

デジタルカメラワークによるサッカー映像生成 Auto-Making Soccer Video Movie by Digital Camerawork

鈴木 健司† 誉田 雅彰† 村上 真† 白井 克彦†
Kenji Suzuki Masaaki Honda Makoto Murakami Katsuhiko Shirai

1. はじめに

近年、放送の多チャンネル化、ブロードバンドの普及に伴い映像コンテンツが増加している。またデジタル技術の進歩に伴い、家庭でも容易に高精度のビデオカメラが利用できるようになった。今後は個人が映像コンテンツの編集を行い、放送局の代わりに映像配信を行う機会も増加することが予想される。そこで現在、映像生成の効率化を求められ、必ずしもプロのカメラマンや編集者を必要としない自動映像生成技術が注目されている。本研究では、自動映像生成の1手法として考えられるデジタルカメラワークを用い、サッカー映像に対し、重要イベントの認識・映像の編集を自動で行う映像生成システムの構築を目指している。本稿ではその第1段階として、TVサッカー映像と同様の形式である試合の流れがわかるサッカー映像の生成法を述べる。

2. デジタルカメラワーク

デジタルカメラワークとは自動撮影法の1つである。固定カメラで撮影した映像を素材とし、その1部分をクリッピングすることで映像を生成する手法である。デジタルカメラワークによるサッカー映像生成には、サッカーフィールド全体像を固定カメラ1台によって撮影した映像を用いる。この映像素材に対し、デジタル処理を行うことで映像生成を行う。デジタルカメラワークならば、試合進行によってクリッピングサイズ、クリッピング位置を変更することで様々な映像を生成することができる。

3. 処理

デジタルカメラワークによるサッカー映像生成の処理の流れを示す(図1参照)。

3.1 メタデータ生成部

入力に対し映像の各フレームごとのボール、選手の絶対座標、イベントの取得を行う。ボール・選手認識は、背景差分による2値化の後、輝度・円形度・面積に基づき領域を抽出する。領域検出後は、3次元アフィン変換

によりボール、選手の絶対座標取得を行う。絶対座標取得後、得られた選手・ボールの絶対座標を元に、独自に決めた認識規則に基づきイベント認識(ドリブル・パス・クロス・シュート・ゴール)を行う[2]。

3.2 映像生成部

メタデータ生成部で取得したメタデータ(ボール位置・選手位置・イベント)を用い映像生成を行う。本稿では映像生成部の中でも、試合の流れがわかる映像の生成に必要である、遠景・近景映像生成、スイッチングの制御法について述べる。今回デジタルカメラワークによる映像生成法の有効性を示すため、メタデータ生成部で未検出、誤認識したメタデータは手動で訂正し映像生成を行った。

3.2.1 クリッピングサイズ・位置の決定

映像生成には、クリッピングサイズと位置を決定する必要がある。先行研究[1]よりTVサッカー映像は、ほぼボールを中心にしたカメラワークを行っていることが明らかになった。よって遠景映像、近景映像ともに、ボールの画面上の座標を追跡することでクリッピング位置を決定する。またクリッピングサイズは、本稿の実験用の映像(HDV1440×1080でフィールド半分を撮影)の場合、遠景映像は元映像の2倍、近景映像は8倍とし映像を生成する。映像の拡大には拡大縮小アルゴリズムである16点バイキュービック補間を用いる。

3.2.2 カメラワーク補正処理

ブレのない映像を生成するため、カメラワークに補正処理を行う。図2のように画面の中央に補正枠をおく。この補正枠内にボール座標が存在する場合はクリッピング位置を移動させない。逆に補正枠の外にボール座標が動いた場合、クリッピング位置を移動させる。この際、画面の中心をボールの座標に移動させるとカメラワークが急激に変化することになる。そこで補正枠の外に出た場合、画面の中心からボール座標までの距離を計測し、その距離に応じて1フレームあたりの移動量を決定する。

3.2.3 スイッチング

これまでの研究で、TVサッカー映像(7試合630分)の検証を行った結果、イベントとシーンの継続時間がスイッチングにおいて重要な要素であると判明した

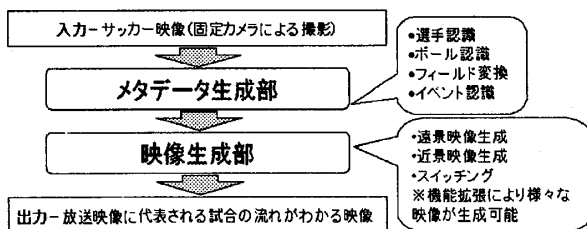


図1: 処理フロー

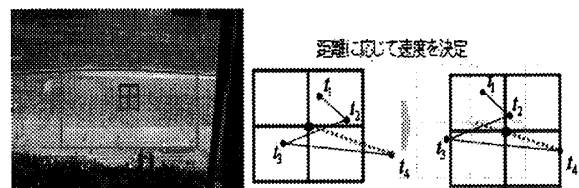


図2: カメラワーク補正処理

†早稲田大学

‡東洋大学工学部

[3]. よってイベントと継続時間長を考慮したスイッチングの状態遷移モデルを生成し、遷移確率に基づいてスイッチングの自動制御を行う。イベントに関してはイベントの組み合わせを2つ1組として考え、組の始めのイベントが発生した時の切替確率を状態遷移確率とする。また継続時間長については、映像継続時間の累計切替確率を状態遷移確率とする。表1, 2に各状態遷移確率の代表例を示す。これら各遷移確率を利用し、モデルの状態遷移確率を算出する(式1参照)。各状態遷移確率は、TVサッカー映像(7試合630分)によって学習したものをを用いる。

$$\begin{aligned}
 S &= \{S_F, S_N\} \\
 a_{ij} &= P(S_i, S_j) \quad i, j \in F, N \\
 E &= \{e_{lm} \mid l, m \in \{PS, PF, CS, CF, D, S\}\} \\
 P(a_{ij} \mid e_{lm}, d) &= P(a_{ij} \mid e_{lm}) \times P(a_{ij} \mid d) \\
 \left(\begin{array}{l} P(a_{ij} \mid e_{lm}) : \text{イベントの状態遷移確率} \\ P(a_{ij} \mid d) : \text{継続時間長の状態遷移確率} \\ P(a_{ij} \mid e_{lm}, d) : \text{モデルの状態遷移確率} \end{array} \right) & \quad (1)
 \end{aligned}$$

表1: イベントの状態遷移確率

event	ff(遠→遠)	fn(遠→近)	nf(近→遠)	nn(近→近)
D+CS	0.74	0.26	1.00	0.00
D+CF	0.75	0.25	0.91	0.09
D+S	0.90	0.10	1.00	0.00
PF+D	0.90	0.10	0.51	0.49

表2: 継続時間長の状態遷移確率

遠景継続時間 (s)	切替合計	fn	ff	近景継続時間 (s)	切替合計	nf	nn
5	130	0.26	0.74	1	284	0.33	0.67
10	252	0.50	0.50	3	647	0.75	0.25
15	375	0.75	0.25	5	802	0.93	0.07

4. 評価

4.1 一対比較法による評価実験

生成映像の評価に主観評価法である一対比較法を試みる。実験用の入力映像は、固定カメラ(HDV1440×1080)2台で2005年度早慶サッカー戦を撮影した映像を使用する。比較映像は以下の通りである。これら6つの映像に対し、同様のシーン(60sec)、映像サイズ(720×540)で、画質を考慮しないという条件下のもと「試合の進行のわかりやすさ」という項目で一対比較を行った。なお被験者は大学生6名である。

- M1 - 本手法のスイッチングモデル(2sec周期)で生成した映像
- M2 - スwitchingを手動で切り替えた映像(放送映像を模倣)
- M3 - スwitchingを時間で制御した映像(遠景18秒, 近景3秒)
- M4 - 遠景映像(switchingなし, カメラワーク補正あり)
- M5 - 遠景映像(switchingなし, カメラワーク補正なし)
- M6 - 入力映像(固定映像)

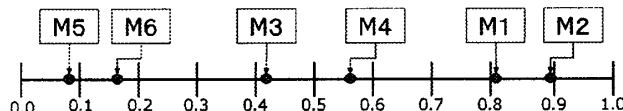


図3: 一対比較による映像評価

勝率に基づき、映像を数直線上に変換したものが図3である。結果より、本手法で生成した映像M1の有効性のある程度示すことができた。また最も評価の高かったM2とM1の数値が近似していることから、switchingに関して本手法の映像は、放送映像同様人間が手動で切り替えを行った映像と、遜色のない映像が生成することができたと考えられる。他にもM4とM5の数値に大きな差があることから、本手法のカメラワーク補正処理が映像生成をする上で非常に有効であることがわかる。

4.2 アンケートによる評価実験

4.1節のM1の映像に対しアンケートによる評価を行う。アンケート項目は以下の通りであり、それぞれ5段階(1. 許容できない, 2. 少し厳しい, 3. まあ許容できる, 4. 許容できる, 5. 十分に許容できる)で評価する。

- a - 遠景映像の画質
- b - 近景映像の画質
- c - 近景映像の1シーンの長さ
- d - スwitchingの切り替わるタイミング
- e - 映像全体を通してのswitchingの回数
- f - パンの自然さ(速度, 追跡対象を参考に)
- g - サッカーを視聴するにあたりどうか(画質を考慮しない)
- h - サッカーを視聴するにあたりどうか(画質を考慮する)

各項目の平均値を表3に示す。項目a~fまでの平均値より本手法の生成映像は、近景映像の画質以外は映像を視聴するにあたり許容範囲であることがわかる。また項目g, hより総合的な映像の評価を行った結果、近景映像の画質の劣化を除けば、プロのカメラマンを必要としない本原稿の映像生成手法でも、サッカー映像を視聴するにあたり、十分に許容できるレベルに到達していると判断できる。

表3: アンケート項目の平均値

平均値	a	b	c	d	e	f	g	h
	4.67	2.17	4.00	4.00	3.83	4.67	4.33	2.33

5. まとめと今後の課題

本稿では、デジタルカメラワーク、switchingの自動制御による映像生成法の提案を行い、実際、固定サッカー映像(2005年度早慶サッカー戦)を利用し映像を生成した。生成映像に対し主観評価実験を行った結果、画質を考慮しなければ、プロのカメラマンを必要としない本手法の映像でも、サッカー映像を視聴するにあたり十分許容できるレベルに到達しているとの結果が得られた。今後は、メタデータ生成部の精度向上とともに、映像生成部の機能拡張を行い、デジタルカメラワーク独自の新たなサッカー映像の提示法を検討する予定である。

参考文献

- [1] 加藤大二郎, 山田光穂他”被写体を追尾撮影時の放送カメラマンのカメラワーク特性分析” テレビ学会誌, 50, 12pp. 1941-1948 (1996.12)
- [2] 川口克則, 稲葉大樹他”TVサッカー動画の要約生成のための自動Indexing” 情報処理学会第66回全国大会, (2004.3)
- [3] 加賀明久, 大平茂輝他”サッカー映像における近景切替の要因分析” FIT2004 第3回情報科学技術フォーラム (2004.9)