

能動カメラ群を用いた三次元ビデオのセル分割追跡撮影方式

吉本 廣雅^{†a)} 山口 辰久^{†b)} 牧 淳人^{†c)} 松山 隆司^{†d)}

A Cell-based 3D Video Capturing Method with Active Cameras

Hiromasa YOSHIMOTO^{†a)}, Tatsuhisa YAMAGUCHI^{†b)}, Atsuto MAKI^{†c)}, and Takashi MATSUYAMA^{†d)}

あらまし 3次元ビデオ撮影のためには1)カメラパラメータが既知のカメラで2)対象表面をくまなく、かつ3)高い空間分解能で撮影する必要がある。従来の撮影方式はカメラの視野を固定することでこれらの条件を満たしていた。そのため撮影可能範囲が狭い領域に限定されるという問題があった。そこで我々は能動カメラを用いて広域移動対象を撮影するセル分割追跡撮影方式を提案する。この追跡撮影方式は撮影対象となる空間をセルと呼ぶ部分空間に分割し、セルを単位としてキャリブレーションを行うと共に離散的にカメラの視野を制御することで1から3の条件を満たす。実験により本方式で3次元ビデオの撮影ができることを実証した。

キーワード 3次元ビデオ, 追跡撮影, 能動カメラ, 多視点映像

1. はじめに

任意の視点から対象を鑑賞できる自由視点映像を生成するには、対象の三次元形状と表面テクスチャの時系列データが必要となる。このような時系列データは3次元ビデオと呼ばれ、近年盛んにその生成方法が研究されている[1]~[5]。これら一連の研究では、対象のシルエットに基づいて形状を計算するShape-from-Silhouette (SFS)、または視体積交差法[6]とも呼ばれる手法を基本的枠組みとして用いることが多い。つまり、対象を取り囲むように配置された多数のカメラを用いて多視点ビデオを撮影し、得られた動画像を、別途計測されたカメラ間の幾何学的・光学的関係を表すカメラパラメータを用いて解析することで、三次元ビデオを計算する方式が取られている。

このような多視点ビデオから三次元ビデオを計算する方式では、撮影に用いるカメラが以下の二つの条件を満たしていることが必要条件となる。

条件1 カメラパラメータが既知であること

条件2 対象表面をくまなく撮影していること
さらに表面テクスチャが精度良く得られるには、各カメラの

条件3 空間分解能が十分高いことも重要である。

これら3つの条件が同時に満たされる空間を三次元ビデオの撮影可能範囲と定めると、従来の三次元ビデオ撮影方式では撮影可能範囲が狭く、歩行する人物のような広域移動対象の三次元ビデオを撮影することが非常に困難になっていた。この原因は従来方式が

- 条件1を満たすために、カメラを固定する
- 条件2,3を満たすために、カメラ台数を増やす

という受動的な撮影方法を取った点にある。条件1で必要となるカメラパラメータはカメラの位置・撮影方向・画角に応じて変化するため、従来方式ではまずカメラの位置・撮影方向・画角を事前に決定し、その状態にカメラを固定した上でカメラキャリブレーションを行うことで、厳密にカメラパラメータを計測する。ところが、カメラを固定すると、条件2と条件3の間にトレードオフの関係が生じてしまう。なぜならカメラの画角の広さと空間分解能の間にはトレードオフの関係があり、たとえば条件2を満たすためにカメラの画角を広くすると空間分解能が低下するため条件3が満たせなくなってしまうからである。そこで従来の3次元ビデオ撮影方式では、カメラ台数を増やすことで

[†] 京都大学大学院情報学研究所, 京都府
Graduate School of Informatics, Kyoto University, Yoshida
Honmachi, Sakyo-ku, Kyoto, 606-8501 Japan

a) E-mail: yosimoto@vision.kuee.kyoto-u.ac.jp

b) E-mail: yamaguti@vision.kuee.kyoto-u.ac.jp

c) E-mail: maki@vision.kuee.kyoto-u.ac.jp

d) E-mail: tm@vision.kuee.kyoto-u.ac.jp

条件 2 と条件 3 の両立を計ってきた．ところがこの方式では撮影可能範囲が広がるにつれ膨大な台数のカメラが必要となり，物理的・コスト的な限界が生じてしまう．その結果，従来法では広域移動対象の三次元ビデオ撮影が非常に困難になっていた．

我々はこの問題を解決するために能動カメラを用いた三次元ビデオの追跡撮影方式を提案する．これは，各カメラの視線方向，ズームを動的に制御し，3 つの条件が満たされる領域を移動対象に追従させる撮影方式である．この能動的な撮影方式を用いると，カメラの利用効率を高めることができ，カメラ台数を増加させることなく，三次元ビデオの撮影可能範囲を広域化することができる．

しかしながら能動カメラを導入すると新たな問題が二つ生じる．一つ目の問題は能動カメラのキャリブレーションである．条件 1 を満たすには逐次状態が変化する能動カメラを高精度にキャリブレーションする必要がある．そのようなキャリブレーション方法としては文献 [7] などが提案されているが，これら手法を用いてもズームが大きく変化する能動カメラについて，レンズの焦点距離やレンズ歪みなどを高精度で頑健にキャリブレーションすることは難しい．二つ目の問題は，処理時間の問題である．追跡撮影では，複数台の能動カメラによる撮影画像から対象の三次元位置を計算し，その三次元位置に応じて条件 2 と条件 3 が満たされるように能動カメラを制御しなければならない．そして継続して対象を追跡し続けるには，これらの処理を実時間で行わなければならない．つまり

条件 4 条件 2, 3 を満たしながら実時間で対象を追跡していること

という新たな条件も満たす必要がある．

以上 4 つの条件を用いると，追跡撮影方式が解くべき問題は，対象の移動に応じて常に 4 つの条件を満たすように能動カメラの視野を制御する最適化問題として定式化できる．ところがその最適化には，4 つの条件間には複雑な相互関係があり，対象の 3 次元形状と各カメラの視野の組み合わせからなる広大な解空間の全探索が必須となる．その計算は条件 4 の観点からも極めて困難である．

そこで本論文は現実的な解法の一つとしてセル分割追跡撮影方式を提案する．これは対象の移動経路 L と条件 3 に関して最低限必要とする空間分解能 r とが与えられた場合について準最適な視野制御を行うものであり，以下の 4 つの処理で構成される．①まず L

に沿った三次元空間をセルと呼ぶ部分空間に分割し，②各セル毎に最低限 r の空間分解能を保証しつつセル全体を撮影するように能動カメラの設定値を定め，設定値毎にカメラキャリブレーションを行う．ここまでの処理によりセルを注視した状態にある能動カメラは条件 1, 3 を満足することが保証できる．換言すると，各能動カメラは対象の移動に伴ってセルから次のセルへと視線を変化させるが，視線移動中は条件 1, 3 を満足できないため，撮影された映像は三次元ビデオ生成には使用しない．③次に，移動経路上の対象の位置 x 毎に，(i) 条件 2 を満たすように x が含まれるセルを撮影するカメラ，(ii) 対象の移動に備え次のセルへと視線を移動させるカメラ，(iii) 対象を待ち受けるため次のセルを撮影するカメラ，という 3 つのモードをカメラ群に割り当てる．以上の処理は事前に行い，追跡撮影時は，④計測された対象の三次元位置に基づいて対象が含まれるセルを求め，上記③ で得られたカメラのモード割り当ての事前プランに従って各カメラを制御し，セルを注視しているキャリブレーション済のカメラ群によって撮影された多視点映像から三次元ビデオを生成する．

以下，本論文ではセル分割追跡撮影方式における問題の定式化，およびその解法であるアルゴリズムについて詳述し，実験の結果によりその有効性を示す．

2. 問題の定式化

セル分割追跡撮影方式における各処理過程を 4 つの条件の観点で整理すると，条件 1 はセル単位でカメラキャリブレーション，条件 3 はセル単位でのズーム値の計算，条件 4 はセル単位での離散的な追跡撮影処理により満足できる．このように条件 1, 3, 4 を満たすための処理は，それぞれセル単位での処理として定式化できる．

一方，条件 2 について考えると，これは 3 次元ビデオ生成の原理上常に満たすことができるとは言えない．なぜなら，対象表面上のカメラから撮影できる領域は対象の表面形状とカメラの視野の幾何関係によって大きく変動し，場合によっては対象表面上にカメラから撮影できない領域が生じるためである．

そこで本論文では現実的な解法として，条件 1, 3, 4 を保証した上で，出来るだけ条件 2 が満足されるように視野を最適化する方式を提案する．そのために対象とカメラの視野の幾何関係から条件 2 が満たされる度合を推定する評価関数を定義し，条件 2 を満たすた

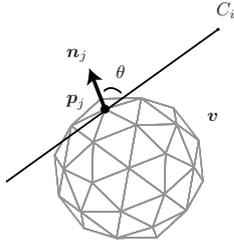


図 1 各要素の幾何関係
Fig. 1 Geometric relations

めの処理を最適化問題として定式化する．また条件 2 と条件 3 のトレードオフについては条件 2 を優先し，条件 3 に関しては空間分解能が要求された r 以上となることのみを保証し，それ以上の最適化は行わない．

以下評価関数とそれを用いた条件 2 に関する視野の最適化について詳述するにあたり，記号の定義を行う．まず三次元空間を \mathbb{S} ，条件 3 で要求される空間分解能を $r(\text{mm/pixel})$ と定義する．ある時刻において，カメラの内部・外部パラメータが既知の N_E 台のカメラ $\{C_i\}_{i=1, \dots, N_E}$ が空間中の領域 $v \subseteq \mathbb{S}$ を占めている対象を撮影している状態を考え，各 C_i について，条件 1 で必要となるカメラの内部・外部パラメータを E_i ，視野を \mathbb{F}_i とする．ここで \mathbb{F}_i は，最悪の場合でも r の空間分解能で画像として撮影できる空間中の領域とする．

これら $v, \{(E_i, \mathbb{F}_i)\}_{i=1, \dots, N_E}$ を用いて条件 2 に関する評価関数を次のように定義する． v 表面上の点群 $\{p_j\}_{j=1, \dots, N_j}$ を考える． v を三次元ビデオとして忠実に撮影するには，各 p_j が以下の二つの要求を同時に満たす必要がある．

- テクスチャの計測という観点からは，対象表面に対して垂直な方向から画像を撮影する．
- 視体積交差法による形状の計測という観点からは，対象表面に対して平行な方向から画像を撮影する．前述のように v の表面形状の複雑さや自己遮蔽を考慮すると，すべての点でこれら二つの要求を同時に満たすことは困難である．そこで，まず (E_i, \mathbb{F}_i) の状態にある C_i と p_j の間で，各要求が満たされている度合として，評価式をそれぞれ以下のように定義する．

$$q_t(p_j, (E_i, \mathbb{F}_i)) = \text{visible}(p_j, (E_i, \mathbb{F}_i)) \cos \theta$$

$$q_s(p_j, (E_i, \mathbb{F}_i)) = \text{visible}(p_j, (E_i, \mathbb{F}_i)) |\sin \theta|$$

これらの評価式は，その値が大きくなるほど，より要求が満たされていることを示す．ここで 図 1 に示すように θ は点 p_j における対象表面の法線ベクト

ル n_j と， p_j から C_i の投影中心に向かうベクトルの成す角である．また， $\text{visible}(p, (E_i, \mathbb{F}_i))$ は点 p を (E_i, \mathbb{F}_i) の状態にある C_i で観測可能な場合 1，不可能な場合 0 となる関数とする．これらの評価式が示すように，一台のカメラだけでは二つの要求を同時に満たすことはできない．しかしカメラが複数台あることを考えると，いずれかのカメラでそれぞれの要求が満たされれば，3次元ビデオとして忠実に撮影できたことになる．そこで各点 p_j について各要求が満たされている度合を次の式で表す．

$$q'_t(p_j, \{(E_i, \mathbb{F}_i)\}_{i=1, \dots, N_E}) = \max_i q_t(p_j, (E_i, \mathbb{F}_i))$$

$$q'_s(p_j, \{(E_i, \mathbb{F}_i)\}_{i=1, \dots, N_E}) = \max_i q_s(p_j, (E_i, \mathbb{F}_i))$$

更に次の式により対象表面上で最も撮影できていない点の評価値を計算する．

$$Q_t(v, \{(E_i, \mathbb{F}_i)\}_{i=1, \dots, N_E}) = \min_j q'_t(p_j(t), \{(E_i, \mathbb{F}_i)\}_{i=1, \dots, N_E}) \quad (1)$$

$$Q_s(v, \{(E_i, \mathbb{F}_i)\}_{i=1, \dots, N_E}) = \min_j q'_s(p_j(t), \{(E_i, \mathbb{F}_i)\}_{i=1, \dots, N_E}) \quad (2)$$

式 (1), (2) による評価値が共に大きくなる程，より対象表面の広い範囲が忠実に撮影できたことになる．そこで条件 2 が満たされる度合を推定する評価関数を式 (3) で定義する．

$$Q(v, \{(E_i, \mathbb{F}_i)\}_{i=1, \dots, N_E}) = Q_t(v, \{(E_i, \mathbb{F}_i)\}_{i=1, \dots, N_E}) Q_s(v, \{(E_i, \mathbb{F}_i)\}_{i=1, \dots, N_E}) \quad (3)$$

式 (3) を用いると $v, \{(E_i, \mathbb{F}_i)\}_{i=1, \dots, N_E}$ から条件 2 が満たされている度合を計算できる．次に対象が広域を移動する場合を考える．まず時刻 t に対象が存在する領域を $\mathbb{V}(t) \subseteq \mathbb{S}$ とし，能動カメラのカメラパラメータと視野をそれぞれ $E_i(t), \mathbb{F}_i(t)$ とする．さらに，視野制御のための指令値からなるベクトルを制御値 $e_i(t)$ と表現すると， $E_i(t) = \bar{E}_i(e_i(t))$ ， $\mathbb{F}_i(t) = \bar{\mathbb{F}}_i(e_i(t))$ となる関数が定義できる．これら定義を用いると，条件 2 に関する処理，すなわち対象表面をできるだけ忠実に撮影するための処理は， $\mathbb{V}(t)$ に応じて $Q(v, \{\bar{E}_i(e_i(t)), \bar{\mathbb{F}}_i(e_i(t))\}_{i=1, \dots, N_E})$ を最大にするように $\{e_i(t)\}_{i=1, 2, \dots, N_E}$ を最適化する処理と定式化できる．

3. セル分割追跡撮影アルゴリズム

2節で述べたように、セル分割追跡撮影方式の処理アルゴリズムは、条件1, 3, 4に関するセル単位の計算と、条件2に関する式(3)を用いた最適化計算として定式化できる。しかし各条件間には複雑な相互関係があり、その計算には全解探索が必須である。また条件2は近似的にしか計算することができない。そこで本論文では準最適視野を求める近似解法を考え、次のようにアルゴリズムを設計する。まず条件2と条件3のトレードオフの関係を考え、条件1, 条件3を確実に保証するように三次元空間をセルへと分割し、セル毎に制御値を定める。その後条件4を保証しつつ条件2をできるだけ満足するように視野の準最適化を行う。この手法は条件1, 条件3の観点のみで先にセルおよびセル毎の制御値を決定してしまうため最適解を求めることは不可能である。しかし条件1, 3, 4を満足することが保証できる点で、現実的な解法である。

詳細を説明するにあたり、アルゴリズムへの入力を以下のように定義する。本論文ではこれらのデータ群を、それぞれシナリオ、リソースと分類する。

- シナリオ

- $\{L(l) \in \mathbb{S} | l \in [0, L]\}$: 対象の移動経路。長さ L の曲線とし、始点から距離 l だけ進んだ点の3次元位置を $L(l)$ とする。

- $\hat{V}(x) \subseteq \mathbb{S}$: 対象の形状モデル。対象重心の3次元位置を $x(t)$ とすると $\forall t(t) \subseteq \hat{V}(x(t))$ となる。

- V_{\max} : 対象の最高移動速度。
- r : 条件3に関して要求される空間分解能。

- リソース

- $\tau_i(e_i, e'_i)$: カメラの動特性。 C_i の状態をある状態 e_i から次状態 e'_i へ変更する際に要する時間を求める関数。

- $K_i(v)$: 点の集合 v に対して条件3を満足する e_i を計算する関数。条件3を満足する e_i が存在しない場合は $K_i(v) = \phi$ とする。

- $\text{visible}(p, (E_i, F_i))$: 可視判定を行う関数。
- τ_{cam} : 多視点ビデオの撮影周期。
- τ_{proc} : 対象の3次元位置 x の計測に要する時間。

なお $K_i(v)$ の計算アルゴリズムは、使用する能動カメラの種類・特性に従い e_i を一意に決定するように与える。例えば、能動カメラとして準視点固定カメラ [8] を用いる場合、 e_i はパン・チルト・ズーム・フォーカスの4次元ベクトルとなり、 $K_i(v)$ の計算アルゴリズム

は次のようになる。まず v 中でカメラの投影中心から最も遠い点の空間分解能が r となるようにズーム値を定め、次にフォーカス値を v 全体がカメラの被写界深度に含まれるように定め、最後に v 全体を画像中央に撮影するようにパン・チルト値を設定する。また一つでも設定不可能な制御値があれば $K_i(v) = \phi$ とする。

セル分割追跡撮影アルゴリズムは上記のシナリオ・リソースを入力とし、以下の5段階の処理で、広域移動対象の3次元ビデオを出力する。

- (1) セル分割
- (2) カメラキャリブレーション
- (3) スケジューリング
- (4) 追跡撮影
- (5) 3次元ビデオ生成

以下、各処理について詳述する。

3.1 セル分割

本節では、各セルが満たすべき条件、セルの定義を述べ、その計算アルゴリズムについて述べる。

まず2節で定式化したように条件1, 3を満足するための処理はセル単位で行われる。ここで条件3は各セルに含まれるすべての点がいずれのカメラにおいても空間分解能 r 以上で撮影できれば満足できるので、各セルは

$$\mathbb{C}_k \subseteq \bar{\mathbb{F}}_i(\bar{e}_i(k)) \quad (4)$$

となる $\bar{e}_i(k)$ が存在するように決定されなければならない。次に条件1が必要とするカメラキャリブレーション作業を考えると、セルの総数 N_k は少ないほど良い。換言すると個々のセルは大きいほど良い。

セル分割アルゴリズムは、移動経路に沿った3次元空間を、上記二つの条件を満たすようにセルへと分割する。詳細を記述にあたりセルを次のように定義する。移動経路を $\Delta l = \tau_{\text{cam}} V_{\max}$ 単位で分割した N_f 個の小区間 $\{f(n) = \{L(l) | l \in [\Delta ln, \Delta l(n+1)]\}\}_{n=0, \dots, N_f-1}$ を考え、各小区間 $f(n)$ に対象が位置するとき対象が存在する可能性のある領域を処理単位 $\mathbb{W}(n)$ として次式で定義する。

$$\mathbb{W}(n) = \bigcup_{x \in f(n)} \hat{V}(x) \quad (5)$$

さらに、セル \mathbb{C}_k を $\mathbb{W}(n)$ の和集合

$$\mathbb{C}_k = \bigcup_{n=c_k}^{c_{k+1}} \mathbb{W}(n)$$

```

k ← 1
c_k ← 0
while c_k < N_f - 1 do
  for i = 1, 2, ..., N_E do
    n_i ← c_k
    while n_i < N_f - 1 and K_i (∪_{n=c_k}^{n_i+1} W(n)) ≠ ∅ do
      n_i ← n_i + 1
    end while
    if n_i = c_k then
      return ∅
    end if
    e_i(k) ← K_i (∪_{n=c_k}^{n_i} W(n))
  end for
  c_{k+1} ← min{n_i | i=1, 2, ..., N_E}
  k ← k + 1
end while
N_k ← k - 1
return {N_k, {c_k}, {e_i(k)} | k=1, 2, ..., N_k}

```

図2 セル分割アルゴリズム
Fig. 2 Cell dividing algorithm

として定義する。ただし c_k は $0 \leq c_k < c_{k+1} \leq N_f - 1$ を満たす整数である。

図2にセル分割アルゴリズムを示す。このアルゴリズムは式(4)を満たすセル分割が存在する場合は $\{N_k, \{c_k\}, \{e_i(k)\} | k=1, 2, \dots, N_k\}$ を、存在しない場合には \emptyset を返す。ここで \emptyset は、与えたシナリオ・リソースでは条件3を満たす撮影ができないことを表す。そこでセル分割アルゴリズムが \emptyset を出力した場合セル分割追跡撮影方式は撮影不可能という判定を行い、処理を完了する。

なお式(4)が示すように、 C_k を注視したカメラは C_k 外の領域も撮影できる場合がある。換言すると C_k を注視したカメラは C_{k-1} や C_{k+1} の一部でも条件1、条件3を満たした撮影が可能となる。このような領域は C_k のみの撮影では余分であるが、本アルゴリズムはこのような冗長な領域を敢えて残しておくことで条件2に関するスケジューリングでより良い視野切り換えスケジュールが生成できるようにしている。

3.2 カメラキャリブレーション

条件1を満たすには、セル毎にカメラキャリブレーションを行うことで、 $\bar{E}_i(e_i(k))$ を測定するだけでよい。セル単位では C_i は能動カメラではなく固定カメラと見なすことができるため、固定カメラ群に対する従来のキャリブレーション手法を用いて高精度なカメラキャリブレーションが可能となる。

3.3 スケジューリング

前述のようにスケジューリングは条件2に関する処理となる。条件2を満たすには、対象の移動に応

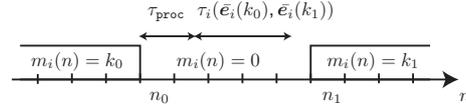


図3 $\tau_i(e, e')$, τ_{proc} と $m_i(n)$ の関係
Fig. 3 Relation among $\tau_i(e, e')$, τ_{proc} and $m_i(n)$

じて式(3)を最大にする $\{e_i(t)\}_{i=1, \dots, N_E}$ を計算しなければならない。本アルゴリズムでは、そのような $\{e_i(t)\}_{i=1, \dots, N_E}$ の計算を以下に述べるように小区間単位で行う。

まず撮影対象がある小区間 $f(n)$ 内に存在する場合を考える。そしてあるカメラ C_i が $f(n)$ 内に存在する対象を撮影する際の $e_i(t)$ を表現するために、

$$m_i(n) = \begin{cases} k & e_i(k) \text{ を用いて } C_k \text{ を注視している状態} \\ 0 & \text{注視先セルを切替中の状態} \end{cases}$$

となる変数 $m_i(n)$ を定義する。 $\tau_i(e, e')$ と τ_{proc} を考慮すると、各 C_i に対して $m_i(n) = 0$ となる区間が次のように計算できる。時刻 t に C_{k_0} を注視していたカメラ C_i が、対象が $f(n_0)$ に来たときに C_{k_1} へと注視先を切り換える場合を考える。 $f(n_0)$ に来た対象を C_i が撮影してから対象位置を計測し、切り換えを開始するまでには τ_{proc} の時間が掛かる。更に C_i が切り換えを完了するまでには $\tau_i(e_i(k_0), e_i(k_1))$ の時間が掛かる。つまり、この切り換えには合計 $\Delta T = \tau_{\text{proc}} + \tau_i(e_i(k_0), e_i(k_1))$ の時間が掛かることになる。この間に対象は最大で $V_{\text{max}} \Delta T$ の距離を進む。よって図3に示すように n_0 から $n_1 = n_0 + \lceil \frac{V_{\text{max}} \Delta T}{\Delta T} \rceil$ までの小区間で、 C_i は注視先を切り換え中の状態、つまり $m_i(n_0) = 0, \dots, m_i(n_1 - 1) = 0$ となる。

さらに、セルを注視している状態にあるカメラ群を

$$g(n) = \{i | m_i(n) \neq 0\}$$

と表す。 $g(n)$ に属さないカメラはいずれのセルも注視しておらず、条件1、3を満たすことができない。そのため、条件2に関する評価式は、条件1と3を満たす $g(n)$ に属するカメラのみで

$$Q(\mathbb{V}(t), \{\bar{E}_i(e_i(m_i(n))), \bar{F}_i(e_i(m_i(n))) | i \in g(n)\}) \quad (6)$$

と計算しなければならない。

次に、式(5)と $\hat{\mathbb{V}}$ の定義より $\mathbb{V}(t)$ と $\mathbb{W}(n)$ には

$\mathbb{V}(t) \subseteq \mathbb{W}(n)$ という関係がある．そこで， $\mathbb{W}(n)$ に対して式 (3) を最大にする $\{\bar{e}_i(m_i(n)) | i \in g(n)\}$ は， $\mathbb{V}(t)$ に対しても式 (3) を最大にすると考え，対象が $f(n)$ に存在する場合の評価値を以下のように計算する．

$$Q(\mathbb{W}(n), \{(\bar{E}_i(\bar{e}_i(m_i(n))), \bar{F}_i(\bar{e}_i(m_i(n)))) | i \in g(n)\}) \quad (7)$$

移動経路全体でできるだけ条件 2 を満たすには，各小区間 $f(n)$ で式 (7) が最大となるように $\{m_i(n)\}_{i=1, \dots, N_E, n=0, \dots, N_f}$ を最適化すれば良い．そこで，次に示す組み合わせ最適化問題を解く．

変数 $\{m_i(n)\}_{i=1, \dots, N_E, n=0, \dots, N_f}$

目的関数

$$\min_{n=0, \dots, N_f-1} Q(\mathbb{W}(n), \{(\bar{E}_i(\bar{e}_i(m_i(n))), \bar{F}_i(\bar{e}_i(m_i(n)))) | i \in g(n)\}) \quad (8)$$

制約条件 n' で注視先を C_{k_0} から C_{k_1} に切り換える場合には $n' + 1$ から $n' + \lceil \frac{V_{\max}(\tau_{\text{proc}} + \tau_i(\bar{e}_i(k_0), \bar{e}_i(k_1)))}{\Delta t} \rceil$ までの n について $m_i(n) = 0$ となる

なおこの最適化には全解探索が必須である．なぜなら $g(n)$ が変数 $\{m_i(n)\}$ の組み合わせにより変化するためである．ここで $\{m_i(n)\}_{i=1, \dots, N_E, n=0, \dots, N_f}$ の組み合わせを考えると，その解の候補の数は $O(N_k^{N_f N_E})$ になる．4.2 節で示す例のように，我々の想定するシナリオとリソースでは N_f は 100 以上に， N_E は 10 以上になる．このようにセルを導入し，問題を部分問題に分割してもなお，解の候補の数は膨大であり，全解探索による解法は計算量の観点から非現実的である．そこで本アルゴリズムは，対象が C_k から C_{k+1} へ移動する際に，全てのカメラが 1 度だけ必ず C_k から C_{k+1} に注視先を切り換えるという制約を加え，問題を隣接セル対 (C_k, C_{k+1}) 間の部分問題へと分割する．ここでその分割方法はそれ自身が困難な最適化問題となる．そこで本アルゴリズムは各セルの中央で経路を分割した区間 $[A_k, A_{k+1}] (k = 1, \dots, N_k - 1)$ を考え，その区間毎に部分問題へと分割する．ただし

$$A_k = \begin{cases} c_1 & (k = 1) \\ \lfloor \frac{c_k + c_{k+1}}{2} \rfloor & (2 \leq k \leq N_k - 1) \\ c_{N_k+1} & (k = N_k) \end{cases}$$

である．なおこの分割方法は次に述べる経験則に基

づいている．まず式 (8) を最大化するには一般に対象を注視するカメラ台数が多いほうが良い．ここで対象が区間 $[A_k, c_{k+1} - 1]$ 中にある際の各カメラの状態を考えると，各カメラは C_k を注視するカメラ， C_{k+1} に先回りするカメラの二つに分類できる．同様に区間 $[c_{k+1}, A_{k+1}]$ では C_k を注視する後追いのカメラ， C_{k+1} を注視するカメラの二つに分類できる．そしてこれら先回りや後追いをするカメラが同時に注視先を切り換える頻度が高くなるほど，対象を注視するカメラ台数は減少する．そこで本アルゴリズムでは， C_k の中央に A_k を設定し $[c_k, A_k - 1]$ ， $[A_k, c_{k+1} - 1]$ の長さを可能な限り長くすることで，カメラが同時に注視先を切り換える頻度の低減を図る．

以上まとめると，条件 2 に関する処理は，次に示す組み合わせ最適化問題となる．

変数 $\{m_i(n)\}_{i=1, \dots, N_E, n \in [A_k, A_{k+1}]}$

目的関数

$$\min_{n \in [A_k, A_{k+1}]} Q(\mathbb{W}(n), \{(\bar{E}_i(\bar{e}_i(m_i(n))), \bar{F}_i(\bar{e}_i(m_i(n)))) | i \in g(n)\})$$

制約条件 各カメラは 1 回だけ注視先を C_k から C_{k+1} に切り換えるものとし，その切り換え開始位置を $n'_i (A_k < n'_i < A_{k+1} - \tau_i(\bar{e}_i(k), \bar{e}_i(k+1)))$ とすると $m_i(n)$ は

$$m_i(n) = \begin{cases} k & (A_k \leq n \leq n'_i) \\ 0 & (n'_i < n \leq n'_i + \tau_i(\bar{e}_i(k), \bar{e}_i(k+1))) \\ k+1 & (n'_i + \tau_i(\bar{e}_i(k), \bar{e}_i(k+1)) < n \leq A_{k+1}) \end{cases}$$

となる．

部分問題の解候補の数は $O((A_k - A_{k-1} + 1)^{N_E})$ となる．ここで $A_k - A_{k-1}$ は平均するとおよそ $\frac{N_f}{N_k}$ となる．したがって全体の計算量は全解探索の場合で $O(N_k (\frac{N_f}{N_k})^{N_E})$ にまで削減できる．本アルゴリズムではこの問題を GA を用いて解く．

3.4 追跡撮影

事前のスケジューリングにより対象の三次元位置に対応する準最適な視野割り当て $\{\bar{e}_i(k)\}$ と $m_i(n)$ がすでに求まっている．そのため追跡撮影は，対象の三次元位置 x の計測と，計測結果に応じたカメラの視野制御を並行して処理するだけで実現できる．

図 4 は追跡撮影処理の概要を時間軸に沿ってまとめたものである．追跡撮影は，首振りカメラが接続され

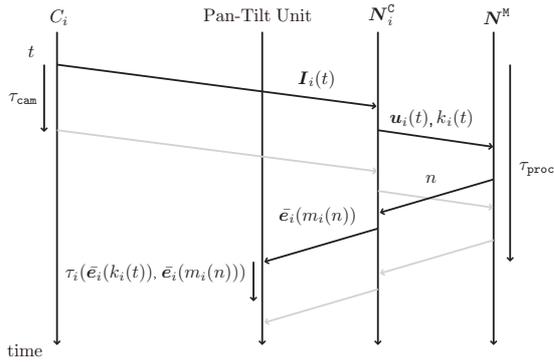


図 4 追跡撮影処理の概要
Fig. 4 Overview of tracking process

た N_E 台のカメラノード $\{N_i^C\}$ と 1 台のマスターノード N^M が撮影対象の三次元位置 $x(t)$ を共有するために通信を行いながら動作する。

対象の三次元位置の計測は以下の手順で行われる。以下、時刻を t と表記する。 t は全ノード間で十分な精度で同期しているものとする。

N_i^C 側の処理

以下の処理を周期 τ_{cam} で繰り返す。

(1) 同期撮影を行うことで時刻 t の撮影画像 $I_i(t)$ を得る。

(2) カメラ C_i がいずれかのセルを注視した静止状態であれば、画像平面上での対象の重心位置 $u_i(t)$ を追跡する。追跡には Condensation アルゴリズム [9] を用いる。

(3) $u_i(t)$ が計測できた場合は、 $\{t, u_i(t), k_i(t)\}$ の組を N^M へ送信する。ここで $k_i(t)$ は C_i が注視していたセルの番号とする。

N^M 側の処理

以下の処理を周期 τ_{cam} で繰り返す

(1) $\{\{t, u_i(t), k_i(t)\} | i = 1, \dots, N_E\}$ の組のうちある閾値 N_{Th} 個以上の組が受信できた場合は、 $u_i(t)$ 、 $\bar{E}_i(e_i(k_i(t)))$ を用いてステレオの原理で対象重心の三次元位置 $x(t)$ を計算する。

(2) $x(t)$ が計算できた場合は、 $x(t)$ に対応する L 上の小区間 n を求める。求めた n は、すべての $\{N_i^C\}$ へ送信する。

カメラの視野制御は以下の手順で行われる。

N_i^C 側の処理

以下の処理を逐次繰り返す

(1) 対象を含む小区間 n を新たに受信した場合は、

予め計算した $m_i(n)$ に応じて、対応する $e_i(m_i(n))$ へ能動カメラの状態を更新する。

(2) C_i がいずれかのセルを注視した静止状態であれば、撮影画像と注視したセル $C_{k_i(t)}$ の組として、 $\{t, I_i(t), k_i(t)\}$ を保存する。

3.5 3次元ビデオ生成

以上のアルゴリズムを用いると、 N_i^C が保存した $\{t, I_i(t), k_i(t)\}$ から、4つの条件を満たした $\{I_i(t), \bar{E}_i(e_i(k_i(t)))\}$ の時系列が得られ、3次元ビデオの生成が可能となる。

4. 実験

実験では、約 8m 四方の撮影スタジオに能動カメラとして 23 台の能動カメラを設置した。これら能動カメラは、ズーム・フォーカス制御が可能なカメラ (SONY DWF-VL500) をパン・チルト制御のためにカメラ雲台 (Directed Perception Inc. PTU-46) と組み合わせ作成した準視点固定型パン・チルト・ズームカメラ [8] である。図 5 に撮影スタジオと各カメラの配置関係を示す。図中の黒点と数字は能動カメラ C_i を表し、斜線領域はスタジオ中の座標 $(-1500, 0, 0)$ にて直立する人物を撮影する際の C_{16} の視野 \mathbb{F}_{16} の例を表す。このように一般的な能動カメラ・撮影スタジオを用いる場合、各カメラの視野つまり条件 3 を満たす撮影が可能となる領域はスタジオ内の一部分に限定される。

以下本節ではこのスタジオを用いて行った実験について詳細を述べる。なお提案手法が必要とするリソースとしては、能動カメラの動特性を計測することで $\tau_i(e_i, e'_i)$ を与え、 $K_i(v)$ は 3 節で述べたアルゴリズムを与えた。visible($p, (E_i, \mathbb{F}_i)$) については、本スタジオではカメラから見て対象が他の物体に遮蔽されることは無いため、自己遮蔽のみを計算する関数を与えた。また計算機としては、 N_i^C 用計算機を 23 台、 N^M 用計算機 1 台、合計 24 台の計算機 (Xeon 3.6GHz) を用意した。

以下の本節ではこの実験環境の下で行った 3 つの実験について詳細を述べる。

4.1 評価関数の定量評価

まず我々が設計した評価関数について、対象表面を撮影する度合いが正しく推定でき、条件 2 に関する評価式として適切であることを示す実験を行った。

2 節で述べたように、対象の真の形状 \mathbb{V} と n 台の能動カメラの視野が与えられると、 \mathbb{V} 表面上でどのカメラからも撮影できない点の集合が計算できる。ここ

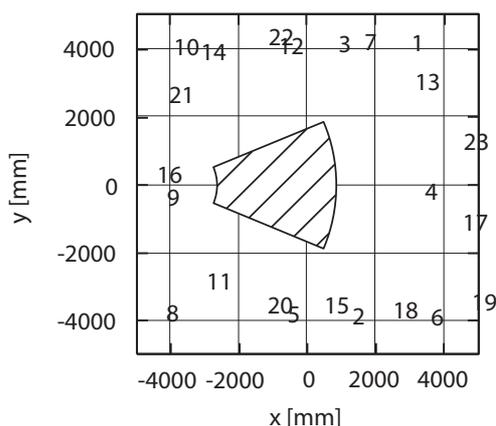


図 5 撮影スタジオとカメラの配置状況
Fig. 5 Studio and camera arrangement

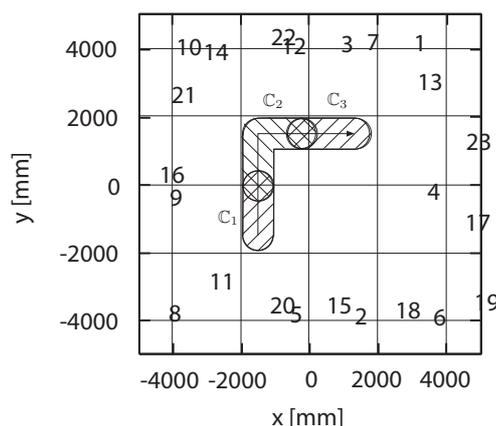


図 7 シナリオ 1 の移動経路とセル分割結果
Fig. 7 Path and computed cells for scenario 1

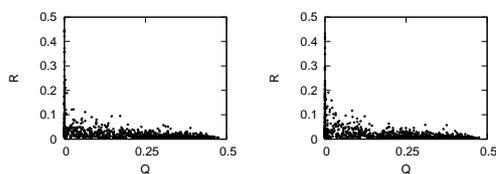


図 6 (Q, R) の分布
Fig. 6 Distribution of (Q, R)

で、全表面に対してそのような撮影できない領域が占める割合を R とし、その際の式 (3) の値を Q と表記すると、条件 2 の観点からは R の値が小さくなるほど、より条件 2 が満たされていると言える。そこで対象の位置・形状と、使用するカメラ台数・その視野の組を一つの撮影パターンとし、ランダムに 1000 通り生成した撮影パターンについて、シミュレーションによる仮想的な撮影を行い Q と R の関係を測定した。

なお各撮影パターンは、人物の 3 次元形状データを \mathbb{V} として与え、まずランダムに対象の位置を決定し、次に使用するカメラ台数をランダムに $n(1 \leq n \leq N_E)$ 台と決め、最後に $K_i(v)$ を用いて視野が設定できるカメラの中からランダムに n 台撮影に使用するカメラを選択することで生成した。また Q の計算で使用する対象の仮の形状 $\hat{\mathbb{V}}$ は、4.2 節で後述する理由により \mathbb{V} 中の上半身のみを内包する球を使用した。

実験は 2 種類の異なる形状データに対して行った。結果を図 6 に示す。図中の横軸は Q 、縦軸は R を表し、黒点はそれぞれが一つの撮影パターンに対応する。まず幾何関係や形状の差異に応じて条件 2 が満たされる度合は大きく変化するため、 Q に対する R の値は

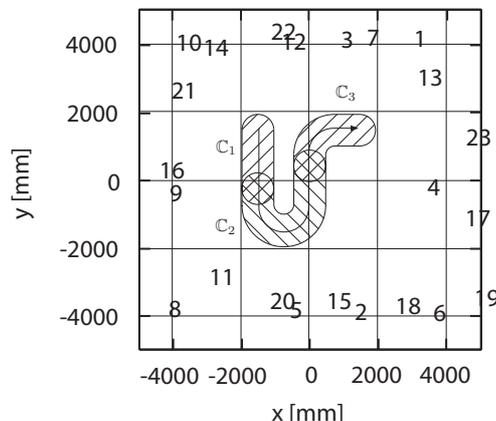


図 8 シナリオ 2 の移動経路とセル分割結果
Fig. 8 Path and computed cells for scenario 2

分散が大きい。また原理的に床面に接する足の裏などカメラから撮影できない面が存在するため、すべての面がくまなく撮影できる場合つまり R が 0 となる場合も存在しない。しかしながら R の最大値に着目すると、 Q の値が大きくなるほど R の最大値は減少している。つまり式 (3) を用いると真の形状が未知の物体に対して条件 2 が満たされる度合の下限を計算できる。よってその下限がより大きくなるように各能動カメラの視野を最適化すると、より条件 2 を満足する多視点ビデオが撮影できると言える。

4.2 実証実験

次に提案手法の有効性を実証するために、広域を歩行する人物を撮影対象とした 3 次元ビデオの追跡撮影を行った。

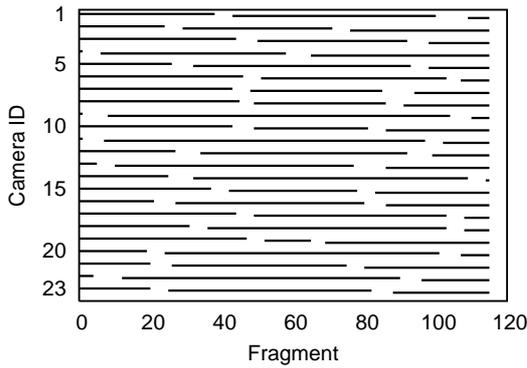


図 9 シナリオ 1 スケジューリング結果
Fig. 9 Optimized schedule for scenario 1

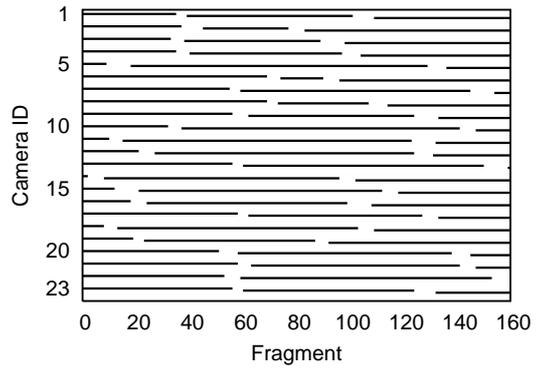


図 10 シナリオ 2 スケジューリング結果
Fig. 10 Optimized schedule for scenario 2

セル分割追跡撮影アルゴリズムへの入力として、これらのカメラに関する $K_i (i = 1, \dots, N_E)$, $\tau_i(e_i, e'_i)$, τ_{cam} をリソースとして与え、またシナリオを二通り用意した。図 7, 図 8 の矢印付き実線はそれぞれ対象の移動経路 L を表す。対象人物の身長は約 1.8m と想定し、シナリオはこれら L 上を人物が最速で 0.5m/sec で自由に進むと設定した。要求空間分解能は $r = 8[\text{mm}/\text{pixel}]$ とした。

まず対象人物全身の 3 次元ビデオを撮影するために、 $\tilde{V}(x)$ として半径 0.45m の球 2 個を積み重ねた 3 次元形状を用い、セル分割アルゴリズムを適用した。結果、我々の用意したスタジオのカメラ群では、与えたシナリオに対して条件 3 を満たすことが保証できないと判断された。これは、実験で使用したスタジオ・カメラでは、人物の位置によってはそもそも条件 3 を満たす撮影が行えない状況が生じるためである。例えば座標 $(-1500, 0, 0)$ に直立している人物を C_{16} で撮影する場合など撮影対象がカメラの近くにいる場合は、カメラの画角の制約上人物全身を撮影することができず、条件 3 を満たせない。

そこで以下の実験では、対象人物の上半身のみの 3 次元ビデオ生成を保証した追跡撮影を行い、上半身の 3 次元ビデオ生成結果を検証した。まず $\tilde{V}(x)$ は対象上半身を内包するような半径 0.45m の球とし、その他のリソース、シナリオは前述のものをそのまま利用した。計算には Xeon 3.6GHz の計算機を用い、所要時間は 5 秒であった。結果両シナリオにおいて、それぞれ 3 つのセルが得られた。得られたセル C_k を図 7, 図 8 中に網掛け領域として示す。そしてこれら C_k 毎にカメラキャリブレーションを行うことで、カメラパ

ラメータ $\{\bar{E}_i(\bar{e}_i(k))\}_{i=1, \dots, N_E}$ を測定した。

次に各シナリオについて、得られたセルを基にカメラの視野制御のスケジューリングを行った。計算では個体数を 2000, 世代数を 2000 とした。計算時間はそれぞれ約 80 時間であった。結果を図 9, 図 10 に示す。図中の縦軸は各カメラ C_i を、横軸は移動経路上での対象の位置 n を表す。図中の横線は n に応じたカメラの状態 $m_i(n)$ を表し、線分はカメラがいずれかのセルを注視した静止状態にある区間、空白部分は次のセルに視野を切り換えるためにカメラが運動状態にある区間を表す。

これらセル分割結果とスケジューリング結果を基に追跡撮影を行い、三次元ビデオ生成を行った。図 11 に撮影した多視点ビデオを、図 12 に生成した三次元ビデオの表示例を示す。このように多視点ビデオの全フレームで上半身が撮影されており、提案手法を用いることで上半身の高精細な 3 次元ビデオが生成できることを実証した。

なお対象の下半身に関しては図 12 に示すように一部のカメラで下半身を撮影できないフレームが生じた。これはカメラの画角には機械的な制限があり、対象がカメラに近づき過ぎた場合に対象全体を同時に撮影できなくなるためである。その結果、生成された 3 次元ビデオではテクスチャ・形状情報に欠損が生じる。この現象は図 12 の 3 次元ビデオの表示例でも生じており、足元の部分が欠けている。

実証実験の結果をまとめると、提案手法は以下に述べる二つの事を実現したと言える。まず提案手法を用いると広域移動対象の 3 次元ビデオ生成が可能である。また実験の過程でセル分割アルゴリズムが対象全身の



C_{12} による撮影画像 C_{20} による撮影画像
 Images captured by C_1 Images captured by C_6

図 11 15,54,104,151 フレーム目の撮影画像
 Fig.11 Captured images at 15, 54, 104 and 151 frame

3次元ビデオ生成は保証できないと判定したように、提案手法を用いると追跡撮影前に4つの条件が満足できない状況の検出が可能となる。

4.3 定量評価

次に撮影結果の定量評価を行うためにカメラの利用効率の観点から従来法との比較を行った。そのために各時刻での視野配置に対して、以下に示す二つの利用率を定義した。

$$\text{視点利用率} = \frac{\text{対象を撮影しているカメラ台数}}{\text{全カメラ台数}}$$

$$\text{画素利用率} = \frac{\text{撮影画像中で対象が占める画素数}}{\text{全画素数}}$$

双方の利用率が高い視野配置ほど、3次元ビデオの生成に利用できる情報の多い画像が撮影できるため、復元結果である3次元ビデオの精度は向上する。しかし従来の固定カメラを用いる撮影方式では、撮影対象が

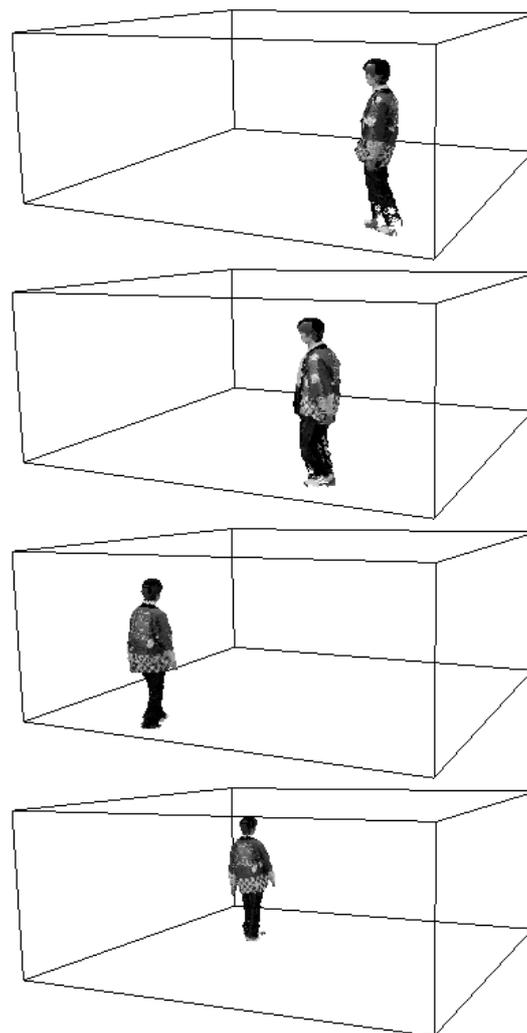


図 12 3次元ビデオの表示例
 Fig.12 An example of rendered 3D Video

広域を移動する場合、二つの利用率にトレードオフの関係が生じる。そこで、従来の固定カメラによる撮影方式として、視点利用率を優先した視野配置による撮影方式と画素利用率を優先した視野配置による撮影方式の2方式を用意し、セル方式追跡撮影と利用率の比較を行った。なお、これら利用率は撮影時の対象の位置や形状により変動する性質を持つ。そこで同一のシーンに対する撮影画像を得るために、まず提案手法により追跡撮影を行い3次元ビデオを生成した。従来法による撮影画像は、この3次元ビデオからシミュレーションにより合成した。

具体的には、図7、図8の二つのシナリオに対して、

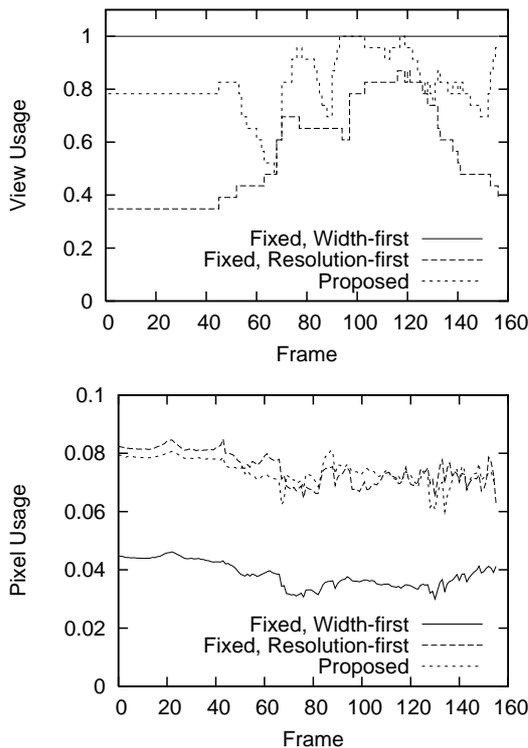


図 13 シナリオ 1 での視点利用率と画素利用率
Fig. 13 View and pixel usage in scenario 1

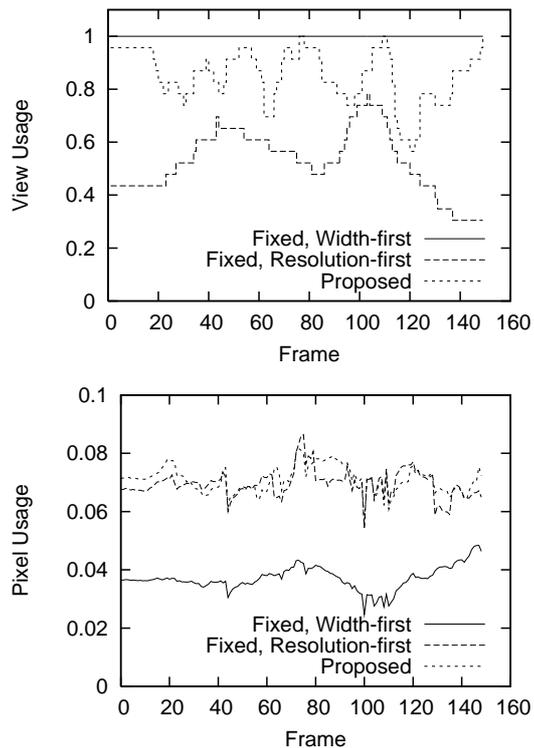


図 14 シナリオ 2 での視点利用率と画素利用率
Fig. 14 View and pixel usage in scenario 2

視点利用率を優先した固定カメラの視野配置，画素利用率を優先した固定カメラの視野配置をそれぞれ仮想的に用意した．前者の視点利用率優先の視野配置では，各カメラの視野が対象の通過領域全体を収めるように広角に設定した．一方後者の画素利用率優先の視野配置では，セル分割追跡撮影で設定したセルを流用し，各カメラはいずれかのセルを注視し続けるように視野を配置した．ここで 23 台ある各カメラが注視するセルの組み合わせは 3^{23} 通りもある．そこでランダムに 10^{12} 通りのセルの組み合わせを生成し，その組み合わせの中で式 (3) を最大にするセルの組み合わせを最善の視野配置として採用し，画像を合成した．

各利用率の計測結果を図 13，図 14 に示す．従来の固定カメラによる撮影方式では二つの利用率の間にトレードオフの関係がある．たとえば視点利用率を優先した固定カメラの視野配置を行うと，図 13 が示すように視点利用率は高くなるが，対象は画像上で小さく撮影されるため画素利用率は低下する．同様のトレードオフは図 14 から読み取れる．このように従来の

固定カメラによる撮影方法では，視点利用率か画素利用率のいずれかを犠牲にした撮影しか行えない．一方，提案手法は図 13，図 14 が示すように固定カメラによる画素利用率優先の撮影方式と同等の画素利用率を保ったまま，視点利用率を高めることができる．視点利用率は従来方法・提案手法とも対象の位置により大きく変動する性質を持つが，利用率の最悪値で比較を行うと，二つのシナリオともに従来の分解能優先方式よりも視点利用率が向上していることが判る．このように視点利用率，画素利用率の観点からも提案手法が有効であることが定量的に示された．

5. まとめと今後の課題

従来の三次元ビデオ撮影方式はカメラの視野を固定して撮影を行うため，撮影可能範囲と空間分解能の間にトレードオフの関係が生じ，広範囲を移動する撮影対象の高分解能な三次元ビデオを撮影することができなかった．この問題に対して，我々は能動カメラを用いた広域移動対象の追跡撮影方式が満たすべき条件を

整理し、追跡撮影方式における問題の定式化を行った。そして、その現実的な解法の一つとしてセル方式追跡撮影方式を提案した。さらに実証実験を行い、本方式は移動経路が既知の移動対象を、高分解能な三次元ビデオが生成できるように追跡撮影できることを示した。

ただし提案手法は準最適な視野制御を行う現実的な手法であり、最適な視野制御を行うものではない。また移動経路を既知とする強い制約もある。今後の課題としては、本論文で示した問題の定式化の下で、移動経路が未知の移動対象を最適な視野制御で撮影できる理想的な解法アルゴリズムを検討することが挙げられる。

謝辞 本研究の一部は、文部科学省「知的資産の電子的な保存・活用を支援するソフトウェア技術基盤の構築」事業「大型有形・無形文化財の高精度デジタル化ソフトウェアの開発」プロジェクトの補助を受けて行った。

文 献

- [1] T. Kanade, P. Rander, and P. Narayanan, "Virtualized reality: Constructing virtual worlds from real scenes," *IEEE Multimedia*, vol.4, no.1, pp.33-47, 1997.
- [2] T. Matsuyama, X. Wu, T. Takai, and S. Nobuhara, "Real-time 3D shape reconstruction, dynamic 3D mesh deformation, and high fidelity visualization for 3D video," *Computer Vision and Image Understanding*, vol.96, no.3, pp.393-434, 2004.
- [3] C. Esteban, and F. Schmitt, "Silhouette and stereo fusion for 3D object modeling," *Computer Vision and Image Understanding*, vol.96, no.3, pp.367-392, 2004.
- [4] J. Starck, and A. Hilton, "Virtual view synthesis of people from multiple view video sequences," *Graphical Models*, pp.600-620, 2005.
- [5] Y. Furukawa, and J. Ponce, "Carved visual hulls for image-based modeling," *European Conference on Computer Vision (ECCV)*, pp.564-577, 2006.
- [6] A. Laurentini, "The visual hull concept for silhouette based image understanding," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.16, no.2, pp.150-162, 1994.
- [7] 和田俊和, 浮田宗伯, 松山隆司, "視点固定型パン・チルト・ズームカメラとその応用," *信学論 (D-II)*, vol.J81-D-II, no.6, pp.1182-1193, June 1998.
- [8] J. Kondou, X. Wu, and T. Matsuyama, "Calibration of partially-fixed viewpoint active camera," *情処学研報 (コンピュータビジョンとイメージメディア)*, vol.2003-CVIM-137 No.36, no.137-19, pp.149-156, 2003.
- [9] M. Isard, and A. Blake, "Condensation - conditional density propagation for visual tracking," *International Journal of Computer Vision*, vol.29, no.1,

pp.5-28, 1998.

(平成 xx 年 xx 月 xx 日受付)



吉本 廣雅 (正員)

2000年九大工学部卒。2002年九大大学院システム情報科学府修士課程修了。2005年同博士後期課程単位取得退学。2005年より京大大学院情報学研究所産学官連携研究員。コンピュータビジョン、実時間並列画像処理に関する研究に従事。



山口 辰久

2005年京大・工・情報卒。2007年京大大学院情報学研究所修士課程修了。現在、同研究科博士後期課程在学中。



牧 淳人 (正員)

1991年京大・工・電気卒。1993年東大大学院修士課程了。1996年スウェーデン王立工科大学(KTH)大学院博士課程了。Ph.D. 同年(株)東芝入社。関西研究所、研究開発センターに勤務。2003年京大芸術情報メディアセンターを経て翌年より同大学院情報学研究所に勤務(准教授)。2008年 Toshiba Research Europe Ltd 入社。現在、同ケンブリッジ研究所に勤務。コンピュータビジョン、時系列画像認識の研究に従事。1995年 SICE 学術奨励賞、2000年 電気通信普及財団テレコムシステム技術賞、2005年 IAPR MVA Most influential paper over the decade award その他受賞。IEEE、計測自動制御学会各会員。



松山 隆司 (正員:フェロー)

1976年京大大学院修士課程修了。京大助手、東北大助教授、岡山大教授を経て1995年より京大大学院電子通信工学専攻教授。現在同大学院情報学研究所知能情報学専攻教授。2002年学術情報メディアセンター長、京都大学評議員、2004年情報環境機構長。工博。画像理解、分散協調視覚、3次元ビデオの研究に従事。最近「人間と共生する情報システム」の実現に興味を持っている。1980年情報処理学会創立20周年記念論文賞、1990年人工知能学会論文賞、1993年情報処理学会論文賞、1994年電子情報通信学会論文賞、1995年第5回国際コンピュータビジョン会議 Marr Prize、1999年電子情報通信学会論文賞、2000年画像センシングシンポジウム優秀論文賞。2004年、2005年

FIT 優秀論文賞 . 国際パターン認識連合 , 情報処理学会 , 電子
情報通信学会フェロー . 日本学術会議連携会員

Abstract This paper proposes a 3D video capturing method with active cameras, which enables us to produce 3D video of a moving object in a wide spread area. Most of existing capturing methods use fixed cameras and have a strong restriction on allowable object motion; an object cannot move widely. To solve this problem, our method partitions a studio space into a set of subspaces named 'cells,' and conducts the camera calibration and control for object tracking based on the cells. We first formulate our method as an optimization problem and then propose an algorithm to solve it.

Key words 3D video, Tracking, Active camera, Multiple view video