

対象オブジェクトの周辺情報を考慮したズーム映像の生成

渡邊 健司[†] 新田 直子[†] 馬場口 登[†]

[†] 大阪大学大学院工学研究科 〒 565-0871 大阪府吹田市山田丘 2-1

E-mail: [†]{watanabe,naoko,babaguchi}@nanase.comm.eng.osaka-u.ac.jp

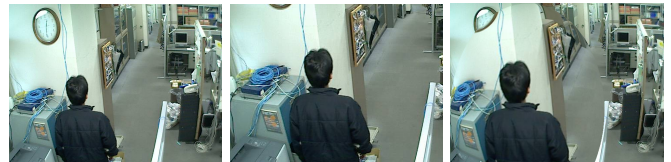
あらまし 本稿では、閲覧者の負担を軽減するための映像サーベイランス技術として、周辺情報を考慮したズーム映像の生成手法を提案する。提案手法では、映像中においてズーム表示の対象とするオブジェクトの領域を対象オブジェクト領域、それ以外の全ての領域を周辺領域と定義し、対象オブジェクト領域をズーム表示する際、重要度に応じて周辺領域も変形させ表示することにより、対象オブジェクトの詳細と周辺領域の内容、つまり周辺情報を同時に把握可能な映像を生成する。具体的には、直線エッジを考慮して入力フレームを表す三角形メッシュを導入した後、エッジ強度に基づき決定される画素の重要度に応じて三角形メッシュを変形させることにより、周辺領域の重要部分の変形を抑えた映像の生成を目指す。既存手法との比較実験において、提案手法がより周辺領域の変形を抑制可能であることを確認した。

キーワード 対象オブジェクト、周辺情報、ズーム映像

1. はじめに

近年、安心・安全な生活環境の実現のため、多くの場所で監視カメラによる映像サーベイランスが行われている。サーベイランス映像中には、一般に、閲覧者が注目するオブジェクトが含まれており、そのオブジェクトの詳細を把握するためにはズーム表示が必要である。

画像・映像中のオブジェクトに対するズーム表示手法は、大別すると、full-zoom view を用いる手法 [1], [2], [3], fisheye view を用いる手法 [4], [5], [6] に分けられる。前者は、図 1(b) に示すように、ズーム表示の対象オブジェクトを取り出して表示する手法であり、閲覧者は対象オブジェクトの詳細を把握可能である。しかし、可視範囲が狭まり、対象オブジェクトとその周辺領域とのつながりは失われるため、閲覧者が元の画像・映像の全体の概略を把握することは困難である。他方、後者は、図 1(c) に示すように、画面上で局所的に拡大率を変化させ、対象オブジェクトの周辺領域も変形させ表示する手法であり、閲覧者は対象オブジェクトとその周辺領域の内容を同時に把握可能である。刻々と状況が変化する映像サーベイランスにおいては、対象オブジェクトをズーム表示する際、周辺領域の内容も常に把握できることが望ましく、後者の手法がより有効であると考えられる。但し、fisheye view を用いる手法の多くは、対象オブジェクトからの距離のみに応じて拡大率を決定するため、周辺領域の内容、つまり周辺情報に対する考慮は十分であるとは言えない。周辺領域の重要な部分が変形するほど、周辺情報は失われると考えられる。そこで、本研究では、対象オブジェクトをズーム表示するのみでなく、周辺領域の重要な部分についても変形を抑えることにより、周辺情報を保持したズーム映像の生成を目指す。



(a) 入力フレーム (b) full-zoom view (c) fisheye view

図 1 ズーム表示の例

内容を考慮して情報を保持するという観点から、近年、画像・映像のリターゲットング手法 [7], [8], [9], [10], [11], [12], [13], [14] が数多く提案されている。リターゲットングとは、画像・映像のサイズが表示画面よりも大きい場合に、情報量をなるべく減らさずに効率よくサイズを変更する手法である。例えば、Avidan ら [8] は、画像情報に基づき定義されたシームと呼ばれる画像の継ぎ目を順に切り取ることにより、縮小された画像の違和感を減らす、シームカービングと呼ばれる手法を提案している。また、Wolf ら [10] は、別のアプローチとして、サイズ変更後の各画素の座標を連立一次方程式から求め、各画素を移動させる手法を提案している。ズーム映像の生成においても、ズーム表示の対象オブジェクトとその周辺領域を限られた画面内に同時に表示させることが必要になるという点で、このリターゲットングの概念を取り入れることが可能であると考えられる。

以上の背景をふまえて本稿では、対象オブジェクトをズーム表示するという目的のもとに、内容を考慮して情報を保持するというリターゲットングの概念を取り入れることにより、周辺情報を保持したズーム映像の生成を試みる。提案手法では、まず、サーベイランス映像中から対象オブジェクトを抽出する。次に、限られた画面内に、ズーム表示した対象オブジェクトと、その周辺情報

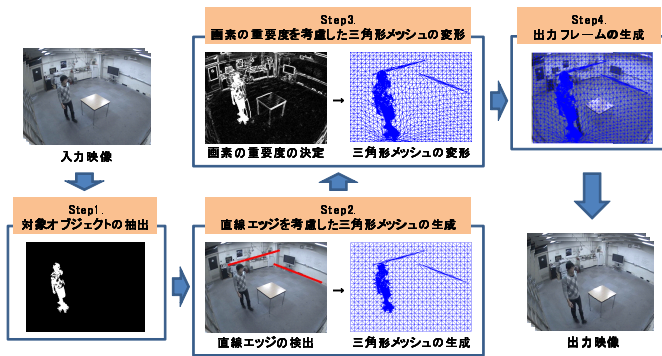


図 2 全体図

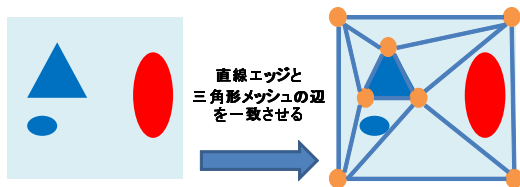


図 3 直線エッジを考慮した三角形メッシュの生成

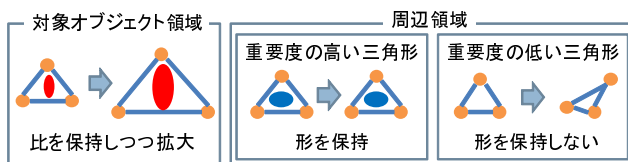


図 4 画素の重要度を考慮した三角形メッシュの変形

を同時に表示するために、周辺領域を内容に応じて変形する．具体的には、周辺領域のエッジ強度に着目し、エッジ強度の低い部分を変形させる一方、エッジ強度の高い部分については変形を抑制する．このように、対象オブジェクトからの距離のみに応じてではなく、内容を考慮した変形を行うことにより、周辺情報を保持したズーム映像の生成を試みる．

2. 手法概要

図 2 に提案手法の全体図を示す．本研究は、映像中の対象オブジェクトを、周辺情報を保持しながらズーム表示することを目的としているが、同一画面上に、ズーム表示された対象オブジェクトとその周辺情報を同時に表示するためには、周辺領域を変形する必要がある．ここで、周辺領域の中には、変形すると映像として見づらくなる部分と、変形しても見え方に影響が少ない部分があると考えられる．中でも特に、エッジ強度が高い部分を変形するほど、見づらい映像となることが推察される．そこで、提案手法では、周辺領域を変形する際に、エッジ強度が高い部分については変形を抑えるという方針をとる．具体的にはまず、周辺領域のうちエッジ強度が高い部分が直線となっている、直線エッジに着目し、図 3 に示すように、直線エッジとメッシュの辺を一致させるよう、入力フレームを表す三角形メッシュを生成する．三角形メッシュをどう変形させたとしても、三角形の 1

辺は直線となるため、メッシュの変形後も直線エッジが保持されることが期待される．次に、エッジ強度が高い部分のうち、直線エッジ以外の部分についても形を保持した変形を実現するために、エッジ強度に基づいて周辺領域の重要度を画素ごとに決定した後、図 4 に示すように、含まれる画素の重要度に応じて三角形メッシュを変形する．重要な画素、すなわちエッジ強度の高い画素を多く含む三角形ほど重要度を高く設定し、重要度が高い三角形ほど形を保持するよう、メッシュを変形させることにより、直線エッジ以外のエッジ強度が高い部分についても変形が抑えられることが期待される．

上記のように、周辺領域の中でも変形すると映像として見づらくなると考えられる部分であるエッジ強度が高い部分の変形を抑えることにより、周辺情報を保持したズーム映像を生成する．ズーム映像の生成は、以下の 4 ステップで構成される．

Step1. 対象オブジェクトの抽出

入力フレームから、ズーム表示の対象とするオブジェクトを抽出する．抽出されるオブジェクトの領域を対象オブジェクト領域、それ以外の領域を周辺領域と定義する．

Step2. 直線エッジを考慮した三角形メッシュの生成

周辺領域から、直線エッジを検出する．直線エッジと三角形の 1 辺を合わせるように、入力フレームを表す三角形メッシュを生成する．

Step3. 画素の重要度を考慮した三角形メッシュの変形
入力フレームの各画素に対して、エッジ強度に基づき画素ごとに重要度を決定する．決定した重要度をもとに、重要な部分の変形を抑えつつ、三角形メッシュを変形する．三角形メッシュの変形は、変形後三角形メッシュが満たすべき条件を表した連立一次方程式を解くことにより行う．

Step4. 出力フレームの生成

変形前後の三角形メッシュの対応する三角形ごとにアフィン変換を施すことにより、入力フレームを領域ごとに変形させ、出力フレームを得る．

3. 周辺情報を考慮したズーム映像の生成

本章では、前章で述べた各ステップのうち、提案手法の要となる Step1～Step3 について詳細を述べる．

3.1 対象オブジェクトの抽出

サーベイランス映像において、人物などの動物体は、閲覧者が特に注目するオブジェクトであると考えられる．そこで提案手法では、映像を撮影するカメラは固定されていると仮定し、入力映像中の動きのあるオブジェクトの領域を対象オブジェクト領域として抽出する．具体的には、前景領域の抽出、ノイズの除去からなる．

- 前景領域の抽出

まず始めに、入力映像に対してガウス混合分布による背景モデル構築法 [15] [16] を用いて背景を推定する．次

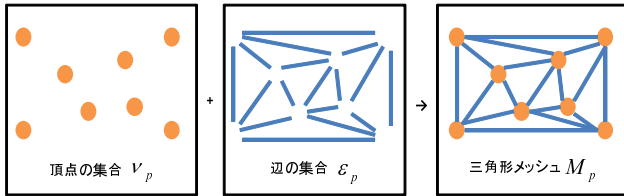


図5 三角形メッシュ

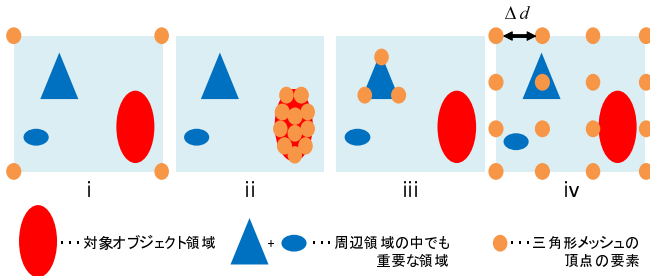


図6 V_p の要素

に、推定した背景と入力フレームとの差を前景領域として抽出する。

- ノイズの除去

抽出された前景領域には、背景の推定誤差によるノイズが含まれている可能性があるため、2 値化した後、独立な領域のうち面積が一定値 τ_m より大きい前景領域のみを対象オブジェクト領域として抽出する。ここで、 τ_m は対象オブジェクト領域の面積の下限を表す閾値である。

3.2 直線エッジを考慮した三角形メッシュの生成

3.2.1 直線エッジの検出

提案手法では、周辺領域のオブジェクトの直線エッジに、三角形メッシュの辺を合わせることで、出力フレームにおいてもオブジェクトの直線エッジの形を保持する。直線エッジは、フレームに対して Canny フィルタによりエッジを検出した後、確率的ハフ変換 [17] を用いることにより検出される。確率的ハフ変換により得られた線分のうち、長さが一定値 τ_l より長い線分のみを対象オブジェクトの直線エッジとして検出する。なお、提案手法では、映像中の背景オブジェクトは移動しないと仮定し、映像の始めの 1 フレームを用いて直線エッジを検出する。

3.2.2 三角形メッシュの生成

入力フレームを表すための三角形メッシュ M_p を生成する。図 5 に示すように、 M_p は、頂点の集合 V_p と、辺の集合 E_p からなる。従って、 M_p を生成するためにはまず、 M_p を構成する V_p を決定する必要がある。 V_p の要素として、入力フレームの全ての画素を含める方法も考えられるが、 V_p の要素数が増加するほど、三角形メッシュの変形に要する処理時間は増加する。そこで、提案手法では、入力フレームの代表的な画素のみを V_p の要素とすることにより処理時間の増加を防ぐ。図 6 に示すように、 V_p の要素は、以下の i ~ iv の画素の集合とする。

i. 入力フレームの四隅の画素

入力フレーム全体を覆う M_p を生成するために、入力フレームの四隅の画素を V_p に含める。

ii. 対象オブジェクト領域の全ての画素

対象オブジェクトをより正確にズーム表示するため、対象オブジェクト領域の全ての画素を V_p に含める。

iii. 周辺領域の画素のうち、直線エッジの端点の画素

周辺領域の中でも、特に画素値の変化が大きい部分が直線的につながっている部分である直線エッジが変形するほど見づらい映像となることが推察される。そこで、直線エッジに着目し、直線エッジの端点の画素を V_p に含める。入力フレームの直線エッジと、三角形メッシュの辺が一致するように、三角形メッシュを生成することにより、入力フレーム中の直線エッジは、出力フレームでも形が保持されることが期待される。

iv. 周辺領域から、 Δd の間隔でサンプリングした画素
周辺領域の画素のうち、 Δd の間隔でサンプリングした画素を V_p に含める。 Δd の値を変化させることにより、周辺領域をどの程度の細かさで見るとかを調整可能である。

上述の通り、 V_p を決定した後、 V_p に対して、ドロネー三角形分割を行う。この際、入力フレームの直線エッジと、三角形メッシュの辺が一致するように三角形分割に制約を加える [18]。ドロネー三角形分割により得られた辺の集合を E_p とする。

以上の操作により、入力フレームを表す三角形メッシュ M_p が生成される。

3.3 画素の重要度を考慮した三角形メッシュの変形

3.3.1 画素の重要度の決定

提案手法では、周辺情報を保持したズーム映像を生成するために、三角形メッシュの各三角形内の画素の重要度に応じて、三角形メッシュを変形する。そのため、まず各画素に対して重要度を決定する。画素 (i, j) の重要度 $S_{i,j}$ の決定は、対象オブジェクト領域、周辺領域の画素についてそれぞれ以下のように行う。

- 対象オブジェクト領域の画素

ズーム映像において、ズーム表示の対象である対象オブジェクト領域は、閲覧者にとって最も重要な領域であると考えられる。そこで、対象オブジェクト領域 Z の画素の重要度を、

$$S_{i,j} = 1 \quad ((i, j) \in Z) \quad (1)$$

とする。

- 周辺領域の画素

周辺領域の中には、変形すると映像として見づらくなる部分、つまり重要度が高い部分と、変形しても見え方に影響が少ない部分、つまり重要度が低い部分があると考えられる。中でも特に、画素値の変化が大きい部分であるエッジが変形するほど見づらい映像となることが推察される。そこで、提案手法では、エッジに着目し、画素値の変化量、つまりエッジ強度 $E_{i,j}$ にもとづいて周辺領

域の画素の重要度を決定する．なお，エッジ強度の算出には，Sobel フィルタを使用する．

$$E_{i,j} = \left| \sum_{k=0}^2 \sum_{l=0}^2 I_{i,j} V_{k,l} \right| + \left| \sum_{k=0}^2 \sum_{l=0}^2 I_{i,j} T_{k,l} \right|$$

但し，

$$V = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}, T = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix}$$

である．周辺領域の画素の重要度は，

$$S_{i,j} = \frac{E_{i,j}}{\max(E_{i,j})} \quad ((i,j) \notin Z), \quad (2)$$

と定義する．但し， $\max(E_{i,j})$ を $E_{i,j}$ の最大値と定義する．

これらをまとめて，画素の重要度 $S_{i,j}$ を

$$S_{i,j} = \begin{cases} 1 & ((i,j) \in Z) \\ \frac{E_{i,j}}{\max(E_{i,j})} & ((i,j) \notin Z) \end{cases} \quad (3)$$

とする．

3.3.2 三角形メッシュの変形

前節で得られた M_p を変形し，変形後の三角形メッシュ M_q を得る．ここで， M_p, M_q の頂点の集合 $\mathcal{V}_p, \mathcal{V}_q$ の要素をそれぞれ $\{P_k = (x_k, y_k) | k = 1, \dots, n\}, \{Q_k = (u_k, v_k) | k = 1, \dots, n\}$ とする．但し， n は要素数である．なお， $Q_k = (u_k, v_k)$ は，メッシュ変形後の $P_k = (x_k, y_k)$ の座標を表す．よって，三角形 (P_k, P_l, P_m) と，三角形 (Q_k, Q_l, Q_m) は， M_p と M_q の対応する三角形となる．

以降， M_q を求めるために，その要素 $\{Q_k | k = 1, \dots, n\}$ が満たすべき条件を立式し，その解を求める．具体的には， Q_k の x 座標 u_k ， y 座標 v_k についての連立一次方程式をそれぞれ立て，その解を求める問題とする．

条件は，フレームの端の画素位置に関する条件，画素間の距離に関する条件の2つを考える．なお， u_k, v_k は独立に考えることができ，同様に計算できるため，まず u_k についてそれぞれの条件式を以下のように設定する．以降， \mathcal{V}_p に含まれる頂点のうち，最左列，最右列にある頂点の集合をそれぞれ $\mathcal{V}_{p,L}, \mathcal{V}_{p,R}$ とする．

- フレームの端の画素位置に関する条件

$\mathcal{V}_{p,L}, \mathcal{V}_{p,R}$ に含まれる画素はそれぞれ，出力フレームの最左列，最右列に位置することが望まれる．それぞれの画素に対して，条件式を

$$u_k = 1 \quad (P_k \in \mathcal{V}_{p,L}), \quad (4)$$

$$u_k = L \quad (P_k \in \mathcal{V}_{p,R}), \quad (5)$$

と設定する．これらの式は，入力フレームの境界が出力フレームの境界となるよう働く．これにより，可視範囲が狭まることを防ぐ．



(a) 入力フレーム例 (b) 画素の重要度

図7 入力映像

- 画素間の距離に関する条件

M_p の辺の集合 \mathcal{E}_p のある要素の両端点を P_k, P_l とし， P_k, P_l は x 軸方向に r だけ離れて位置しているとする． α をズーム倍率とした場合， M_q において，対応する点である Q_k, Q_l は，対象オブジェクト領域では x 軸方向に αr ，周辺領域では r だけ離れて位置することが望まれる．辺の集合の全ての要素の両端点に対して，条件式を，

$$W_{k,l}(u_k - u_l) = \begin{cases} W_{k,l} \times \alpha r & (P_k, P_l \in Z) \\ W_{k,l} \times r & (\text{otherwise}) \end{cases} \quad (6)$$

と設定する．但し， $W_{k,l}$ は，この条件式に対する重みである． $W_{k,l}$ には，辺 (P_k, P_l) を含む三角形のうち，最も重要度が高い三角形の重要度が設定される．三角形の重要度には，三角形内に含まれる画素の重要度 $S_{i,j}$ の総和を $0 \sim 1$ に正規化した値が設定される．すなわち，重要度が高い画素を多く含む三角形ほど，重要な三角形ということになる．これにより，対象オブジェクト領域を拡大するとともに，周辺領域を重要度に応じて変形し，重要度の高い部分については形を保持する．

変形後の三角形メッシュは，上記の条件式を全て合わせた連立一次方程式 $Au = b$ の最小二乗解を求めることにより得られる． v_k についても同様に求められる．

以上の計算により，変形後の三角形メッシュ M_q が生成される．出力フレームは， M_p と M_q の対応する三角形ごとにアフィン変換を施すことにより生成される．

4. 実験

4.1 実験概要

提案手法の有用性を確認するために映像生成実験を行った．実験では，パラメータを $\tau_m = 100, \tau_l = 120$ とし，固定カメラを使用して予め撮影された映像を入力映像として使用した．映像サイズは，全て $320(\text{幅}) \times 240(\text{高さ})$ である．また，実験を行ったコンピュータ環境は，OS: Microsoft Windows 7 Professional, CPU: Intel(R) Xeon(R) CPU E5520 @ 2.27GHz, メモリ: 3.00 GB である．連立一次方程式の最小二乗解を求めるために，数値解析ソフトウェア MATLAB R2010b を使用した．生成された映像を比較し，特性を調べた．また，既存手法との比較実験を行った．

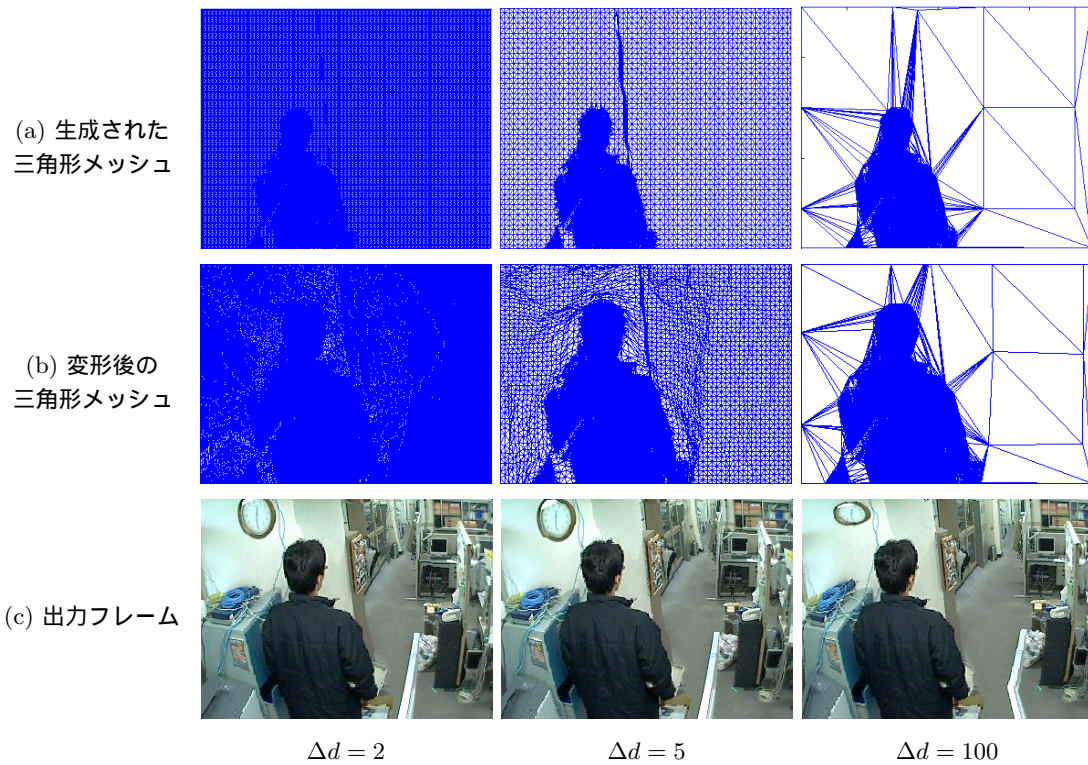


図 8 Δd の値と出力映像のフレーム例の比較

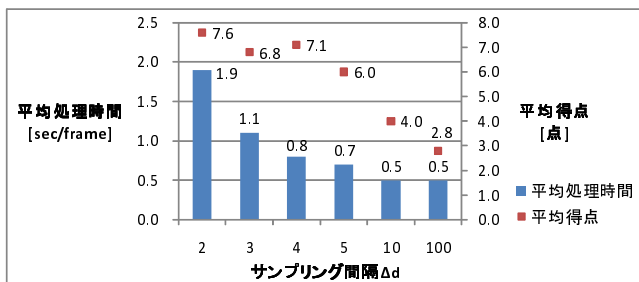


図 9 Δd の値と平均処理時間や平均得点との関係

4.2 三角形メッシュの密度の違いによる影響

提案手法では、 Δd の値を変化させることにより、三角形メッシュの密度を変化させることが可能である。三角形メッシュの密度の変化が、出力映像の質や処理時間に与える影響を検証するために、図 7 に示す入力映像を用いて、 Δd の値をそれぞれ $\Delta d = 2, 3, 4, 5, 10, 100$ の 6 種類に変化させた映像を生成し、平均処理時間を計測した。なお、全ての場合において、 $\alpha = \sqrt{2}$ とした。また、生成映像を被験者 6 名に提示し、映像の質に関して、最低点を 1、最高点を 10 として 10 段階評価させた。出力映像のフレーム例を図 8 に、映像生成に要した平均処理時間と、平均得点を図 9 にそれぞれ示す。

まず、 Δd の値と平均処理時間の関係について考察する。図 9 に示すように、 Δd の値を減少させるにつれて、平均処理時間は増加するという結果を得た。 Δd の値を減少させるほど、図 8(a) に示すように、三角形メッシュ M_p の頂点集合 \mathcal{V}_p の要素数は増加し、生成される三角

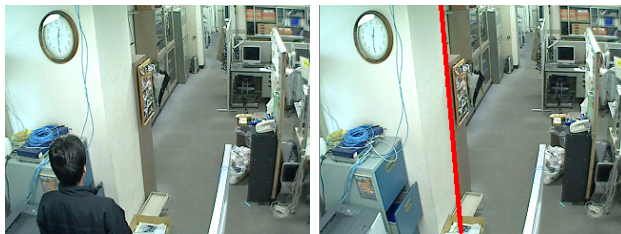
形メッシュの密度は高まる。そのため、図 8(b) に示す変形後の三角形メッシュを得る際に行う、最小二乗解の計算により時間がかかる。このような理由から、平均処理時間が増加したものと考えられる。

次に、 Δd の値と出力映像の質の関係について考察する。図 9 に示すように、 Δd の値を減少させるにつれて、平均得点は上昇するという結果を得た。図 8(c) より、対象オブジェクトである人物は、 Δd の値がいずれの場合もズーム表示されているが、周辺領域について比較すると、図 7(b) において重要度が高い部分、例えばフレーム左上に映る時計の形などが、 $\Delta d = 2$ や $\Delta d = 5$ の場合には保持されているのに対し、 $\Delta d = 100$ の場合には大きく変形していることが確認できる。一方、重要度が低い部分、例えば床面などは、 $\Delta d = 2$ や $\Delta d = 5$ の場合の方が変形が大きい。このように、 Δd の値を減少させるほど、図 8(a)、(b) に示すように、三角形メッシュの密度が高まるため、周辺領域について重要度に応じた細かい変形が可能となる。このような理由から、平均得点が上昇したものと推察される。

平均処理時間の観点からは、 Δd の値を増加させる方が望ましいが、 Δd の値を増加させるほど、平均得点は減少する。そこで、平均処理時間と、主観評価により得られた平均得点とのバランスを考え、以降の実験では、 Δd の値を $\Delta d = 4$ と設定する。

4.3 直線エッジの考慮の有効性

周辺領域のオブジェクトの直線エッジに合わせた三角形メッシュ M_p を生成することにより、出力フレームに



(a) 入力フレーム (b) 検出された直線エッジ



(c) 直線エッジの考慮なし (d) 直線エッジの考慮あり

図 10 直線エッジの考慮の有無による結果の例比較

どのような効果が現れるのかを検証するために、図 7 に示す入力映像を用いて、直線エッジを考慮する場合、考慮しない場合についてそれぞれ映像を生成した。なお、 $\alpha = \sqrt{2}$ とした。図 10 に出力映像のフレーム例を示す。図 10 の (a), (b), (c), (d) はそれぞれ、入力フレーム例、検出された直線エッジ（赤線）、直線エッジの考慮なしの場合の出力フレーム例、直線エッジの考慮ありの場合の出力フレーム例である。

図 10 より、直線エッジを考慮していない (c) においては、対象オブジェクト領域、この場合においては人物の近辺の柱に曲線的な変形が見受けられるのに対して、考慮している (d) においては、そのような柱の変形が抑えられ、違和感が少なくなっている様子が確認できる。この結果より、直線エッジを考慮することにより、考慮した直線エッジについては、出力フレームでも形が保持される傾向があることを確認した。一方、図 10(d) の人物前方、通路の奥に注目すると、図 10(c) よりも変形が大きい。この理由として、検出された柱の直線エッジを保持するために、他の領域に歪みが現れたことが挙げられる。

4.4 既存手法との比較

内容を考慮するという点で同様の考え方をもち、シームカービング [8] を比較対象とし、提案手法により生成された映像の主観評価を行った。主観評価では、対象オブジェクトとして室内を歩く人物が含まれている 2 種類の映像を入力映像として用いた。それぞれ、映像 1、映像 2 と呼ぶ。なお、シームカービングによるズーム映像の生成は、図 11 に示すように、以下の 2 ステップにより行った。

- Step1. 入力フレームを線形変換により α 倍に拡大
- Step2. 対象オブジェクト領域を保持しながら、シームカービングによりフレームサイズを $1/\alpha$ 倍に縮小

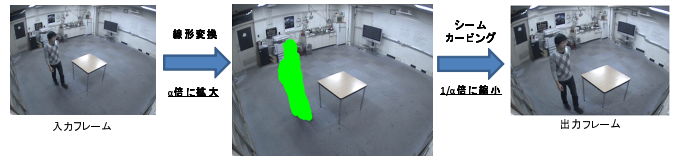


図 11 シームカービングによるズーム映像の生成

表 1 対象オブジェクトの詳細の把握 [点]

		被験者				
		A	B	C	D	平均
提案手法	映像 1	5	5	3	2	3.7
	映像 2	3	5	2	3	3.2
Scaling+SC	映像 1	5	5	4	4	4.5
	映像 2	4	5	4	4	4.2

表 2 周辺領域の変形の抑制 [点]

		被験者				
		A	B	C	D	平均
提案手法	映像 1	4	4	4	5	4.2
	映像 2	4	5	4	4	4.2
Scaling+SC	映像 1	1	3	2	2	2.0
	映像 2	3	3	2	4	3.0

$\alpha = \sqrt{2}$ とした。以降、この手法を Scaling+SC と呼ぶ。被験者 4 名に対し、入力映像、Scaling+SC による結果、提案手法による結果を並べて提示し、以下の 2 点について、最低点を 1、最高点を 5 とする 5 段階評価を行った。

- 対象オブジェクトの詳細の把握
- 周辺領域の変形の抑制

主観評価の結果を表 1、表 2 に示す。また、出力映像のフレーム例を図 12 に示す。

まず、対象オブジェクトの詳細の把握に関しては、提案手法による結果と比較し Scaling+SC による結果に対する平均得点が高かった。この点に関連して、被験者から、提案手法による結果において、対象オブジェクトである人物のアスペクト比が変化している場合があるというコメントがあった。図 13 にアスペクト比の変化の一例を示す。この理由として、提案手法では、三角形メッシュの変形の計算の際、 x 座標と y 座標を独立に計算していることが挙げられる。処理時間の短縮の観点からは、条件式は少ない方が望ましいが、対象オブジェクトの詳細把握の観点から、ズーム映像の質を向上させるためには、 x 座標と y 座標の関係について、アスペクト比が保たれるように条件式を追加する必要があると言える。

一方、周辺領域の変形の抑制については、提案手法により生成したズーム映像が、映像 1、映像 2 に対して共に平均 4.2 点の高い評価を得た。Scaling+SC によるズーム映像の生成では、入力フレームを一度線形変換により拡大するため、フレームに映るオブジェクトは一樣に拡大される。そこから、シームカービングにより重要な画素から順に切り取り縮小するため、対象オブジェク

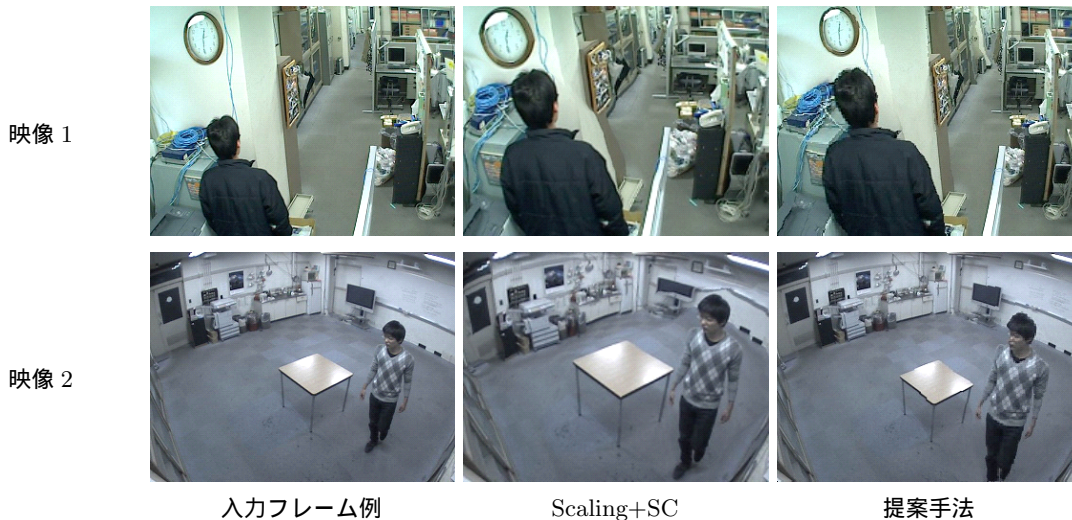
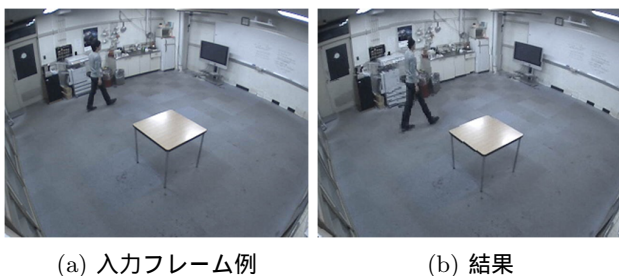


図 12 出力映像のフレーム例



(a) 入力フレーム例 (b) 結果
図 13 対象オブジェクトのアスペクト比の変化

ト領域のみでなく、周辺領域のうち重要な部分は結果的に全て拡大される。一方、提案手法では、対象オブジェクト領域と周辺領域の画素を分けて考えているため、対象オブジェクト領域のみを拡大することが可能である。そのため、周辺領域により多くのスペースを与えることが可能であり、その結果、変形を抑えられたものと考えられる。

4.5 映像生成例

主観評価に用いた映像以外の映像を入力映像として、Scaling+SC と提案手法をそれぞれ適用した場合の結果例を図 14 に示す。結果例の比較により、提案手法の汎用性を示す。 $\alpha = \sqrt{2}$ とした。

図 14(a)、図 14(b) の映像において、どちらも対象オブジェクトは室内を歩く 1 名の人物である。Scaling+SC、提案手法の結果ともに、人物がズーム表示されていることが確認できる。一方、周辺領域について、図 14(a) においては室内中央の机に、図 14(b) においては人物右側の柱に注目すると、提案手法と比較し、Scaling+SC の結果に、より変形が目立つ。

図 14(c) の映像において、対象オブジェクトは道路を走る 1 台の車であるが、同様に、周辺領域について、道路に注目すると、Scaling+SC の結果と比較し、提案手法の結果の方が、より道路の形が入力映像に近い形で保

持されていることが確認できる。

図 14(d) の映像においては、対象オブジェクトは室内を歩く 3 名の人物である。周辺領域について、室内中央の机や、室内奥のホワイトボードに注目すると、Scaling+SC の結果に、より変形が目立つ。対象オブジェクトが複数の場合でも、このように、提案手法は適用可能である。

5. おわりに

本稿では、周辺領域の重要度を考慮することにより、周辺情報を保持したズーム映像を生成する手法を提案した。提案手法では、まず、映像中から動きのあるオブジェクトをズーム表示の対象オブジェクトとして抽出した。次に、その周辺領域について、直線エッジを考慮し、入力フレームを表す三角形メッシュを生成した。最後に、エッジ強度に基づき決定された画素の重要度に応じて、三角形メッシュを変形させることにより、対象オブジェクトをズーム表示するとともに、周辺領域の重要部分については変形を抑えた映像を生成した。

実験では、提案手法を用いて実際に生成した映像を使用し、三角形メッシュの密度とズーム映像の質や処理時間との関係について考察した後、直線エッジの考慮の有効性について明らかにした。その結果、三角形メッシュの密度を高めることにより、ズーム映像の質が向上する一方、処理時間は増加する傾向を確認した。また、直線エッジを考慮することにより、周辺領域のオブジェクトの考慮した直線エッジが保持されたズーム映像を生成可能であることを確認した。さらに、主観評価により他手法との比較を行った。その結果、周辺領域の変形の抑制に関しては、提案手法が高い評価を得たが、ズーム映像において、対象オブジェクトのアスペクト比が変化するなどの問題点も明らかとなった。今後、アスペクト比の保持に関する条件を式として追加することを検討する必要がある。また、動画を対象とするため、時間的な連続性も考慮することが重要な課題となる。



(a) (b) (c) (d)

図 14 映像生成例．上段：入力フレーム，中段：Scaling+SC による結果，下段：提案手法による結果

文 献

- [1] B. B. Bederson and J. D. Hollan: “Pad++: A zooming graphical interface for exploring alternate interface physics”, Proc. 7th Annual ACM Symposium on User Interface Software and Technology, pp. 17–26 (1994).
- [2] H. El-Alfy, D. Jacobs and L. Davis: “Multi-scale video cropping”, Proc. 15th ACM International Conference on Multimedia, pp. 97–106 (2007).
- [3] P. Chiu, K. Fujii and Q. Liu: “Content based automatic zooming: Viewing documents on small displays”, Proc. 16th ACM International Conference on Multimedia, pp. 817–820 (2008).
- [4] G. W. Furnas: “Generalized fisheye views”, Proc. ACM SIGCHI Conference on Human Factors in Computing Systems, pp. 16–23 (1986).
- [5] Y. Hosokawa, N. Kimura and N. Takahashi: “An implementation method of a location-based active map transformation system”, Proc. 6th ACM International Conference on Mobile Data Management, pp. 13–21 (2005).
- [6] G. Wang, T.-T. Wong and P.-A. Heng: “Gpu-friendly warped display for scope-maintained video surveillance”, ACM/Springer Multimedia Systems Journal, **12**, 3, pp. 169–178 (2006).
- [7] F. Liu and M. Gleicher: “Automatic image retargeting with fisheye-view warping”, Proc. 18th Annual ACM Symposium on User Interface and Software Technology, pp. 153–162 (2005).
- [8] S. Avidan and A. Shamir: “Seam carving for content-aware image resizing”, ACM Transactions on Graphics, **26**, 3, p. 10 (2007).
- [9] M. Rubinstein, A. Shamir and S. Avidan: “Improved seam carving for video retargeting”, ACM Transactions on Graphics, **27**, 3, p. 16 (2008).
- [10] L. Wolf, M. Guttmann and D. Cohen-or: “Non-homogeneous content-driven video-retargeting”, Proc. 11th IEEE International Conference on Computer Vision, pp. 1–6 (2007).
- [11] Y.-F. Zhang, S.-M. Hu and R. R. Martin: “Shrinkability maps for content-aware video resizing”, Computer Graphics Forum, **27**, 7, pp. 1797–1804 (2008).
- [12] P. Krähenbühl, M. Lang, A. Hornung and M. Gross: “A system for retargeting of streaming video”, ACM Transactions on Graphics, **28**, 5, p. 126 (2009).
- [13] Y.-S. Wang, C.-L. Tai, O. Sorkine and T.-Y. Lee: “Optimized scale-and-stretch for image resizing”, ACM Transaction on Graphics, **27**, pp. 1–8 (2008).
- [14] T. Ren, Y. Liu and G. Wu: “Image retargeting based on global energy optimization”, Proc. IEEE International Conference on Multimedia and Expo, pp. 406–409 (2009).
- [15] C. Stauffer and W. E. L. Grimson: “Adaptive background mixture models for real-time tracking”, Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Vol. 2, pp. 246–252 (1999).
- [16] P. KaewTraKulPong and R. Bowden: “An improved adaptive background mixture model for real-time tracking with shadow detection”, Proc. 2nd European Workshop on Advanced Video Based Surveillance Systems, pp. 135–144 (2001).
- [17] N. Kiryati, Y. Eldar and A. Bruckstein: “A probabilistic hough transform”, Pattern Recognition, **24**, 4, pp. 303–316 (1991).
- [18] L. P. Chew: “Constrained delaunay triangulations”, Proc. ACM 3rd Annual Symposium on Computational Geometry, pp. 215–222 (1987).