

ストーリーの流れと顕著なイベントに基づく 個人化ダイジェスト自動生成手法の検討

Study on Automatic Generation of Personalized Digest based on Context Flow and Distinctive Events

宮森 恒†
Hisashi Miyamori

1. まえがき

筆者らは、映像中のストーリー全体の流れと個々の顕著なイベントに基づき、オリジナル映像の意味内容を表現したナレーション付きダイジェストの自動生成手法を提案している[1]。本稿では、テニスを対象とした実装システムの評価のためにスポーツ等の番組制作経験者にダイジェストの正解データを作成してもらい、提案法の結果と比較することでその有効性について考察した。

2. 提案法の概要

図1に、提案法の処理概要を示す。

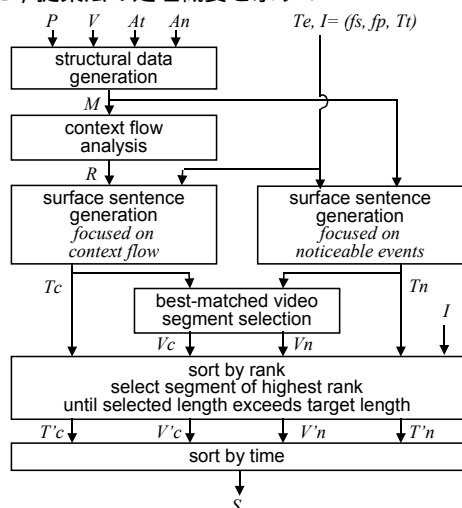


図1. ダイジェストの生成手順

入力データは、スコアデータ P 、映像情報 V 、時間順選手動作 A_t 、顕著な選手動作 A_n 、要約用テキスト要素 T_e 、利用者入力 I の6種類のデータから構成される。

紙面の都合上、詳しい処理手順の説明[1]は省略するが、本手法では、自動取得が可能と考えられるインデックス P 、 A_t 、 A_n [2]に基づき映像の内容把握を行い、 V 、 T_e からナレーション文や対応する映像を生成している。また、生成されるダイジェストの内容は、利用者入力 I によって適応的に変化する。

利用者入力 I は、利用者の興味・嗜好を反映させるための項目で、内容構成 f_s 、着目選手 f_p 、要約時間 T_t から成る。ここで、 f_s は試合展開などストーリーを重視した要約を見たいのか、スーパーショットなど個々のイベントを重視した要約を見たいのか、といった内容構成を指定する項目で、 f_p はどの選手に着目した要約とするかを指定する項目である。 T_t は、生成した要約の再生に許容できる時間を表す。

3. 実験結果

3.1 提案法によるダイジェスト

内容構成 f_s 、着目選手 f_p 、合計時間 T_t の各値を $f_s = \{0: \text{ストーリー重視}, 1: \text{イベント重視}\}$ 、 $f_p = \{0: \text{選手 A}, 1: \text{選手 B}, 2: \text{均等}\}$ 、 $T_t = \{15, 30, 60: \text{単位=秒}\}$ と変化させ、要約映像とナレーションを自動生成した。ナレーション文は、現在、予め用意されている候補文テンプレートに必要な数値等を補って連結することで生成している。各候補文には1つのシーンが対応付けられる。

以下に、 $f_s = 0$ 、 $f_p = 0ka$ 、 $T_t = 30$ の設定で生成されたナレーション文の例を示す。

第1セットは、ノオカのほぼ一方的な展開となり、ノ選手ノオカが6-2で危なげなくとった。ノ第2セットは、Hinomuraが一時有利に試合を進めるもののノほぼ互角の接戦となり、ノ選手ノオカが6-4で逃げ切った。

また、 $f_s = 1$ 、 $f_p = 0ka$ 、 $T_t = 30$ の設定で生成されたナレーションの例を示す。

第1セットは、パッシングやノサーブ&ボレーを決めるなどしてリードしたノ選手ノオカが6-2で危なげなくとった。ノ第2セットは、ノオカがサービスエースやノストロークエースを決めるなどしたが、ノ選手ノオカが6-4で逃げ切った。

また、 $f_s = 1$ 、 $f_p = \text{Hinomura}$ 、 $T_t = 30$ の設定で生成されたナレーションの例を示す。

第1セットは、Hinomuraがサーブ&ボレーやノサービスエースを決めるなどして好プレーを見せたがノ選手ノオカに6-2で危なげなくとられた。ノ第2セットは、Hinomuraがパッシングやノストロークエースを決めるなどして好プレーを見せたがノ選手ノオカに6-4で逃げ切られた。

内容構成や着目選手を変化させることにより、ストーリーの流れを保持しつつ、着目選手を中心とした表現やイベントが多く含まれ、利用者の嗜好に応じてダイジェスト内容を再構成できていることが確認できる。

3.2 番組制作経験者によるダイジェスト

提案法では、内容構成・着目選手・要約時間という3つの利用者パラメータ入力によって、自動生成されるダイジェストの内容が変化する。そこで、生成したダイジェストを評価するための正解データを、同様のパラメータが与えられたとして、番組制作経験者(スポーツを含む)に依頼し作成した。表1に作成概要を示す。

表1. 評価用ダイジェストの作成概要

作成者	スポーツを含む番組制作経験者5名
使用映像	未編集の素材映像3試合 編集済の放送映像2試合
パラメータ	$f_s = \{0, 1\}$ 、 $f_p = \{0, 1, 2\}$ 、 $T_t = \{15, 30, 60\}$

各作成者、各試合毎にパラメータを18通りに変化させ、映

†通信総合研究所, CRL

像とナレーションからなるダイジェストを計 450 通り作成した。

以下に、 $f_s=0$, $f_p=0$, $T_t=30$ の設定で生成されたナレーション文の例を示す。

親善試合男子シングルス, 対戦相手は Hinomura . / 今日の日 Oka, 序盤から積極的に前へ出ます . / 第 1 セット第 5 ゲーム, 粘りのプレーでブレイクすると, / 勢いに乗った Oka, このまま第 1 セットをとります . / 続く第 2 セット, Hinomura に一時追い込まれる場面もありましたが, / 着々とポイントを重ねた Oka . 2-0 のストレートで勝ちました .

また、 $f_s=1$, $f_p=0$, $T_t=30$ の設定で生成されたナレーションの例を示す。

親善試合男子シングルス, 相手は Hinomura . / サーブ&ボレーの早い攻撃 . / いいところでサービス / & リターンエース . / 相手が出てくるところをクロスに, / そしてストレートに抜く . / 最後まで自分のリズムで試合を / 進めた Oka, 2-0 のストレートで勝利 .

また、 $f_s=1$, $f_p=Hinomura$, $T_t=30$ の設定で生成されたナレーションの例を示す。

親善試合男子シングルス, 相手は Oka . / 早い仕掛けの Hinomura . / ネットプレーが次々決まる . / サービス, リターンともに / 切れ味抜群 . / Hinomura の鋭いショットに Oka も翻弄される . / しかし肝心なところでのミスショットが響き, / 0-2 で惜しくも敗れる .

3.3 比較および考察

まず、正解データの定性的な性質について確認する。正解のナレーションには一定の構造的なパターンが見られる。すなわち、第 1 文で試合や選手の紹介、最後の文で結果の報告、そして中間の文で、経過や見所などの説明を行っている（いわゆる起承転結の構造）。これは使用映像や作成者によらない傾向であった。

表 2 に、正解データの最初・中間・最後の文のそれぞれで使用頻度が上位の単語を示す。この表からも使われる単語の種類と起承転結構造に一定の相関があることがわかる。

また、中間文の傾向として、 $f_s=0$ の場合、試合展開を表現する単語が多く使われるのに対し、 $f_s=1$ の場合、スーパープレイを表現する単語がよく用いられることも確認することができる。

ここで、提案法によるダイジェストを正解データと比較すると以下のような共通点・相違点が確認できた。

- (1) 提案法は、途中経過の概略は表現できているが、試合や選手紹介にあたるイントロや全体の結果報告にあたるまとめがない。これは試合や選手を表すメタデータやスコアデータから随時生成可能と考えられる。
- (2) ストーリー展開と個々の特徴的イベントの内容を f_s , f_p , T_t の値に応じて良好に再構成している。
- (3) 要約文に用いられる表現が限られている。正解データに見られる活き活きとした鮮やかな表現と比べると機械的な印象を受ける。これについては自然言語処理技術などを導入し改善する必要がある。
- (4) ダイジェストに使用した一部の映像は共通していた（各セットの最後のプレーやエース等）が、そうでない例も多く存在した。しかし、これらは被験者実験による総合的評価も必要と考えられる。

表 2. 正解文で使用頻度の高かった単語(頻度順)

	$f_s=0$	$f_s=1$
最初の文	試合, シングルス, 親善, オープン, 女子, 男子, ジャパン, 混合, ダブルス	試合, シングルス, 親善, 女子, 対, オープン, 男子, ジャパン, 混合
中間の文	セット, ゲーム, 取る, サービス, しかし, 調子, ポイント, リード, 落とす, ミス, プレー, 出る, ショット, なる, 入る, 結局, 前	セット, サービス, ショット, しかし, 取る, ゲーム, プレー, エース, 相手, ストレート, ボレー, サーブ, ストローク, 決まる, 決める, ネットプレー, ハンド
最後の文	セット, する, 敗れる, ストレート, 勝利, 勝つ, カウント, 結局, 見事, 取る	セット, 勝利, 敗れる, 試合, する, カウント, ストレート, 取る, 見事, 選手

今後は、ダイジェスト生成部分のチューニングを行い、一般の被験者に対して本手法によるダイジェスト視聴実験を行ない、有効性を検証する必要がある。

図 2 に生成したダイジェスト再生の様子を示す。



図 2. 生成されたダイジェスト映像とナレーションの同期再生

4. まとめ

ストーリー展開と個々の特徴的イベントに着目した個人化ダイジェストの自動生成手法の有効性を評価するために、番組制作経験者の作成した正解ダイジェストの特性と提案法による結果の比較を行なった。要約生成部分のチューニングを行い、被験者実験による本手法のダイジェスト視聴の有効性検証が今後の課題である。

文献

- [1] 宮森: “文脈の流れと特徴的なイベントに基づく動的な映像要約手法の提案”, 画像の認識・理解シンポジウム(MIRU2002), II-371 ~ II-378, 2002
- [2] 宮森: “映像と音響情報の協調による内容検索のためのテニス動作自動注釈付け”, 信学論 D-II, Vol. J86-D-II, No. 4, pp. 511-524, 2003