



人気領域を考慮した映像切り替え方式の提案

著者	西岡 欣亮, 住野 優, 井東 大, 上島 紳一
発行年	2006-06-30
その他のタイトル	Proposing Video Editing Scheme Based on Popular Area
URL	http://hdl.handle.net/10112/4540

人気領域を考慮した映像切り替え方式の提案

西岡 欣亮[†] 住野 優[†] 井東 大^{††} 上島 紳一[†]

[†] 関西大学大学院総合情報学研究科 〒569-1095 大阪府高槻市霊仙寺町 2-1-1

^{††} 関西大学総合情報学部 〒569-1095 大阪府高槻市霊仙寺町 2-1-1

E-mail: †fb4m133@edu.kansai-u.ac.jp, yutaka.s@mb.newweb.ne.jp,

ueshima@edu.kutc.kansai-u.ac.jp, ††fa20046@edu.kansai-u.ac.jp

あらまし ウェアラブルカメラなどの記録技術の発達や映像機器の普及により、多くの場面で手軽に映像撮影を行うことが可能となった。様々な視点から撮影された映像は、撮影者本人のみならず他人にとっても有益で多くの利用可能性を秘めている。従来、日常生活やスポーツイベントでの映像に対する効果的な利用法について多くの研究が行われてきた。本論文でも、スポーツイベントにおいて複数台のカメラにより撮影される状況を研究背景とし、一般撮影者間で共有した映像データに対する切り替え方式の提案を行う。提案方式では、撮影カメラの視線に基づき人気領域の特定を行い、その領域と各カメラの位置関係による幾何学計算によって時刻ごとのカメラ選択を行う。これによりイベントにおける多数の映像撮影者の興味を考慮した映像を生成する。これを行うために各映像に対して、GPS と電子方位計より取得した位置・方向情報をメタデータとして用いる。

キーワード 映像切替え, 位置・方向情報, 映像共有, MPEG

Proposing Video Editing Scheme Based on Popular Area

Yoshiaki NISHIOKA[†], Yutaka SUMINO[†], Dai ITOH^{††}, and Shinichi UESHIMA[†]

[†] Graduate School of Informatics, Kansai University Reizenji 2-1-1, Takatsuki, Osaka, 569-1095 Japan

^{††} Faculty of Informatics, Kansai University Reizenji 2-1-1, Takatsuki, Osaka, 569-1095 Japan

E-mail: †fb4m133@edu.kansai-u.ac.jp, yutaka.s@mb.newweb.ne.jp,

ueshima@edu.kutc.kansai-u.ac.jp, ††fa20046@edu.kansai-u.ac.jp

Abstract Due to the rapid advancements of storage technologies and popularization of video appliances, we have more chances to use video-recording for create documentaries of various events, concerts, and/or to keep records in our daily life, with digital handycams, wearable cameras and so on. So far, many researches have been done on effective reuse of such video data that are reorded in sports events as well as in our daily life. In this paper, we focus on handling of those video data that have been recorded by multiple handycams from different aspects and angles. For such video data, every single recorded video data is considered to be valuable not only to the recorder himself/herself but also to the public. In this context, we generate a serialized video data that are of interest to the audience. For this purpose, this paper proposes a framework to share these video data among users over the Internet and also provides a video switching scheme based on popular area. In our scheme, we estimate popular areas by considering gazing points of handycams of audience. Then we provide a scheme to switch video cameras according to geometrical calculation using the camera locations, directions obtained by GPS and Gyrocompass and the estimated popular areas.

Key words video edit, locational and direction data, shared video, MPEG

1. はじめに

デジタルビデオなどの普及と高性能化に伴い、誰もが手軽に映像撮影を行えるようになった。また、ウェアラブルカメラの一般利用も実現されつつあり、近い将来には日常生活全てを記録することが可能になる。このように一般撮影者個人による情報処理が充実し、多くの情報が生み出されるようになるにつれ、新たな枠組みが求められる。膨大な量の映像情報を効率良く利用するには、その中から必要または価値の高い情報を取り出すことが必要である。従来、大量または長時間の映像に対して利用価値の高い部分映像を抽出するために、多くの研究が行われてきた。例えば、ウェアラブルカメラにより撮影した映像から生理データを用いて、撮影者が興味を感じた部分の映像検索やダイジェスト作成する手法 [1], [2], 撮影位置や方向、被写体までの距離を計測するセンサを用いて複数のカメラの中から興味対象がより良く写っている映像を検索する手法 [3], [4], 映像の内容を記述したメタデータを付与することで、キーワードにより所望の部分映像を抽出する手法 [5] などがある。また、社会レベルでは一般利用者によるマルチメディアコンテンツの制作から流通、消費に至るまでを考慮した MPEG-21 も標準化されており、マルチメディア資源を有効に活用することができる環境が整いつつある。本研究でも、ウェアラブルカメラなどを用いて、多くの人が手軽に映像撮影を行う状況を研究背景とし、サッカー、フィギュアスケートや陸上競技などの全方位から観戦されるスポーツイベントを適用モデルとして提案を行う。これらのイベントにおいて撮影された映像の交換・共有を想定し、撮影カメラの位置・方向情報を用いて、視線交差に基づく人気領域の特定を行う。人気領域とは、多数の撮影者の興味を惹いている領域とする。本研究では、共有した複数の映像に対して、この人気領域と撮影カメラの位置関係の幾何学計算による切り替え方式について提案を行う。生成した映像を視聴する状況としては、家族団楽での視聴や、友人らと一緒に見るなど、多人数で共に楽しむケースを想定する。2. では研究動機について、3. では人気領域の特定方法とそれに基づくカメラ選択方法の説明について、4. では共有システムについて、5. では関連研究について、6. では今後の課題について述べる。

2. 研究動機

ユビキタスな映像撮影環境の中で撮影された長時間の映像、または多視点から撮影された大量の映像の中から、所望の部分映像を検索するためには、①有効な映像メタ情報と付与の簡便化が必要である。これまでに行われた研究では、撮影場所や時間などの内容記述、センサ機器から取得したデータをメタ情報として用いるものが多く、本研究もこれまでに、映像撮影と同時に取得した位置・方向情報を用いた幾何学計算による検索手法を提案してきた [6], [7]。以前に提案した手法の問題点は、部分映像の検索、表示に留まり、ユビキタス環境で撮影された映像に対して効果的な利用を行っていないことである。そこで本研究では、一般撮影者間での②映像共有を背景に、同一のイベントにおいて各撮影者により様々な視点から撮影された③映

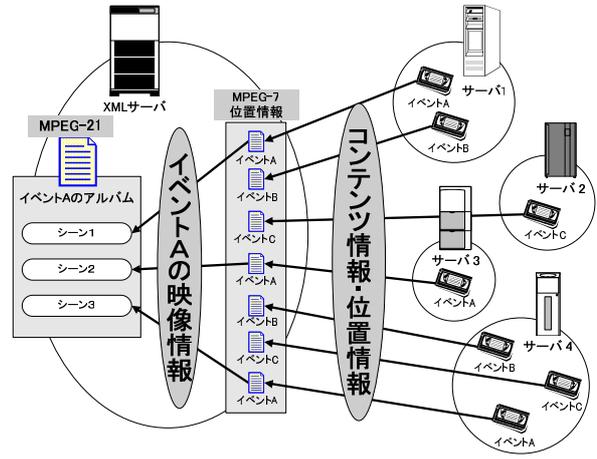


図 1 共有モデル

位置情報	方向情報
<pre><Location> <Latitude> <NorthLatitude sec="0">34°52'43.26" </NorthLatitude> <NorthLatitude sec="1">34°52'43.25" </NorthLatitude> : </Latitude> <Longitude> <EastLongitude sec="0">135°34'66.6" </EastLongitude> <EastLongitude sec="1">135°34'66.6" </EastLongitude> : </Longitude> </Location></pre>	<pre><Direction> <Azimuth angle> <Azim sec="0">77</Azim> <Azim sec="1">63.7</Azim> : </Azimuth angle> <Elevation angle> <Pitch sec="0">4.7</Pitch> <Pitch sec="1">3.6</Pitch> : </Elevation angle> <Rotational angle> <Roll sec="0">3.6</Roll> <Roll sec="1">3.9</Roll> : </Rotational angle> </Direction></pre>

図 2 位置・方向情報の記述例

像片の利用として、これまでの検索手法を用いた映像切り替えを行う。本研究で提案する映像切り替え手法は、以下の特徴を持つ。

- 一般撮影者間での共有映像の利用
- 不特定多数の撮影者の興味による切り替え
- 撮影カメラの位置・方向情報の索引付け

従来スポーツ中継する TV 局は、複数のカメラによって撮影された映像に対して、映像の切り替えや編集といった作業によって視聴者が見やすい番組を作り上げてきた。しかし、スポーツイベントにおいて多くの人が手軽に映像撮影を行うことができ、個人による情報発信や情報交換が可能になった現在において、この方法は一般撮影者においても適用可能である。本研究では、図 1 のように動画メタ情報と位置・方向情報を用いて各撮影者が撮影した映像を共有し、それらを用いて映像切り替えを行う。映像の切り替え方法を考える上で、何に重点を置くかを決める必要がある。例えば、好きな選手の動きに合わせて切り替えを行ったり、より近い距離で撮影しているカメラに重点を置くことによって、切り替えを行うことができる。センサから取得した生理データを用いた体験映像のダイジェスト作成 [1], [2] や TV 局のスポーツ中継などでは、撮影者もしくは TV 局のスイッチャーやカメラマンなどの興味により、大量の映像から“体験者の興味の割合が高い映像”や“視聴者にとって興味が湧くと考えられる映像”を選び出してきた。一方、本研究では多くの

撮影者が興味を惹きつけられている領域を考慮した映像切り替えを想定している．これを行うために本研究では，イベントを撮影している不特定多数の撮影カメラの視線データを用いる．これにより，イベントにおける興味の対象や盛り上がっている場面などを個人で選択するよりも正確に反映することを目指す．カメラが映している場所を求めることも考えられるが，現在のビデオカメラでは，常に変化する焦点距離を取得するのは困難である．そのため本研究では，視線交差によって興味を注がれている領域を特定する．

3. 人気領域の特定と映像切り替え手法

本研究では，イベントを通してより客観的な興味を考慮した映像の生成することにより，多くの視聴者と共に楽しめる映像の生成を目指している．ここでは，多数の視線による人気領域の特定方法と，人気領域を撮影した映像からより良い映像を選択するためのカメラ選択手法について述べる．

3.1 人気領域の特定

家族や友人達と共に楽しむ映像は，視聴者全員の興味を惹きつける映像が好ましいが，視聴者の興味対象は人によって様々である．例えば，親子でも好きな選手や見たいプレーは異なる場合がある．従って，多人数で共に視聴する映像は，特定の視聴者の興味に片寄ることなく，より多くの視聴者の興味を反映することが重要である．視聴者の興味対象は，視聴者が“興味を持つ場所”を求めることで推定できる．多くの視聴者が“興味を持つ場所”は，撮影者にとっても魅力的であり，多くの撮影者が同じ場所を撮影している可能性が高い．例えば，サッカーの試合で視聴者の見たい場所は，ボールを競り合ったり，人気選手がいる場所であり，撮影者もボールや人気選手を中心に撮影しているケースが多い．その結果，最も多くの視聴者が興味を持つ場所は，最も多くのカメラで撮影された場所と，ある程度一致すると考えられる．本稿では，最も多くの撮影者が撮影した場所を人気領域と呼ぶこととする．人気領域は，以下の手順で特定する．

- (1) カメラ視線の交点座標を算出
- (2) フィールド上の交点の最小包含円を求め人気領域とする

(2) の人気領域を撮影している部分映像を抽出し，時系列に繋ぎ合わせることで，多くの視聴者が楽しめる映像を生成する．

3.1.1 視線交差

本来，人気領域はカメラ焦点の位置関係を用いて特定するのが望ましい．一方で，焦点位置を特定するには，常に変化するビデオカメラのズーム情報を記録する問題がある．現在，市販のビデオカメラからズーム情報を取得するのは困難であるため，焦点距離は考慮しないこととした．本研究では，焦点に代わる方法として，カメラ視線の交点に着目する．視線の交差は，カメラの台数が n 台 ($n \geq 2$) のとき，最大 nC_2 個の交点が期待できる．視線交差の位置座標は，図4のようにカメラAとBの緯度，経度と方位角を元に，球面三角法の余弦公式及び正弦余弦公式を用いて求める．カメラの緯度，経度，方位角は，カメラに取り付けられたGPSと電子方位計から取得する．

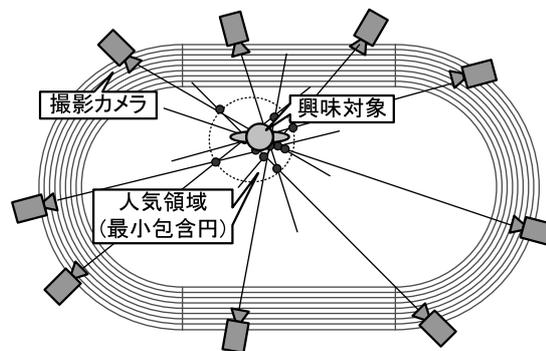


図3 視線交差と人気領域

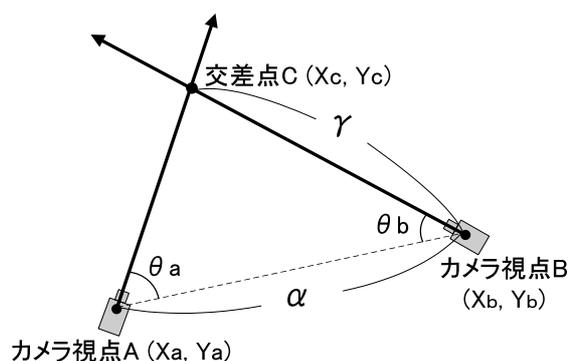


図4 交点計算

図4の交点 C を求める計算式を以下に示す．なお， $\theta_a, \theta_b, \alpha$ は2台のカメラの緯度，経度と方位角より求める．

- (1) 地点 B と交点 C の角距離 γ を求める

$$\gamma = \arctan \left(\frac{\sin \theta_b \cos \theta_a + \cos \theta_b \cos \alpha}{\sin(\arccos(-\cos \theta_a \cos \theta_b + \sin \theta_a \sin \theta_b \cos \alpha))} \right) \quad (1)$$

(但し， α は点 A と点 B の角距離とする．)

- (2) 角距離 γ を用いて交点 C の緯度 X_c と経度 Y_c を求める

$$X_c = \arcsin(\sin X_b \cos \gamma + \cos X_b \sin \gamma \cos \theta_b) \quad (2)$$

$$Y_c = \frac{Y_b + \arcsin(\sin \gamma \sin \theta_b)}{\cos X_c} \quad (3)$$

式(1)~(3)により全ての視線同士の交点を求める．求めた交点の緯度，経度を平面直角座標系に変換し，3.1.2節の最小包含円アルゴリズムの適用や3.2節のクラスタリングを行う

3.1.2 最小包含円による人気領域の特定

フィールド上にある交点から人気領域を特定するために最小包含円を用いた．最小包含円は，平面上にある点集合を包含する最小の円であり，以下の性質を持っている．

- 最小の円は一意に決まる
- 点の位置の偏りに影響を受けない
- 点のばらつきにより大きさが決まる

全方位からの観戦が可能なスポーツイベントでは，興味をひかれた撮影者の全ての視線が平面座標 (x, y) で表されるようなただ1点で集まることは考えにくく，興味対象となるものの周りに多くの交点ができる．交点の位置により，人気領域を特定す

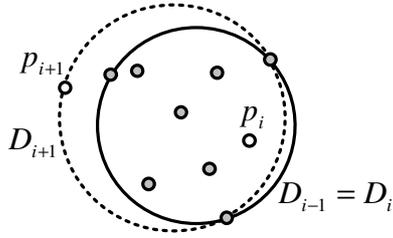


図 5 最小包含円の生成過程

るには、重心法により人気領域の中心点を求める方法も考えられるが、視線交差が起こったとき、交点が被写体の前方、後方、右側、左側のどこで多く起こるかは不確定であるため、重心法は適さない。また興味対象について考えてみると、例えばサッカーの試合において、選手が孤立した状態でボールを保持している場合と、ゴール前で多くの選手が競り合っている場合は、興味対象の大きさが異なるため、視線を向ける領域も広がり、交点もばらつく。本研究では、視線交差領域に最小包含円アルゴリズムを適用し、生成された最小包含円を人気領域とする。これによって、時刻 t のフィールド上にただ 1 つの人気領域を特定し、交点のばらつきにより興味対象の大きさを人気領域の大きさを反映させる。以下に最小包含円アルゴリズムについて述べる。

3.1.3 最小包含円アルゴリズム

最小包含円は、平面の n 個の点集合 P を包含する最小の円である。 P の点のランダムな順列を p_1, \dots, p_n としたとき、ここでは p_1, \dots, p_i の点集合を $P_i := \{p_1, \dots, p_i\}$ で表す。最小包含円アルゴリズムは点集合 $P_i := \{p_1, \dots, p_i\}$ について、 P_i の最小包含円 D_i を管理しながら、点を順次添加していく。図 5 のように、点 p_i を加えたときに現在の最小包含円 D_{i-1} に含まれるのならばそのままが良いが、点 p_{i+1} を加えたときに、現在の最小包含円 D_i に含まれないのならば、点集合 P_{i+1} の最小包含円 D_{i+1} は点 p_{i+1} を必ず境界上に持つという制約の下で生成される。以下に具体的なアルゴリズムの流れについて述べる。

MiniDisc(P)

入力：平面上の n 点の集合 P

出力： P に対する最小包含円

- (1) P のランダムな順列 p_1, \dots, p_n を求める。
- (2) D_2 を $\{p_1, p_2\}$ に対する最小包含円とする。
- (3) for $i \leftarrow 3$ to n
- (4) do if $p_i \in D_{i-1}$
- (5) then $D_i \leftarrow D_{i-1}$
- (6) else $D_i \leftarrow$
 $MiniDiscWithPoint(\{p_1, \dots, p_{i-1}\}, p_i)$
- (7) return D_n

点集合 $P_n := \{p_1, \dots, p_n\}$ の最小包含円を生成する過程で、点集合 P_n の中の 1 つの点 p_i が $p_i \notin D_{i-1}$ のとき、 p_i を境界上に含む最小包含円を求める。これを行うために $q = p_i$ とし、もう一度同じ枠組みを用いる。すなわち、 P_{i-1} の点をランダムな順序で加え、点 q は境界上にないといけないう制約の下

に $P_{i-1} \cup \{q\}$ の最小包含円を管理する。このとき、次のサブルーティンを用いる。

MiniDiscWithPoint(P, q)

入力：平面上の n 点の集合 P と、 q をその境界上に含むような P に対する最小包含円が存在するような点 q

出力： q をその境界上に含むような P に対する最小包含円

- (1) P のランダムな順列 p_1, \dots, p_n を求める。
- (2) D_1 を q と p_1 をその境界線上に含むような最小包含円とする。
- (3) for $j \leftarrow 2$ to n
- (4) do if $p_j \in D_{j-1}$
- (5) then $D_j \leftarrow D_{j-1}$
- (6) else $D_j \leftarrow$
 $MiniDiscWith2Point(\{p_1, \dots, p_{j-1}\}, p_j, q)$
- (7) return D_n

ここでも点集合 $P_i := \{p_1, \dots, p_i\}$ の中の 1 つの点 p_j が、 $p_j \notin D_{j-1}$ のとき、 p_j を境界上に含む最小包含円 D_j を生成する。最小包含円アルゴリズムは、 $p_i = q_1, p_j = q_2$ としたとき、2 点 q_1, q_2 が境界上に含むという制限の下でも同じ方法を再度適用する。これを次のサブルーティンで行う。

MiniDiscWith2Point(P, q_1, q_2)

入力：平面上の n 点の集合 P と、2 点 q_1 と q_2 。ただし、 q_1 と q_2 を境界上に含むような P の最小包含円が存在する

出力： q_1 と q_2 をその境界上に含むような P の最小包含円

- (1) D_0 を q_1 と q_2 に対する最小包含円とする。
- (2) for $k \leftarrow 1$ to n
- (3) do if $p_k \in D_{k-1}$
- (4) then $D_k \leftarrow D_{k-1}$
- (5) else $D_k \leftarrow q_1, q_2, p_k$ を円周上に含む円
- (6) return D_n

ここでも同様に、点集合 P_j の中の 1 つの点 p_k が $p_k \notin D_{k-1}$ ならば、 p_k を境界上に含むという制約が働く。このとき、3 点 q_1, q_2, p_k を円周上に含む円として一意に決まる。以上のアルゴリズムにより時刻 t における人気領域を特定する。

3.2 カメラ選択

これまでに本研究では、注視点を考慮した映像検索方法を提案してきた。[6] 注視点検索は、“どこを撮影しているのか”ということを考慮した検索方式であり、被写体の位置を検索条件として検索を行うことができる。ここでも人気領域を注視点とし、注視点検索により人気領域を映した部分映像を抽出する。3.1 節で述べた手法により時刻ごとに特定した人気領域は、多くの視聴者にとって興味があり、その領域を映している映像は価値が高いという考えから、人気領域を考慮した以下の方法によりカメラ選択を行う。選択条件は、時刻 t において最も近くで人気領域全体を映しているカメラを選んだ。具体的な選択方法については以下の手順 1~3 で述べる。

手順 1 人気領域の中心点が映っているカメラに絞り込む
 図 6 のように、二次元での撮影カメラの視線ベクトル v_c の式を $y = ax + b$ で表す。撮影カメラの画角に合わせて伸ばしたベ

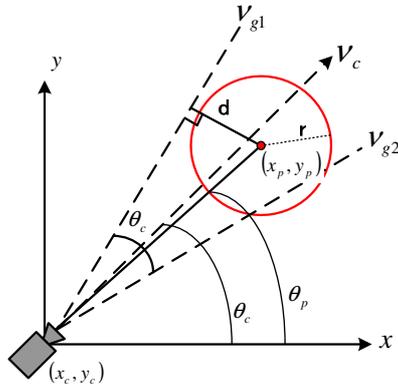


図6 視線と人気領域

クトル v_{g1}, v_{g2} は、視線ベクトルの傾き角度 $\theta_a(\arctan a)$ とカメラ画角 θ_c より求めた以下の式 (4), (5) となる。なお焦点距離が得られないため、直線は撮影カメラの位置から一定距離伸ばすものとする。

$$y = a_1x + b_1 \quad \text{但し, } a_1 = \tan(\theta_a + \theta_c/2) \quad (4)$$

$$y = a_2x + b_2 \quad \text{但し, } a_2 = \tan(\theta_a - \theta_c/2) \quad (5)$$

撮影カメラの座標を (x_c, y_c) とし、人気領域の中心点すなわち最小包含円の中心座標を (x_p, y_p) とするとき、撮影カメラと中心点を結んだ直線の傾きは θ_p は、以下の式より求まる。

$$\theta_p = \arctan \frac{y_p - y_c}{x_p - x_c} \quad (6)$$

これより撮影カメラの画角内に人気領域の中心点が映っている条件は、以下になる。

$$\theta_a - \theta_c/2 \leq \theta_p \leq \theta_a + \theta_c/2 \quad (7)$$

手順 2 式 (7) を満たしたもののの中から人気領域を映しているものに絞り込む。

画角にあわせて伸ばしたベクトル v_{g1}, v_{g2} が、人気領域と交わるかどうかを調べる。人気領域の中心点 (x_p, y_p) とベクトル v_{g1}, v_{g2} を表す直線との距離を調べ、人気領域の半径 r との比較によって求める。つまりベクトル v_{g1} の場合、以下の式を満たすものとなる。

$$d = \frac{|a_1x_p - y_p + b_1|}{\sqrt{a_1^2 + 1}} \geq r \quad (8)$$

手順 3 式 (7), (8) を満たしたもののの中から撮影距離が最も近いカメラを選ぶ

手順 1, 2 を満たしたカメラの中から、最も人気領域に近い距離で撮影しているカメラを選び出す。つまり、なるべく画面いっぱい人気領域が映っている映像を選択する。カメラから人気領域の中心点までの距離は、以下の式より求める。

$$D = \sqrt{(x_p - x_c)^2 + (y_p - y_c)^2} \quad (9)$$

これまでの手順で人気領域全体を映すカメラがある場合、時刻 t においてカメラ選択を行うことができる。人気領域は、視線の交点のばらつきにより大きさが決まる。全ての視線がただ一

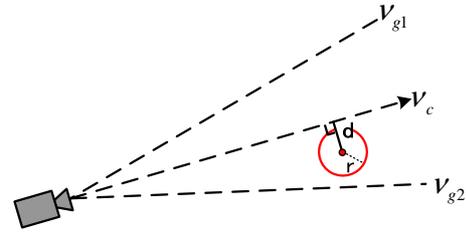


図7 人気領域が小さい場合

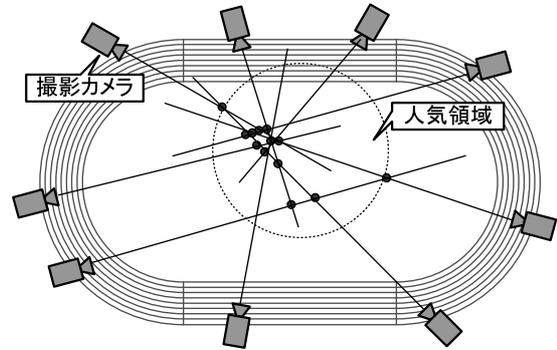


図8 視線がばらついた場合

点で交差することは考えにくい⁶、限りなくそれに近いような狭い領域で全ての交差が起こった場合、生成される人気領域は非常に小さいものになる。この場合、手順 1 ~ 3 によりカメラを選択した場合、図 7 のように人気領域を画面の端で捉えているようなカメラを選択することがある。人気領域をなるべく画面の中央で捉えたような映像を見たい場合には、視線ベクトル v_c と人気領域の位置より以下の式を条件として絞り込むことができる。

$$d = \frac{|ax_p - y_p + b|}{\sqrt{a^2 + 1}} \leq r \quad (10)$$

手順 4 式 (8) を満たすものがなかった場合

人気領域が大きい場合、全てのカメラの画角内に納まらない場合が考えられる。全方位から観戦可能なイベントを考えた場合、興味対象となるものの周りには多くの交点が見れ、それらを円で包含したとき、興味対象そのものは円の中心に近い位置にある可能性が高い。そこで人気領域全体を映すものがなかった場合は、手順 1 を満たしたカメラの中から、視線ベクトル v_c と人気領域の中心の距離 d が最も短いカメラを選択する。距離 d は、手順 2 と同様に式 (8) より求める。以上に述べた方法により時刻ごとにカメラ選択を行い、映像切り替えを行う。

3.3 c- 平均法による視線交差領域の抽出

サッカー、フィギュアスケートを考えた場合、多くの観客はボールがある場所や演技をしている選手を見る。映像撮影者も同様に、それらの場所を映すと考えられる。注目度が高い場所は、数多くのカメラで撮影されるため、多くの視線交差が起こる。しかし、全ての撮影者が必ずしもそうであるとは言い切れない。例えば、自分の好きな選手をずっと見ている場合や、同時刻のフィールド上において興味対象が複数存在する場合などが考えられる。その場合、図 8 のように、一人の撮影者が他の場所を見る視線によって、特定の興味対象の周り以外に視線

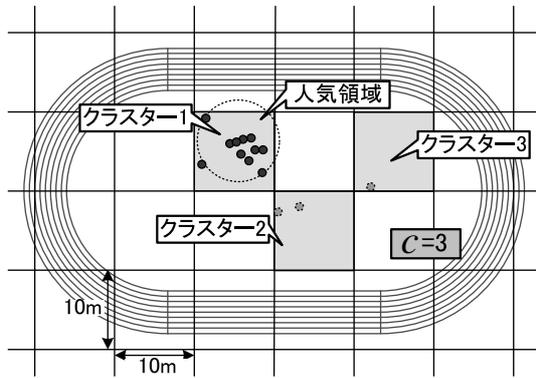


図 9 視線交差領域の抽出

交差の点ができる．視線が複数の箇所に向けられれば，それだけイベントが行われるフィールド上のあちこちに視線交差が起り，そのときの人気領域は非常に大きいものになる．仮にフィールド全体にばらつけば，人気領域はフィールド全体を覆うものになり，適切にカメラ選択を行うことができない．このため本研究では，イベントが行われるフィールド上にある交点の中から多くの交点が集まっている領域，つまり視線交差領域を抽出するために c -平均法を用いた． c -平均法は，あらかじめクラスター (c 個) の個数を指定し，対象データを c 個のクラスに分割する．非類似度はユークリッド距離の 2 乗をとり，クラスターの中心と各データとの間の非類似度を分類の基準とする． c -平均法は初期分割で選んだ個体により適切なクラスタリングが得られないことがあるので，本研究ではあらかじめ交点の位置によりある程度近い距離どうしの交点を同じクラスターに初期分割する．適用モデルとするスポーツイベントを考えた場合，興味対象となるものは人物やボールなどが多いので，フィールドを $10 \times 10m$ の領域で区切り，視線交差が起っている領域の数を c 個とし，クラスター分割を行った．図 9 の場合，クラスター個数 $c = 3$ となり，最も交点が存在するクラスター 1 が視線交差領域となる．アルゴリズムを以下に示す．

[アルゴリズム]

1. c 個のクラスター中心あるいは，初期分割をランダムに与える．
2. 各対象を最も近いクラスター中心に割り当てる．
3. 全ての対象の割り当てが一つ前のステップと変化がなければ終了する．そうでなければ，各クラスターの重心を新しい中心として 2 に戻る．

END

以上の方法により，イベントが行われるフィールド上から視線交差領域を抽出し，その領域にある交点に対し，最小包含円アルゴリズムを用いて人気領域を特定する．サッカーの試合を想定した場合の興味対象の一連の動きとカメラの位置を図 10 に示す．このときカメラ 1 ~ 3 が人気領域を常に撮影し，興味対象が図 10 のように動いたと仮定すると，提案手法による映像切り替えは，図 11 のようになり，興味対象を常に近い距離で捉えている映像を生成することができる．

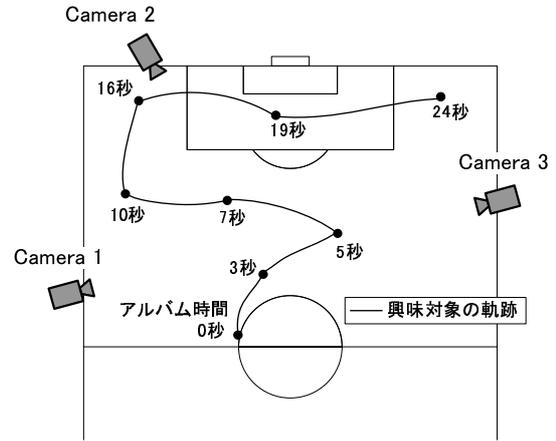


図 10 撮影状況



図 11 映像切り替え

4. 共有システム

想定する一般撮影者間での共有システムの説明，映像を交換・共有するためのメタデータについて述べる．

4.1 共有利用

一般撮影者が，同一のイベントに参加し，撮影した映像を共有する場合，その映像がいつ，どこで何を撮影したのかという撮影状況が重要になる．大量の映像データの中から，必要とする映像を検索するためにメタデータを付与する手法が多く用いられている．本研究では，これまでに映像に位置・方向情報と動画メタ情報を付加する手法を提案してきた [6], [7] 位置・方向情報は，撮影と同時に表 1 で示すパラメータを GPS，電子方位計より取得し，図 2 で示すような XML で記述する．動画メタ情報は，撮影場所や日時，あらすじなどの映像内容記述と，コーデックや解像度，オーディオ情報などの映像データ形式に関する記述を作成した入力フォームより記述することで，手軽に生成できる．本研究では，これらの位置・方向情報と動画メタ情報を用いて，図 1 のような共有利用を考えた．映像撮影者は，映像ファイルのアップロード時に次の作業を行う．

- (1) 撮影者が選んだサーバへ映像ファイルをアップする
- (2) 指定された XML サーバへ動画メタ情報と位置・方向情報を登録する

次に，サーバにアップされた複数の映像の中から自分が見たいイベントの映像切り替えを行うには，切り替えを行うイベントの日時と場所を登録することにより，自動的に XML サーバ内の動画メタ情報と位置・方向情報から切り替えに使用する映

表 1 位置・方向情報

	使用パラメータ	計測機器
位置情報	緯度, 経度	GPS
方向情報	方位角, 仰角, 回転角	電子方位計

像を選択する。選択した映像のメタ情報を図 12 に示すような MPEG-21 形式で切り替え情報を生成する。なお、本論文では生成された切り替え情報のことを複数の映像をまとめることからアルバム情報と呼ぶことにする。作成されたアルバム情報は XML サーバ内に保存される。アルバム情報には複数の映像ファイルの URL が記述されるため、視聴者は XML サーバ内のアルバムを参照するだけで、複数のサーバに格納されている映像ファイルを、あたかも 1 つの映像ファイルのように再生することができる。また、アルバム情報は XML サーバに蓄積されるため、アルバム作成者以外の人も利用し映像を見ることができる。なお、映像を交換・共有した場合、著作権などの問題が重要となるが、ここでは議論しないものとする。

4.2 アルバム情報

アルバム情報は、3.2 節の方式で選択したカメラ映像を、時系列に表した情報である。この時系列情報に従って映像を順次再生することで、複数のカメラで撮影した映像を、あたかも 1 つのメディア・リソースのように再生できる。この仮想的に繋ぎ合わせた映像を、より多くの人と共有し楽しむためには、アルバム情報を国際標準規格で記述することが望ましい。そこでアルバム情報の記述形式には、国際標準規格である MPEG-21 を使用する。MPEG-21 は複数のパートから構成されており、アルバム情報にはパート 2 の Digital Item Declaration を用いる。以下に、アルバム情報で用いる記述子について説明する。

- DIDL 記述子: Digital Item Declaration のルート記述子
- ITEM 記述子: 1 つのメディア・リソースの COMPONENT, ANCHOR 記述子を包括する
- COMPONENT 記述子: DESCRIPTOR, RESOURCE 記述子を包括する
- ANCHOR 記述子: アルバムに使用するメディア・リソースの部分映像を時間で指定
- RESOURCE 記述子: メディア・リソースの URL を記述
- DESCRIPTOR 記述子: メディア・リソースの MPEG-7 (メタ・コンテンツ情報) を記述

図 11 の映像切り替えをアルバム情報で記述すると図 12 になるこのアルバム情報を読み込んで、多くの視聴者が楽しめる映像を生成する。

5. 関連研究

ウェアラブル機器を装着し撮影した長時間の体験映像や複数のカメラで撮影した膨大な映像に対して、それらを効果的に利用するために映像の要約やダイジェスト作成、メタデータ付与、見たいシーンの検索の研究が行われている。

5.1 多視点映像に関する研究

[4] では、スポーツイベントにおいて興味対象を撮影時刻と撮影範囲をメタデータとして有する多視点映像の中から、被写

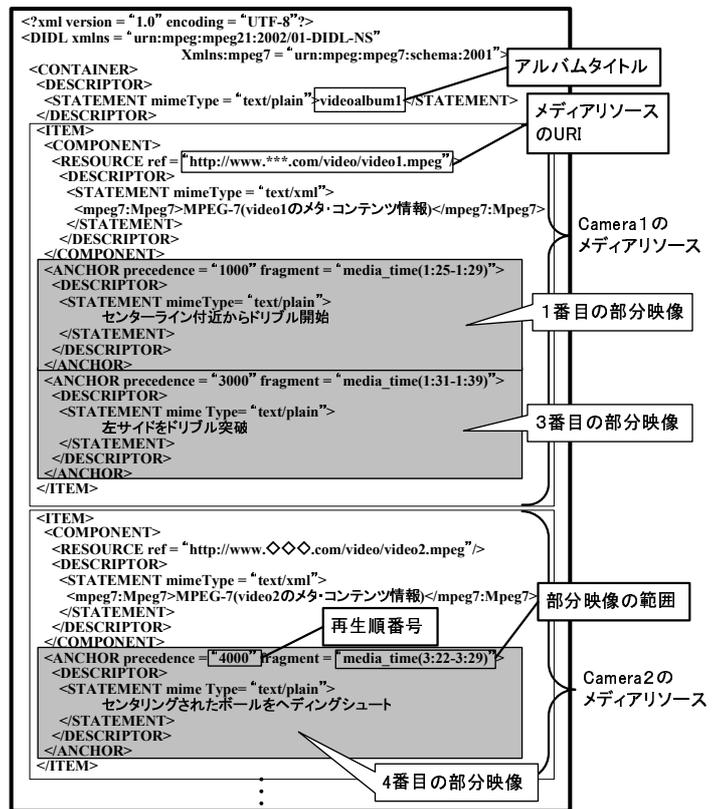


図 12 アルバム情報

体がどの程度良く撮影されているかを示す捕捉状態に基づいて "よりよく" 写している映像の集合を検索している [3] では、メタデータと捕捉状態を用いて意味的な関連を求め、それに基づき注釈情報を継承することで多視点映像全体にキーワードを付けている [8] では、TV 会議システムを有効に行うために、映像の各フレームごとに各参加者の視線方向のラベルを付け視線が多く集まっている人の映像を選択している [3], [4] では、スポーツイベントを適用モデルとし、本研究と同じく撮影カメラの位置と方向を用いて多視点カメラの中から良い映像を検索している。本研究との違いは、個人による興味を反映させている点である。本研究では、多くの撮影者の視線から興味のある領域を特定し、その領域が映っている映像を良い映像としている。

5.2 体験映像に関する研究

[1] では、ウェアラブル機器により撮影した長時間の体験記録映像に対して、脳波により個人の主観的な状態を反映させたインデキシングを行い、映像要約を行っている [2] では、体験映像に対して体験時の感情を日記として付与することが可能なインタフェースにより体験の想起を効率的にしている [9] では、本研究と同じく GPS と電子方位計よりカメラの位置と方向を取得している。本研究では、これらを用いてカメラと人気領域の位置関係による幾何学計算によりカメラ選択を行った [9] では、撮影者との距離とカメラの方向を指標として重みを付与し、それらの重みと滞在時間を考慮して地理オブジェクトの重要度を定義し、映像要約の際のシーン選択の判定に用いている。

5.3 メタデータに関する研究

[5] では、撮影時間、撮影位置と方向、地理オブジェクトと

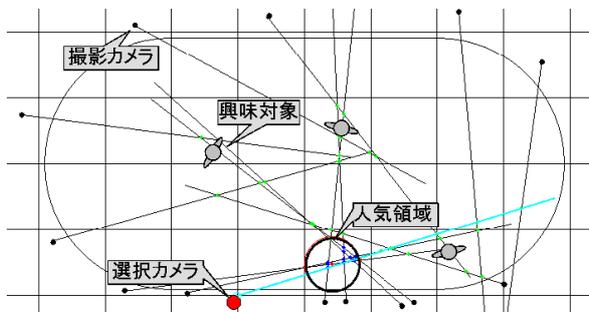


図 13 適切な人気領域が得られない場合

その重要度などを MPEG-7 ファイルに記述し、日付や地理オブジェクト、キーワードなどから所望の場面の検索を可能にしている。また、長時間の映像に対するダイジェスト作成を行っているが、特定の地理オブジェクトについてであり、時刻ごとに対象が変わる本研究とはこの点で異なる。

5.4 注視点検索に関する研究

本研究では、映像に対して位置・方向情報を付加しているが、[10] では、写真に空間メタデータを付与し、注視点を重視した自然言語による検索や、地図インタフェースによる検索を提案している。

6. 終わりに

本稿では、視線交差による人気領域の特定方法と人気領域に基づいた映像切り替え手法についての提案を行った。ここでは、今後の課題について述べる。

6.1 今後の課題

Java 2 SDK で作成したシミュレータを用いたところ、図 13 で示すように興味対象が複数ある時に、視線の向けられ方により、興味対象となるものが存在しない場所に人気領域を特定する場合がある。この場合、興味対象を映したカメラ選択を行うことができない。よって、このような時の対処法が必要となる。また、以下の3点についても課題とする。

● 最小包含楕円の適用

本論文では今回、人気領域を特定するのに最小包含円アルゴリズムを用いた。しかし、視線の交点の配置によっては、必要以上の領域を包含することがある。これを改善するために、最小包含楕円を適用したい。また楕円の場合、円とは異なり方向によって見える形状が異なるので、楕円の形状による選択方法も可能になる。

● クラスタリング

視線交差による点を c-平均法を用いてクラスター分けを行い、視線交差領域を求めた。しかし、 $10 \times 10m$ の領域で初期分割をしているため、完全なクラスタリングを行えているとは言えない。今後、複数の興味対象があった場合、興味対象ごとに交点を正確に分割できる手法を用いたい。

● 3次元へ適用

本論文で述べた視線交差に基づく人気領域の特定は2次元でのものである。しかし、スポーツイベントを考えた場合、見上げたり、高い位置から見下ろし撮影を行ったりする場合がある。

よって、3次元への適用を考えたい。

文 献

- [1] 相澤 清晴, 石島 健一郎, 椎名 誠, “ウェアラブル映像の構造化と要約:個人の主観と考慮した要約生成の試み”, 電子情報通信学会論文誌, Vol.86, No.6, pp.807-815, (2003).
- [2] 志村 将吾, 平野 靖, 梶田 将司, 間瀬 健二, “体験記録における日記を用いた感情記録インタフェース”, 情報処理学会研究報告(ヒューマンインタフェース), HI-115, pp. 61-68 (2005).
- [3] 中西 吉洋, 廣瀬 竜男, 田中 克己, “多視点映像データの概念モデリングと代表映像検索”, 情報処理学会論文誌, Vos.43, No.SIG 5(TOD 14), pp.54-65, (2002).
- [4] 秦 淑彦, 廣瀬 竜男, 中西 吉洋, 田中 克己, “カメラメタファによる多視点映像の検索”, 情報処理学会論文誌, Vos.42, No.SIG 4(TOD 9), pp.14-26, (2001).
- [5] 兵 清弘, 天笠 俊之, 吉川 正俊, 植村 俊亮, “MPEG-7 を利用したウェアラブルカメラ映像の索引付け手法”, 電子情報通信学会データ工学ワークショップ (DEWS2002) C2-17
- [6] 西岡 欣亮, 住野 優, 大西 真晶, 上田 真由美, 上島 紳一, “位置・情報付き映像の検索方式の提案”, 電子情報通信学会データ工学ワークショップ (DEWS2005) 3-A-02
- [7] 住野 優, 明石 康久, 上島 紳一, “3D 仮想空間を用いた動画蓄積システムの試作”, 電子情報通信学会データ工学ワークショップ (DEWS2004) 2-C-03
- [8] 竹前 嘉修, 大塚 和弘, 武川 直樹, “対面の複数人対話を撮影対象とした対話参加者の視線に基づく映像切替え方法とその効果”, 情報処理学会論文誌, Vol.46, No.7, pp.1752-1767, (2005).
- [9] 上田 隆正, 天笠 俊之, 吉川 正俊, 植村 俊亮, “位置情報と地理情報を用いたウェアラブルカメラ映像のダイジェスト作成”, 研究報告:データベースシステム, No.125, pp.177-184, (2001).
- [10] 藤田 秀之, 有川 正俊, 岡村 耕二, “注視点を考慮したデジタル写真の検索インタフェース”, 電子情報通信学会データ工学ワークショップ (DEWS2003) 6-P-03
- [11] 宮本 定明, クラスタ分析入門, 森北出版 (1999)
- [12] M. オーバマーズ, M. ドバーク, M. ファン クリベルド, O. シュワルツコップ, コンピュータ・ジオメトリ, 近代科学社 (2000)
- [13] 国土地理院 <http://www.gsi.go.jp/>
- [14] MPEG HOME PAGE <http://www.chiariglione.org/mpeg/>