

動画視聴時の注目点を考慮した動画推薦方法の提案*

夏目 大伍[†] 入部 百合絵[‡] 桂田 浩一[†] 新田 恒雄[†][†]豊橋技術科学大学 大学院工学研究科[‡]豊橋技術科学大学 情報メディア基盤センター

1. はじめに

近年, YouTube[1]などの動画投稿サイトが人気を博している。これら動画投稿サイトでは, 検索キーワードを入力するだけで, 関連する動画を閲覧できる。しかしながら, 投稿サイトは膨大な動画コンテンツを保持しているため, キーワードによっては検索の結果, 大量の動画が提示されてしまう。また, サイトに動画を投稿する際には, 投稿者が動画のタイトル等の付加情報を自由に編集できるため, 検索キーワードと関係のない動画が検索され, 提示されることもある。

こうした課題への対応策として, ユーザの興味に即した動画を推薦する方法が考えられる。先行研究の動画推薦やテレビ番組推薦では, 土屋ら[2]のようにユーザが視聴した動画のタイトルやカテゴリといった, 動画の付加情報を利用する方式が多い。しかしこのような方式は, ユーザが動画を視聴している際に, 実際に動画中のどの部分を注目していたのかを考慮していない。

そこで, 本研究では動画視聴時の視線の動きに注目する。人は興味のあるところを注視し, 関心の高いところに視線を向けることが知られている[3]。このことから, 動画視聴時の注目点が類似するユーザ同士は, その視線の動きも類似していると推測される。

本報告では, まず動画視聴中の視線移動から注目点を抽出し, ユーザ間の移動軌跡の類似性に基づいて動画を推薦する方法を比較検討する。続いて, 注目点を利用して推薦した際の順位と, 実際にユーザが評価した結果を順位相関から評価した結果について述べる。

2. 動画視聴時における注目点

2.1 注目点の定義

人間の視線データから得られる情報として

- 停留点座標, 停留時間
- 瞬きの頻度, 瞬きの間隔
- サッケード距離, サッケード頻度
- 瞳孔径

などがある[4]。視線は, 視線の跳躍運動であるサッケードが起こってから, 次のサッケードが起こるまでの間, 一箇所に停留する。これが停留点であり, 視線の停留が起こっている間は注視点付近から視覚情報を得ている。

このことから, ユーザが動画を視聴する際に発生する視線の停留点は, そのユーザの注目点であると考えられる。以下, 本報告において注目点と記す場合は, 視線の停留点を意味するものとする。

西田ら[5]の報告によれば, 絵画鑑賞においては, 特定の部分に興味, 関心を集中している場合, 300[ms]以上の視線停留が多く観察されることがわかっている。そこで

以下では, 対象が絵画ではなく動画であることを考慮し, 200[ms]以上の間, 視線移動が視角にして 2° 以下であった場合に, その部分を注目点とした。

2.2 注目点の比較

動画視聴時の注目点は時間の経過と共に移動する。そのため, 本研究ではユーザの注目点を, 移動軌跡データとして表現した。具体的にはまず, 2.1 の定義に基づき, 視線データから視線停留点を動画視聴時の注目点として抽出する。続いて, 注目点間を直線近似することにより, 注目点の移動軌跡データを作成する。時刻 t における注目点の移動軌跡データ λ は, $\lambda(t) = (x, y)$ と与えられる。

同じ動画に対する移動軌跡データ同士の比較を, 柳沢らの手法[6]を参考に以下のように行った。時刻 t における移動軌跡 $\lambda(t)$ 上の点 $\mathbf{p}$ と $\lambda(t')$ 上の点 $\mathbf{p}'$ の距離 $Dist(\mathbf{p}, \mathbf{p}')$ は,

$$Dist(\mathbf{p}, \mathbf{p}') = \sqrt{(x - x')^2 + (y - y')^2} \quad (1)$$

となる。また, 二つの注目点移動軌跡データ λ, λ' 間の距離 $Dist(\lambda, \lambda')$ を

$$Dist(\lambda, \lambda') = \sqrt{\sum_{i=1}^n Dist(\mathbf{p}_i, \mathbf{p}'_i)^2} \quad (2)$$

と定義する。

3. 動画視聴時の注目点を用いた動画推薦

本報告では, 動画視聴時の注目点移動軌跡データをもとに, ユーザ間の興味度の合いを評価し, 動画を推薦する方法を検討する。

ある動画に高い評価をしているユーザがいて, 他方, このユーザが示す「注目点の移動軌跡データ」と類似の軌跡データを示すユーザがいる場合, 両者は注目点が類似していると考えてよいだろう。すなわち, 移動軌跡データ間の距離に近いほど, 視聴動画に対する興味度の合いも近いとする。

この仮定をもとに動画推薦を行うシステムを考える。最初に, 評価対象のユーザ A , 他のユーザを $\{B, C, D, \dots\}$ とした時, ユーザ A の注目点の移動軌跡に類似の軌跡を持つユーザを $\{B, C, D, \dots\}$ の中から選択する。ユーザ A に推薦される動画は, この注目点が近いと判断されたユーザ等が視聴した動画のうち, ユーザ A が未視聴でかつユーザ評価が高いものである。以下に具体的な手順を説明する。

まず, ユーザ A とユーザ B 間の興味度の近さ $Dist(U_A, U_B)$ を, 両ユーザが視聴した全ての動画に対する注目点移動軌跡データ間距離の総和から, 次式のよう

$$Dist(U_A, U_B) = \sum_{i \in \text{videos}} Dist(\lambda_i(U_A), \lambda_i(U_B)) \quad (3)$$

*Proposal of Video Recommendation Method Using Eye Mark
D.Natsume[†], Y.Iribe[‡], K.Katsurada[†], T.Nitta[†]

[†] Graduate School of Engineering, Toyohashi Univ. of Tech.

[‡] Information and Media Center, Toyohashi Univ. of Tech

ここで、 $\lambda_i(U_A)$ はユーザ A が動画 i を視聴した際の注目点の移動軌跡データである。 $Dist(U_A, U_B)$ は正値を取り、この値が小さいほど、動画視聴時のユーザ A、B 間の興味傾向が似ており、また値が大きいほど二人の動画に対する興味傾向が似ていないことを示す。

動画推薦の際には、上述の手続きから得た、推薦対象ユーザに興味傾向の近いユーザの集合 O_{sim} に注目する。本報告では、推薦対象ユーザと距離の近いユーザ上位 5 名からなる集合を O_{sim} とする。 O_{sim} に含まれるユーザの評価の合計値をもとに、動画 i の推薦値 $Score(i)$ を次式から算出する。

$$Score(i) = \frac{\frac{1}{N_{user}} \sum_{U \in O_{sim}} Rating_i(U)}{\frac{1}{N_{pair}} \sum_{U_1, U_2 \in O_{sim}} Dist(\lambda_i(U_1), \lambda_i(U_2))} \quad (4)$$

ここで、 N_{user} は O_{sim} に属するユーザ数、 N_{pair} は O_{sim} に属するユーザ対の数、 $Rating_i(U)$ はユーザ U が動画 i を視聴した際の評価値である。(4)式の分母は、動画 i を O_{sim} に含まれるユーザが視聴した際の注目点移動軌跡データ間距離の総和を全ての対について計算したものをユーザ対の数で割ったものである。 $Score(i)$ は、推薦対象ユーザと興味傾向の似ているユーザの評価が高く、ユーザ同士の注目点が類似している動画ほど、高い値を示す。

推薦対象ユーザには、 $Score(i)$ の値が高い動画から順番に推薦を行う。

4. 評価実験

4.1 実験概要

被験者は、大学生及び大学院生 20 名である。各被験者には、頭部が動かないよう顎台で顎を固定した状態で視線測定装置(ナック社製 EMR-8)を装着し、動画を視聴させた。この際の視線の動きを DV カメラで記録し、視線測定装置に付属の解析ソフトウェアを用いて視線データを得た。実験に使用した動画は、YouTube に投稿されている「邦楽」に関連する動画 15 本である。各被験者には、動画を 1 本視聴する度に、視聴した動画に関する興味の度合いを 4 段階で答えてもらった。なお、視聴させた動画の順番は被験者毎に異なる。

実験結果は、20 名の被験者を推薦対象ユーザ 1 名、比較ユーザ 19 名のグループにわけ、交差確認法を適用して評価した。評価指標としてはスピアマンの順位相関係数を用い、提案方法による推薦順位と推薦対象ユーザから得た順位の相関関係を調べた。

4.2 実験結果

YouTube 等の動画投稿サイトを利用する場合、ユーザは自分が興味のある動画を視聴することが多いと考えられる。そこで、実際の利用形態に近い状態で推薦を行うため、推薦対象ユーザの評価が 4 段階で 3 以上の動画のみを用いて、残りの動画を推薦した。例えば、使用動画が 1 本の場合は、その 1 本をもとに残りの 14 本を推薦する。これを、使用動画数が 1 本の場合、3 本の場合、推薦対象ユーザの評価が 3 以上であった動画の総数の半分(平均 4.75 本)を使用した場合について行い、提案方法による推薦順位と推薦対象ユーザが実際につけた推薦順位との順位相関の平均値を算出した。これを表 1 に示す。

表 1 使用動画数と順位相関の関係

	1 本使用時	3 本使用時	半数使用時
順位相関 (標準偏差)	0.50 (0.23)	0.53 (0.23)	0.61 (0.22)

表 2 ユーザ数と順位相関の関係

	ユーザ 10 名	ユーザ 15 名	ユーザ 20 名
順位相関 (標準偏差)	0.52 (0.26)	0.55 (0.22)	0.61 (0.22)

表 1 の結果より、使用動画本数が多いほど、提案方法による推薦結果が向上していることがわかる。これは、使用するデータが多いほど、ユーザ間の類似性が現れやすいためであると思われる。一般に相関が強いといわれている 0.7 以上の値となったユーザの数は、1 本使用の場合で 2 名、3 本使用の場合で 5 名、評価が 3 以上となる動画の総数の半分を使用した場合は 8 名であった。

評価が 3 以上となる動画の総数の半分を使用する推薦に関して、ユーザ数を変化させた場合の順位相関を表 2 に示す。表から、ユーザ数の増加と共に、推薦順位とユーザが実際につけた順位との相関が高くなることが分かる。これらのことから、ユーザ数が増加すれば更に適切な推薦が可能になると考えられる。

5. まとめ

動画投稿サイトにおける動画推薦方法として、ユーザの動画視聴時の視線から注目点移動軌跡データを算出し、推薦対象ユーザと他のユーザとの注目点の類似性をもとに推薦順位を決定する方法を提案した。評価実験から、被験者 20 名、被験者の評価が 4 段階で 3 以上であった動画の総数の半分を使用した場合の推薦において、推薦順位と被験者が実際につけた順位との相関が 0.61 となった。しかし、標準偏差が 0.22 と大きく、推薦結果にばらつきが出ている。今回の結果は、動画に付与したタグ等の付加情報を利用せずに評価した結果であるため、今後、タグ情報を加味することで、精度改善を計りたい。また、ユーザ数の増加と共に、様々な興味傾向に対応できるようになり、推薦結果の向上が期待される。

今後の課題として、動画に付与されているタグとの組み合わせによる精度の向上、複数カテゴリの動画を用いた実験と評価が挙げられる。

参考文献

- [1] YouTube : <http://jp.youtube.com/>
- [2] 土屋 誠司 他 : “TV 番組推薦システムの構築とその有用性の検証”, 情報処理学会研究報告, ヒューマンインターフェース研究報告, Vol.2006, No.3, pp.95-102, 2006
- [3] 大野健彦 : “視線を用いたインターフェース”, 情報処理, Vol.44, No.7, pp.726-732, 2003.
- [4] 大野健彦 : “視線から何がわかるか - 視線測定による高次認知処理の解明”, 日本認知科学会「認知科学」, 9 巻第 4 号, pp.565-576, 2002.
- [5] 西田謙太郎 他 : “停留点の時間的な発生間隔に着目した注目点検出方法の検討”, 情報処理学会研究報告, ヒューマンインターフェース研究会報告, Vol.2007, No.11, pp.121-126, 2007
- [6] 柳沢豊 他 : “移動軌跡データに対する類似度検索手法”, FIT2002 概要集, pp.37-38, 2002