

体験記録の特徴箇所抽出に関する性能評価

Evaluation of Significant Experience Selection

山口 仁[†] 太田 昌克[‡] 山口 正泰[‡]
 Hitoshi Yamaguchi Masakatsu Ohta Masayasu Yamaguchi

[†]NTTネットワークサービスシステム研究所

[‡]NTT未来ねっと研究所

1. はじめに

今日、個人の体験を記録するための環境は大きく変化しつつある。例えばデジタルカメラやデジタルビデオカメラが普及し、以前はフィルムやビデオテープに記録されていたデータがデジタルデータとして記録され、活用が容易になってきている。

「入力機器の変化」や「記録媒体の変化」は、記録スタイルにも大きな影響を与えている。従来は「必要なものだけを厳選して記録する」という手法が主流であったが、最近のデジタルカメラによる写真撮影などでは「多く記録を行い、利用時に選択する」という手法に変わってきている。このような傾向は、今後ユビキタス環境が整いウェアラブルカメラや健康を管理するセンサなどが普及すると「記録者が記録行為を意識しない」という記録も可能になるため、一層強まると考えられる。したがって、大量の体験記録から重要な箇所を効率的に取り出す手法が必要である。

本稿では、このような大量の個人体験記録を有効に利用するための手法として、連続データからの特徴箇所抽出法を提案し、抽出性能に関する評価実験の結果を示す。

2. 時系列データからの特徴箇所抽出法

個人の体験記録から特徴箇所を抽出する手法として、これまで(a)記録済みコンテンツ自体を解析する方法、(b)記録時に並行して測定した体験者の生体反応を利用する方法などが提案されてきたが^[1]、これらは (i)対象コンテンツや生体反応に対する多くの専門知識、(ii)環境変化や反応の個人差への対応が必要である。これらの課題を解決する手法として、我々はこれまで人工免疫系の特異点検出法^[2]に基づいた特徴箇所抽出について提案を行ってきた^[3]。ここではその概要を説明する。

2.1 人工免疫系の特異点検出技術

免疫系では「自己の体内組織(self)」と「ウイルスなどの外敵(non-self)」とを判別し、未知の外敵(non-self)から自己を防御するシステムを形成している。この免疫系のメカニズムを利用した人工免疫系の特異点検出法が提案されている(図1)^[2]。

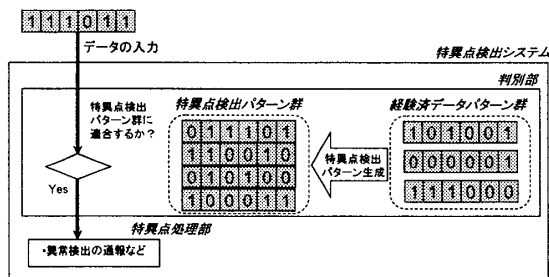


図1 人工免疫系における特異点検出法

図1の判別部において、システムは入力データパターンを self と non-self に判別する。システムは、内部に蓄積された「経験済データパターン群(self)」から複数の特異点検出パターンを生成し、それらの検出パターンに適合する入力データを特異点(non-self)と判別する。

2.2 人工免疫系を用いた特徴箇所抽出

上述の人工免疫系の手法を連続データに応用し、連続データにおける特異点を特徴箇所として抽出するのが、本稿での特徴箇所抽出法である。この手法を用いた個人体験からの特徴箇所抽出の一例としては、「動画像と一緒に生体センサ情報を記録し、生体センサ情報における特異点を特徴箇所として抽出することで動画像にマーキングを行なう」といったことが挙げられる。

システムにはデータを分割・単純化し、データパターンを作成する前処理部が必要になる。動画像からの特徴箇所抽出を例にすると、「動画をフレームごとに静止画像に分割する」、「分割した静止画を単純化してデータパターンを作成する」等が前処理部で行われる。

このように作成されたデータパターンは、人工免疫系の特異点検出により特異点のデータパターン(以下、特徴データパターン)とそれ以外(以下、平常データパターン)とに判別される。そして、特徴データパターンが検出された箇所を、特徴箇所として抽出する。

2.3 特徴箇所抽出における課題

このような特徴箇所抽出においては、「経験したデータパターンをいつまで保持するか」が大きな問題となると考えられる。極端な例として、一度経験した特徴データパターンを永久に保持するシステムを想定すると、同じデータパターンを持つ特徴箇所は2度と検出されないため、時間が経過するにつれて特徴箇所が検出されなくなると考えられる。逆に保持可能なデータパターン数を極端に小さく見積もると、平常データパターンがすぐに削除されてしまうため、非常に多くの箇所が特徴箇所として抽出されてしまうと考えられる。

そこで、保持するデータパターン数を変化させて特徴抽出の実験を行い、その検出精度に対する評価を行なった。

3. 実験

実験では、特徴箇所を抽出するためのセンサ情報として、以下の理由から「瞬目」を用いた。

- 瞬目頻度は体験者の興味を反映するので^[4]、体験記録からの特徴箇所抽出に有効であると考えられる。
- 瞬目の閉眼期間は、開眼期間に対して非常に少ない。
- 瞬目については正解/不正解の判定が比較的容易であり、評価が行ないやすい。

瞬目の判定について、その概要を図2に示す。

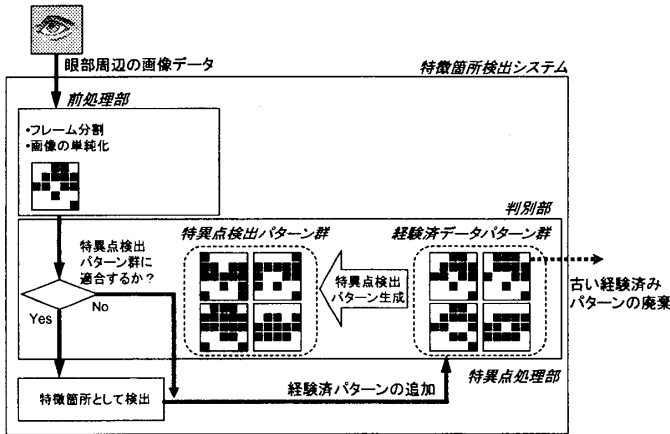


図2 瞬目抽出の概要

実験では、眼周周辺の動画像 (30 フレーム/秒) を入力データとして用い、前処理部において入力動画像をフレーム単位に分割した後 8×8 のブロックに変換し、各ブロックを白もしくは黒に二値化してデータパターンを作成した。

入力されたデータパターンは、特徴データパターンか平常データパターンかに関わらず、経験済データパターン群に追加する。ただし、すでに保持しているデータパターンを追加する際には古いデータパターンを削除し、同一データパターンを重複して保持しないようにする。また、保持できる限界数を超える場合には古いデータパターンから廃棄を行なう。

実験では、保持できるデータパターン数を 20, 40, 60, 80 に設定し、それぞれの場合について瞬目の判定実験を行なった。どの場合でも経験済データパターン群が満たされる 120 秒後以降の判定結果を評価に用いた。

4. 実験結果および考察

図3に、保持するデータパターン数を 20~80 の間で変化させた場合の評価を示す。図3に示すとおり、保持できるデータパターン数を増加させても適合率はあまり変化せず、再現率は大きく減少している。

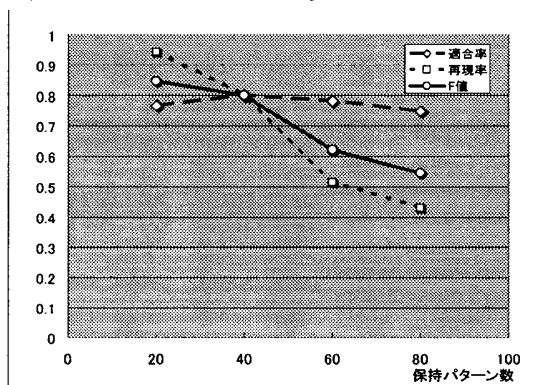


図3 保持するデータパターン数と判別性能

この理由は、以下のように考えられる。

保持できるデータパターン数を増やすと、システム内の経験済データパターン群が大きくなり、その結果、システムが「誤って」瞬目と判定する回数は減少する。ただし、

経験済データパターン群には瞬目のデータパターンも登録されるため、システムが「正しく」瞬目と判定する回数も減少してしまい、適合率は変化しない。また、保持できるデータパターン数を増やすと、実際の瞬目回数が変わらないのに、システムが「正しく」瞬目と判定する回数が減少してしまうため、再現率は減少する。これにより、保持できるデータパターン数を増やすと F 値が減少することになってしまう。

図4に、新規データパターンの累計が時間とともに増加する様子を示す。

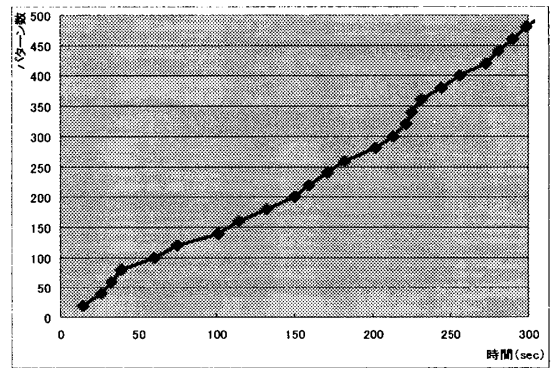


図4 新規データパターンの累計変化

ここでは 120 秒以降もほぼ一定の割合で新規データパターンが発生していることが示されており、保持できるデータパターン数を増やすことが「誤って」瞬目と判定する回数を減少させるのに効果的だと思われる。ただし、単純に経験済データパターン群を大きくするだけでは「正しく」瞬目と判定させる回数も減少させてしまうため、適合率を向上させるためには「データパターンの出現頻度を用いたデータパターン保持」等、平常データパターンを優先的に保持するメカニズムが必要になると考えられる。

5. まとめ

連続データの特徴箇所抽出について、保持できるデータパターン数と性能との評価を行ない、保持できるデータパターン数を増やすだけでは F 値を制御できないことが分かった。今後は、性能を向上させるためのデータパターン保持方法について検討する。

6. 参考文献

- [1] 相澤他, "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 J86-D-II, Jun 2003.
- [2] S. Forrest et al, "Self-Nonself Discrimination in a Computer", Proc. 1994 IEEE Symposium on Research in Security and Privacy, 1994.
- [3] 山口他 「体験記録からの特徴箇所抽出法の提案」2004 信学会総合大会 D-12-11 p.177.
- [4] 石山他 「立体画像観察時における瞬目についての考察」NHK 技研 R&D No.43, Nov 1996.