

ライフログ検索のための異種メディア統合 Media Integration for Life-log Search

牛尼 剛聡[†]
Taketoshi USHIAMA

渡邊 豊英[‡]
Toyohide WATANABE

1. はじめに

近年、ユビキタスコンピューティングおよびウェアラブルコンピューティングに関する技術の発達により、実世界における個人の活動の様子をデジタルデータとして取得することが可能となった。また、Web を中心とした IT 技術の進展と普及により、コンピュータ上で日常生活や業務における様々な活動を電子的に遂行可能となった。これらを背景として、個人の経験をデジタルデータとして記録するための環境が整いつつある。個人の経験を記録したデータはライフログと呼ばれる。近年、利用者のすべてのライフログを蓄積し効果的に管理するシステムが注目されており、MyLifeBits[1] をはじめとして活発に研究されている。

ライフログの特徴として、利用者の活動が複数のメディアによって異なる側面から記録されることがあげられる。すなわち、ライフログを構成する異種メディアのデータは、同一人物の活動を記録しているという点から意味的な関連付けが可能となる。この意味的な関連付けをライフログにおけるメディア統合と呼ぶ。メディア統合により、複数のメディアの特徴に基いて相補的にデータの内容を判断することが可能となり、単一のメディアでは表現が困難な高度な検索を実現できる。例えば、利用者が送受信した e-mail メッセージの記録や GPS により取得した位置情報を利用して、デジタルカメラで撮影した記録画像を、人手によるアノテーションなしで、連想的なキーワードにより検索可能となる。

ライフログ上でのメディア統合において問題となるのは、ライフログを構成する多種多様なメディアデータ間の意味的な関係を厳密に決定することは困難であり、関連付けに存在する不確実性を適切に表現し検索に反映する必要があることである。本論文では、筆者らが提案しているライフログ検索モデル [2] において、異種メディア間の対応付けの不確実性を反映したメディア統合を実現する手法を提案する。

2. ライフログ検索モデル

本研究で利用するライフログ検索モデル [2] の概念図を図 1 に示す。本モデルは、ベイジアンネットワークに基づく代表的な文書検索モデルの一つである推論ネットワーク [3] を拡張したものである。

ライフログ検索モデルは活動ネットワーク (activity network) と質問ネットワーク (query network) 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、どのメディアのデータでどのように表現されているかを示している。質問ネットワークは、利用者の検索要求が表現している内容を示している。

活動ネットワークは、活動 (activity) ノード a_i 、メディア (media) ノード m_j 、メディア概念表現 (media concept representation) ノード r_k から構成される。活動ノードは利用者の活動を表す。メディアノードは活動を表現する様々な種類のメディアデータを表す。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 e-mail メッセージを 1 個のメディアノードとして表現可能である。メディア概念表現ノードはメディアノードによって表現される概念を表す。文書検索における索引語に相当する。例えば、e-mail メッセージに対するキーワードはメディア概念表現ノードとして表現可能である。

メディアノードとメディア概念表現ノードは型によって分類される。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を親として有する。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 (media network) と呼ぶ。

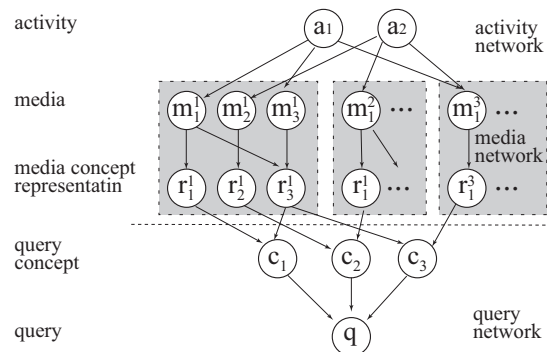


図 1 ライフログ検索モデル

質問ネットワークは質問概念 (query concept) ノード c 、質問 (query) ノード q から構成される。質問表現ノードは利用者が検索に利用する概念を表す。従来の検索システムにおける検索語に相当する。質問ノードは利用者の質問を表す。

図中の円がノードを表し、矢印は依存関係を表している。依存関係の強さは条件付き確率として与え、各ノードに CPT (条件付き確率表) を割り当てることで実現する。

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、 m_i^t の検索要求 q に対する適合度は条件付き確率 $P(q|m_i^t = \text{ture})$ により与えられる。この確率は各ノードの CPT を利用して計算可能である。 M^t に属する全ての要素に対して $P(q|m_i^t = \text{ture})$ を計算することにより、検索結果のランク付けを行う。

3. メディア統合

ライフログ検索モデルでは、活動ノードが複数の子 (メディアノード) を有することができる。これは、単一の活動が複数のメディアデータで表現されることに対応し、活動を介してメディアデータが意味的に関連付けられ統合化されることを意味する。本モデルでは依存関係を持つノード間の条件付き確率を推定し、ノードに CPT を割り当てることにより、検索を実行可能となる。ここで、メディア統合はメディアノードの CPT の決定という処理に還元されることになる。以下に、活動ノードとメディアノードの依存関係を決定するための基本的なアプローチを示す。

人物 p の活動が a であるとき、 p を a の主体であるという。ライフログは利用者を主体とする活動に関する様々な記録 (メディアノード) を集めたものである。本研究では、活動とは「特定の人物おける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、このように考えることにより、時間と空間に基づいたメディア統合を実現できる。

本手法では、すべてのアクティビティとメディアに対して時間・空間に関する確率密度関数を与える。

活動 a の存在を表す確率密度関数を $f_a(x, y, t)$ とする。 x は緯度、 y は経度、 t は時刻を表す。同様に、メディア m の表現対象を表す確率密度関数を $f_m(x, y, t)$ とする。

活動 a とメディア m に確率密度関数が与えられているとき、条件付き確率 $P(m|a)$ を、活動とメッセージの確率密度関数の共通領域の体積として、以下のように定義する。

$$P(m|a) = \int_0^\infty \int_{-p_i}^\pi \int_{-\pi}^\pi \min(f_m(x, y, t), f_a(x, y, t)) dx dy dt$$

[†] 九州大学, Kyushu University

[‡] 名古屋大学, Nagoya University

メディアの特徴に基づいて確率分布関数を推定する。メディアの特徴によっては、時間的な特徴や空間的な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、時間の分布のみを与える確率密度関数 $f_m^t(t)$ を考える。このとき、条件付き確率は以下のように計算する。

$$P(m|a) = \int_0^\infty \min(f_m(t), \int_{-pi}^\pi \int_{-pi}^\pi f_a(x, y, t) dx dy) dt$$

空間的な特徴のみを有するメディアに関しても、空間の分布のみを与える確率密度関数 $f_m^s(x, y)$ を考え、同様に計算するものとする。

4. メディアノードに関する確率密度関数推定

確率密度関数の推定方式はメディアの特徴によって異なる。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 e-mail メッセージを用いて、確率密度関数の推定方法について述べる。

4.1 GPS データ

GPS データは時刻 t_i 、緯度 x_i 、経度 y_i からなる組 (t_i, x_i, y_i) の系列である。GPS の測定誤差を e 、サンプリング周期を s とするとき、データ (t_i, x_i, y_i) が表す確率密度関数を以下のように推定する。

$$f_i^{GPS}(x, y, t) = \begin{cases} \frac{1}{s \times \pi e^2} & t_i - \frac{s}{2} \leq t < t_i + \frac{s}{2} \text{かつ} \\ & (x, y) \text{ と } (x_i, y_i) \text{ 間の距離} \leq e \\ 0 & \text{その他} \end{cases}$$

個々の GPS データを個別のメディアノードとするのは現実的でないため、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確率密度関数は、構成データの確率密度関数の平均として導出する。

4.2 記録画像

デジタルカメラ等で撮影された記録画像から時間概念として撮影時刻を獲得することができるが、空間的な情報を獲得することは困難である。そこで、記録画像に対しては時間に関する確率密度関数のみを考える。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確率密度関数を以下のように推定する。

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

ここで、 d は撮影時刻として区別可能な最小時間である。

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に関しては同様な手法で確率密度関数を推定可能である。

4.3 e-mail メッセージ

現在、e-mail 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ため、e-mail 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、メッセージにおいては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。時間に関する確率密度関数 $f_i^{email}(t)$ と空間に関する確率密度関数 $f_i^{email}(x, y)$ を利用して時間・空間を表す確率密度関数 $f_i^{email}(x, y, t)$ を以下のように推定する。

$$f_i^{email}(x, y, t) = f_i^{email}(t) \times f_i^{email}(x, y)$$

以下に時間と空間それぞれについて確率密度関数の推定方法を述べる。

単一のメッセージには複数の時間概念が含まれる。e-mail メッセージから抽出可能な時間概念は、メッセージの内容として本文やタイトルに含まれる時間表現と、送受信や閲覧等のコンピュータ上の処理時刻を示す時間表現とに大別できる。メッセージに含まれる時間概念は自然言語処理によって抽出する。メッセージ m_i に含まれる時間概念集合を t_1, \dots, t_n とする。

このときメッセージ m_i が表す確率密度関数を以下の式により推定する。

$$f_i^{email}(t) = \frac{1}{\sum_k w_k} \sum_j (w_j \times N(t_j, \sigma_j))$$

ここで、 w_i は t_i に関する重みである。また、 $N(t_j, \sigma_j)$ は平均 t_j 分散 σ_j の正規分布を表す。メッセージに含まれる時間概念は不確実性が高いため、個々の時間概念を正規分布としてモデル化する。分散は時間概念の粒度または不確実性を表している。メッセージ中の時間概念の分散を正確に推定するためにはメッセージの内容の理解が必要である。

メッセージに含まれる地名等の空間概念を利用して空間的な特徴を推定する。空間の確率密度関数は時間に関する確率密度関数と同様な手法で求める。自然言語処理により空間概念を抽出し、個々の空間概念 s_i を辞書を利用して経度と緯度を用いた地点ベクトル μ_i にマッチングし、確率密度関数を 2 次元正規分布 $N(\mu_i, \Sigma_i)$ として表現する。ここで、 Σ_i は共分散行列である。メッセージ全体の空間に関する確率密度関数 $f_i^{email}(x, y)$ は以下の式で求める。

$$f_i^{email}(x, y) = \frac{1}{\sum_k w_k} \sum_j (w_j \times N(\mu_j, \Sigma_j))$$

5. 活動ノードに関する確率密度関数推定

活動は、ライフログ中のデータによって表現される対象である。活動は様々な視点から捉えることが可能であり、データによって表現される活動をあらかじめ活動を想定することは困難である。

そこで、すでに存在するメディアノードに基づいて活動ノードを推定することを考える。メディアデータが活動の時間・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 e-mail 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が、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のに対して、e-mail メッセージから得られる特徴にはノイズが含まれている可能性が高い。

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

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

a_i の確率密度関数として m_i の確率密度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。

6. まとめ

本論文では、ライフログ検索を目的として、異種メディアの時空間概念を確率分布関数として表現しメディア統合を実現する手法を提案した。本論文で示した確率密度関数の推定方法は基本的なアプローチを示したものであり、実際の検索においては目的に応じた調整が必要である。

今後、評価実験により本手法の有効性を検証する予定である。また、ノードのクラスタリング等による処理の高速化を実現する必要がある。さらに、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を検索モデルへの組み込むことを考えている。

参考文献

- [1] 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).
- [2] 牛尼剛聡, 渡邊豊英: ライフログ内の記録画像を対象とした撮影コンテキストに基づく検索, 夏のデータベースワークショップ DBWS2004 (2004), (印刷中).
- [3] 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).