

## Relevance Feedbackを用いた個人行動記録の検索 Personal View Records Retrieval by Relevance Feedback

小泉敬寛†  
Takahiro Koizumi

中村裕一†  
Yuichi Nakamura

亀田能成†  
Yuichi Ohta

大田友一†  
Yoshinari Kameda

### 1. はじめに

頭部装着カメラをはじめとする多種のセンサを用いて個人の行動記録を蓄積する試みが行なわれている [1][2] が、ハードウェアの性能が上がったために膨大な記録の蓄積が可能になっている。一方で、ただ蓄積されただけの映像には、何もせず一箇所で止まっている場面など無駄な部分も大きな割合で含まれている。そのため、膨大な個人行動記録中から必要な部分へ素早くアクセスするための検索手法が必要となる。本稿では、そのための手法として、構造化 Relevance Feedback 法を提案する。

### 2. Relevance Feedback による個人行動記録の検索

#### 2.1 個人行動記録の獲得

本研究では、個人行動記録として、頭部に装着した小型カメラを用いて日常の様子を映像として記録し続けた個人視点映像と環境映像を用いる。

得られる個人視点映像が膨大な量となるため、移動や1箇所で立ち止まっている場面などの冗長な部分を取り除く。そのために、久保田らの手法 [3] を用いて、注目シーンを検出する。また、個人視点からの映像だけでは周囲の状況が把握しにくいいため、我々の手法では、周囲の様子を記録するデータとして、部屋全体を撮影する外部カメラの映像(環境映像)を同時に記録する。さらに、環境映像から人物像を検出することによって、個人視点映像がどこで撮影されたか、つまり、カメラ装着者がどこにいたかがわかる。以下では、個人行動記録を、注目シーンとこの位置情報の組の集まりとして考える。

#### 2.2 Relevance Feedback の利用

ユーザが個人行動記録の中から検索しなければならぬデータは、状況に応じて異なる。例えば、財布がどこにいったのかを探したい場合は財布の映っている映像が必要となり、工具を誰が持っていたのかを知りたい場合は工具置き場に立ち寄った人の映像が必要かもしれない。また、事故が起こった原因を知る必要がある場合にはさらに様々な映像が重要な情報となる。このように、ユーザの目的に応じて有効な情報が異なるため、映像の柔軟な検索手法が必要である。

この問題に対して、本研究では、個人行動記録を検索するために Relevance Feedback を用いる。これによりユーザが欲しい画像をできるだけ速く検索することを目的とする。その大まかな枠組みは以下ようになる。

ユーザは最初に提示された複数の画像または特徴量などのデータから、探したい情報に関係があると思うもの(複数でもかまわない)を選択する。選ばれたデータを基

に、システムが各特徴の重要性を計算し、類似検索のための各特徴の重みを変化させる。例えば、特徴量の分散が小さい特徴はユーザが意識的、無意識的であるに関わらず重要な特徴量であると考えられるため、その特徴の重みを大きくする。

変化した特徴量の重みを使って再度検索を行い、さらに、ユーザが関連すると思うものを選択することを繰り返す。その結果、各特徴量の重みが再計算されていくことで、よりユーザの目的に近いデータが効率よく検索される。

具体的な計算式は以下ようになる。

個人行動記録の1要素を  $O$  とすると、 $O$  には、画像、位置座標、時刻といった特徴の種類が  $1, \dots, l$  の  $L$  個、それぞれの特徴の種類に対応する  $1, \dots, k$  の特徴量の個数が  $K$  個含まれるとする。

ユーザが関連するものとして選んだ個人行動記録の要素の集合を  $Rel$  とすると、個人行動記録の  $i$  番目の要素  $O_i$  が  $Rel$  とどれだけ似ているのかを示す類似度  $S$  は以下のように計算される。

$$S(O_i) = \sum_{O_j \in Rel} D(O_i, O_j)^\alpha \quad (1)$$

$$D(O_i, O_j) = \left( \frac{1}{L} \sum_l w_l d(O_{il}, O_{jl})^\gamma \right)^{\frac{1}{\gamma}} \quad (2)$$

$$d(O_{il}, O_{jl}) = \left( \sum_k (O_{ilk} - O_{jlk})^\eta \right)^{\frac{1}{\eta}} \quad (3)$$

式2は各記録の類似度、式3は各特徴量の距離となる。また、重み  $w_l$  は次の式で更新される。

$$w_l^{new} = \frac{\sigma_l^{all}}{\sigma_l^{rel}} \quad (4)$$

$$w_l = \alpha \times w_l^{new} + (1 - \alpha) \times w_l^{old} \quad (5)$$

ここで、 $\alpha$  は定数であり、 $\sigma_l^{all}$  は特徴  $l$  について全要素間で式3を求めた結果の分散であり、 $\sigma_l^{rel}$  は、 $Rel$  に含まれる要素のみについて同様の計算をしたものである。

この枠組みにより、ユーザが適切なデータぶことを繰り返していけば、ユーザが欲しいデータが効率よく検索されることになる。

### 3. Structuring Filter による個人行動記録の構造化

Relevance Feedback による検索を行っても、検索された画像が自分の欲しい情報に近いものかどうかを判断できない場合がある。例えば、カメラの操作方法が知り

†筑波大学 大学院システム情報工学研究科

†京都大学 学術情報メディアセンター

たい場合に、カメラ自体の外観を知っていても、そのマニュアルの色、形や置場所を知らなければ、画像による検索は難しい。

本研究で提案する構造化 Relevance Feedback では、このような場合でも、関連する可能性のあるものを選んで提示することにより、効率的な検索を可能にすることを目的としている。上記のカメラの例でも、そのマニュアルを見ながらカメラを操作した際の個人行動記録があれば、それを利用してカメラとマニュアルが同時に検索されるような仕組みである。

構造化 Relevance Feedback では、Relevance Feedback によって検索されたものに対して、関連性のあるものが強め合うように、適合度 (Relevance) の再計算が行われる。そのアルゴリズムは以下ようになる。

まず、個人行動記録の各要素間の関連の有無を表わす関連行列を作成する。式6は2つの要素  $O_i, O_j$  間の  $r$  番目の関連の種類毎に定義され、その有無を返す。関連の種類  $r$  での  $i$  番目と  $j$  番目の要素間の関連の有無  $a_{rij}$  を行列の形でまとめたものが式7となり、各要素間の関連を表わす関連行列となる。

$$a_{rij} = \text{assoc}_r(O_i, O_j) \quad (6)$$

$$A_r = \begin{pmatrix} a_{r11} & a_{r12} & \dots & a_{r1n} \\ a_{r21} & a_{r22} & \dots & a_{r2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{rn1} & a_{rn2} & \dots & a_{rnn} \end{pmatrix} \quad (7)$$

関連としては、例えば以下のようなものが考えられ、関連行列も定義した関連性の種類の数だけ作られる。

- ある要素に対して同一の移動軌跡
- ある要素に対して同一停留点付近
- ある要素に対して時刻以前・以降

各要素が類似検索で得られた要素とどれだけ関連があるのかを示す関連度を求める。関連度は各関連行列と式1により得られる結果の集合  $S = \{S(O_0), S(O_1), \dots, S(O_n)\}$  とをかけたものとする。

$$R_r = A_r S \quad (8)$$

ここで、各関連性毎に計算される  $R_r$  は、Relevance Feedback の結果から関連するものだけを強める係数とする。されに、この結果を足し合わせた結果  $R = \sum R_r$  は全ての関連性により強められた結果となる。そこで、得られた関連度  $R$  が閾値以上のものは、ユーザが求める情報に関連する可能性が高いとし、上記の再検索で優先的に選び、提示する。

#### 4. 実験

実験のために、簡易な GUI インターフェースを作成し、40分程度の個人行動記録の中から、ロッカーに機器を取りに行くシーンを検索した。

個人行動記録は環境カメラの設置された実験室内で撮影され、同一人物によるものである。類似検索のための特徴量としては、それぞれ、色ヒストグラム、時刻、人

物の位置座標を用いた。個人視点カメラの装着者が一箇所に静止しているなどして、位置座標が検出できなかった時刻に対しては、前後の位置座標からの線形補完用いている。

検索のキーとして、機器の映っている注目シーン(画像による検索)、ロッカーの周辺(位置による検索)を入力した。その結果、どちらの手法でもロッカーから機器を取り出すシーンを検索結果として得ることができた。しかし、機器の画像による検索の場合は、類似機器を使用する無関係な場所でのシーンが、また、ロッカーの位置による検索の場合はロッカー周辺で行なわれた無関係の作業シーンが多く検索された。どちらの場合でも、さらに、ロッカーで機器を取り出すシーンを選択して Relevance Feedback を繰り返すことで重みが改善され、目的のシーンが多く検索されるようになった。

また、同一の移動軌跡上であるという関連性を用いて構造化 Relevance Feedback を適用すると、機器を取り出す前後のシーンが強められ、上記のような不要なシーンが検索されるのを抑えることができた。また、ロッカーに貼ってある、持ち出しチェック表にチェックを入れるシーンを確認することができた。

#### 5. まとめ

Relevance Feedback を用いて類似検索を行なうシステムを実装し、個人行動記録の類似検索を行なう手法を提案した。類似検索により得られた結果から、さらに関連するデータを強め合うことで、個人行動記録の検索が容易になることを示した。しかし、実験で用いた個人行動記録が量・種類ともに不足しているため、特徴量間に大きな相関が見られ、重みが期待したように変化しない場合があった。今後、実験データを増やし適切な重みの更新処理を考える必要がある。

#### 参考文献

- [1] 河野恭之, 河村竜幸, 上岡隆宏, 村田賢, 浮田宗伯, and 木戸出正継. ウェアラブル日記の実現に向けて - 日常記憶の検索・編集・整理・共有機構 -. In *PRMU2003*, 2003.
- [2] 堀鉄郎, 河崎晋也, 石川尊之, and 相澤清晴. ライフログ応用に向けたコンテキストに基づく映像・データ検索. In *MVE2003*, 2003.
- [3] 久保田敏司, 中村裕一, 大田友一. 個人行動記録システムにおける注目シーンの検出 - 注目シーン検出の高精度化と環境カメラの利用. 信学技報 *PRMU-2002-64*, pp.47-52, 2002.
- [4] Leejay Wu, Christos Faloutsos, Katia P. Sycara, and Terry R. Payne. FALCON: Feedback adaptive loop for content-based retrieval. In *The VLDB Journal*, pages 297-306, 2000.