pp. 807–815 (2003).
- [10] Freeman, E. and Gelernter, D.: Lifestreams: A Storage Model for Personal Data, *SIGMOD Record*, Vol. 25, No. 1, pp. 80–86 (1996).
- [11] Dumais, S., Cutrell, E., Cadiz, J., Janck, G., Strin, R. and Robbins, D. C.: Stuff I've Seen: A System for Personal Information Retrieval and Re-Use, in *Proc. of SIGIR'03*, pp. 72–79 (2003).
- [12] Rekimoto, J.: Time-Machine Computing: A Time-centric Approach for the Information Environment, in *Proc. of UIST'99*, pp. 45–54 (1999).
- [13] Gemmell, J., Bell, G., Lueder, R., Drucker, S. and Wong, C.: MyLifeBits: Fulfilling the Memex Vision, in *Proc of ACM Multimedia'02*, pp. 235–238 (2002).