

## 風景写真における建物領域分割の一検討

藤村 誠<sup>†</sup> 松濤 弘幸<sup>††</sup> 今村 弘樹<sup>†</sup> 黒田 英夫<sup>††</sup>

<sup>†</sup> 長崎大学工学部 〒852-8521 長崎市文教町 1-14

<sup>††</sup> 長崎大学大学院 〒852-8521 長崎市文教町 1-14

E-mail: †{makoto,imamura,kuroda}@cis.nagasaki-u.ac.jp

あらまし 風景の中央に道があり、それに沿って建物があるような画像から擬似的な3次元CG空間を生成するためには、その画像を背景、道、道沿いの左右の建物領域の4領域に分割する必要がある。4領域への分割手法については、いくつか報告されているが、建物領域の上部、すなわち屋根部分と背景の境界線の精度に問題があった。本稿では、建物領域上部の抽出精度を向上する手法を提案している。まず、Hough変換により風景画像中の消失点の検出し、4領域に大まかに分割する。並行して濃度などの特徴量を用いてK-means法によって、領域分割を行う。最後に、両領域分割結果に基づいて、4領域の抽出を行う。評価実験では良好な結果を得ている。

キーワード TIP, 透視投影, 3次元シーン, 消失点, 領域分割

## A Study on Building Region Extraction Method from Perspective Image

Makoto FUJIMURA<sup>†</sup>, Hiroyuki MATSUNAMI<sup>††</sup>, Hiroki IMAMURA<sup>†</sup>, and Hiedo KURODA<sup>††</sup>

<sup>†</sup> Faculty of Engineering, Nagasaki University Bunkyo-machi 1-14, Nagasaki city, Nagasaki, 852-8521 Japan

<sup>††</sup> Graduate School of Science and Technology, Nagasaki University Bunkyo-machi 1-14, Nagasaki city, Nagasaki, 852-8521 Japan

E-mail: †{makoto,imamura,kuroda}@cis.nagasaki-u.ac.jp

**Abstract** Tour Into the Picture is that user walk through the 3-D computer graphics world made from a perspective image. Making 3-D computer graphics world need high accuracy segmentation and extraction of regions from the perspective image. In this paper, we proposed an high accuracy extraction method of 4-region, which are road region, background, left building region, right building region, from a perspective image. Experiments demonstrated good results of our proposed method.

**Key words** Tour Into the Picture, Perspective, 3-D Scene, Vanishing point, segmentation

### 1. ま え が き

近年、仮想的な3次元空間を生成する際には、Image-Based Rendering (以後、IBRと称する) [1][2]手法が良く用いられている。IBRは、様々な視点から撮影した大量の画像を蓄積しておき、視線などの条件に応じて、現実の景観を模した仮想空間を精度高く構築できる。しかし、大量の景観画像を必要とするため、一般ユーザには取扱いが難しい。

一方、文献[3]では、Tour Into the Picture (以後、TIPと称する) という1枚の透視投影画像から3次元シーンを生成する手法が提案されている。この方法は、必要とする画像は少ないが、3次元空間生成に必要な情報、すなわち点、エッジおよび線などのユーザによる指示が必要という問題があった。文献[4][5]では、少数の画像から3次元シーンを生成しているが、消失点と消失線だけによる領域抽出であり、建物群、特に

屋根の部分は隣接する建物間で不揃いであることに対応していないという問題がある。

筆者らは、TIPのように、道が中央に位置している1枚の風景画像から背景、道、左右の建物群の領域を抽出し、その奥行き位置にしたがって配置することで擬似的な3次元シーンを生成する手法を提案している[6]。そのためには1枚の風景画像から、背景、道、左右の建物群の4領域を精度良く抽出する必要があるが、現在は手動で行っていた。

本稿では、単眼カメラによる風景画像を対象として、消失点検出に続いて背景、道、左右の建物群領域に分割し、抽出することで建物群の抽出精度を向上する手法を提案する。まず、画像中の消失点を求め、並行してK-means法による領域分割も行う。次に、消失点を中心とした角度ヒストグラムを作成し4領域に直線境界線による分割を行い、道路領域を抽出する。最後に、背景領域と建物群領域の屋根部分との境界を求めるため、

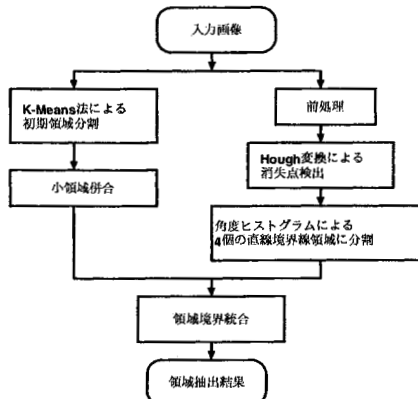


図1 処理の流れ  
Fig.1 Process flow

先の4領域の背景領域部分の直線境界線とK-means法による領域分割結果を統合し、4領域を抽出する。

## 2. 風景画像からの領域抽出

### 2.1 概要

1枚の風景画像からの領域抽出に要する処理の流れを図1に示す。まず、入力画像に対して前処理を行い、消失点および消失線を検出し、画像を4つの領域に分割する。また、これとは別に、K-means法による領域分割を行い、過分割に対する少領域併合を行う。最後に、両領域分割結果の結果を踏まえて、建物領域の上部と背景領域の境界である輪郭線および建物領域の下部と道領域の境界である輪郭線について決定し、4領域を抽出する。

### 2.2 風景写真の消失点検出

前処理では、メディアンフィルタによる平滑化、Sobelフィルタによるエッジ検出、判別分析法による二値化を行う。ただし、単純にこれらの方法を用いると、消失点に向かって伸びる直線上の点だけでなく、建物の柱、扉、窓の線、樹木の幹、電柱などの垂直方向の直線上の点も多数検出する。このような垂直成分の直線は、消失点の検出精度を低下させるため、あらかじめ除去しておく必要がある。すなわち、垂直方向の直線上に存在する点の場合、水平方向にエッジを生じないため、下式を満たすエッジ点は前処理の段階で除去する。

$$\theta(x, y) = \arctan \left( \frac{\partial f(x, y)}{\partial y} / \frac{\partial f(x, y)}{\partial x} \right) \leq \theta_{th} \quad (1)$$

消失点は、Hough変換を用いて画像中の直線が集中する点として検出する。道が中央に位置する風景画像を想定しているため、画像の下半分領域では道路と建物のエッジなど消失点に向かう直線が多いが、上半分領域では道に並ぶ建物の高さや屋根形状は揃っていないため画像の消失点に向かう直線が比較的少ない。このため、画像の下半分領域に対してHough変換を行う。Hough変換の性質から、2次元画像上で直線群が共通の交

点 $(x_0, y_0)$ で交差するならば、それらの直線をパラメータ変換した空間上では、下式で表わされる曲線上に存在するはずである。そこで、最小2乗法によって求めた近似方程式により、消失点座標を得る。

$$\rho = x_0 \cdot \cos \theta + y_0 \cdot \sin \theta \quad (2)$$

### 2.3 領域分割

領域分割にはK-means法を用いており、特徴量には $5 \times 5$ のブロックにおける濃度平均と濃度分散を用いた。

領域分割では1つの物体に対して1つの領域が形成されることが望ましいが、一般には領域の過分割という問題が生じる。この過分割の原因には主として次の2つが挙げられる。第1は、陰影、ハイライト、相互反射やグラデーションなど同じ物体の表面であっても、その位置によって照明条件が変化してしまう光学的な要因である。一様な反射率を持つ物体表面上では一様な色や輝度を持つという仮定が成立しなくなり、1つの物体表面上に複数の領域が形成されることになる。第2は、物理的要因、つまり1つの物体表面が実際に複数の異なる反射率を持つ物体で構成されている場合である。このような場合、1つの物体であっても多数の領域に分割されることになる。

この過分割については、領域併合によって対処する手法がいくつ提案されている。文献[7][8]では、K-means法の初期設定を変化させるとクラスタリングの結果が変化することを利用して、生成された境界線の位置が変化するかどうかを判定基準として領域併合を行う手法が提案されている。これらの手法は、領域の境界線がエッジに一致することを狙った、光学的要因に対処するものである。文献[9]は、輪郭形状、コントラスト、色分布に基づく同一領域度を用いて領域併合を行うという手法であり、光学的要因、物理的要因双方に対応した方法である。文献[10]は、領域の配置情報、隣接情報、面積比情報の3つの併合基準を用い領域併合を行っており、物理的要因に対処した方法である。

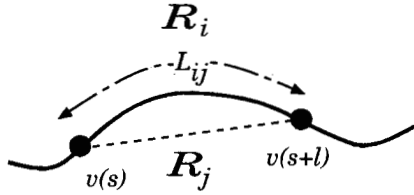
本研究では、対象となる風景画像の特徴を考慮し、包含の特徴量、境界線の滑らかさの特徴量、面積の特徴量の3個の特徴量を基に領域併合する手法を用いることとする。それぞれの特徴量について以下に具体的に説明する。

#### a) 包含の特徴量

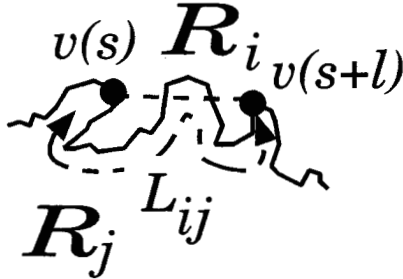
ある物体が別の物体に包含されている場合、その部分領域はある物体の構成部品に対応していることが多い。例えば、風景画像では「建物の領域には窓や扉のような領域が包含されている」などの傾向が見られる。

ここでは、包含の特徴量を注目領域の周囲長と注目領域と隣接している領域との境界線の長さとする。注目領域の周囲長を $L_i$ 、注目領域と隣接領域との境界線の長さを $W_{ij}$ とすると次式のように表すことができる。 $V_{inclusion}(i, j)$ がある閾値 $TH_{inclusion}$ より大きい場合は、領域 $i$ と $j$ は包含関係にあると見なし併合する。

$$V_{inclusion}(i, j) = \frac{W_{ij}}{L_i} \quad (3)$$



(a) 滑らかな場合



(b) 複雑な場合

図 2 境界線の共有部分

Fig. 2 Sharing boundary line between two regions

b) 境界線の滑らかさの特徴量

注目する境界線が、光学的要因によって生じる同一物体領域上の擬似輪郭であるか、あるいは異なる物体間の真の輪郭であるかは、その境界線の滑らかさによって判定できる。すなわち、境界線が滑らかな場合は真の境界線である可能性が高く、境界線が不連続で滑らかでない場合は疑似輪郭の可能性が高い。したがって、共有されている境界線が滑らかでない場合、その境界線は擬似輪郭であると見なし、これらの領域を併合することで物体の領域が得られると考えられる。

図 2 に示すように、境界線の滑らかさは、 $v_i(s)$  を領域  $R_i$  の境界線上の点、 $L_{ij}$  を  $R_i$  と  $R_j$  が共有する境界線の長さ、 $l$  を定数とすると次式で得られる。ここで、 $V_{smoothness}(i, j)$  がある閾値  $TH_{smoothness}$  より大きい場合は、領域  $i, j$  の境界線は疑似輪郭であると見なし併合する。

$$V_{smoothness}(i, j) = \frac{1}{L_{ij}} \int_{\Omega} \frac{\|v_i(s) - v_i(s+l)\|}{l} ds \quad (4)$$

c) 面積の特徴量

分割した領域の面積が数画素程度で小さい場合、隣接する大きな物体の一部分に対応する小さな領域である可能性が高い。この小領域が隣接する大きな物体に含まれる場合、もっとも特徴の近い隣接領域に併合することが妥当である。

d) 併合の順序

小領域の併合にあたっては、特徴量の性質より併合に用いる順序を検討する必要がある。滑らかさの特徴量は境界線情報に基づき算出するため、小さな領域が多数残っている場合には短

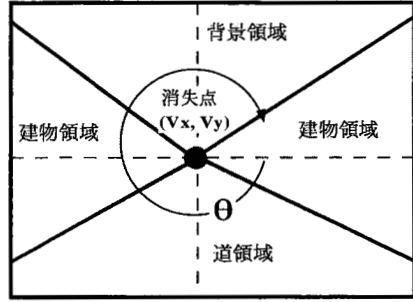


図 3 各領域の角度範囲

Fig. 3 Angle range of four regions

い境界線に対して計算してもあまり信頼できる結果は得られない。

したがって、併合の順序は次のようになる。まず、面積の特徴量を用いて微小領域を併合する。次に、包含の特徴量を用い包含関係にある領域を併合し、最後に滑らかさの特徴量を用いて併合処理を行う。

2.4 領域抽出

風景画像の構造は消失点を中心として 4 領域に分かれている。消失点座標  $(V_x, V_y)$  から角度  $\theta$  方向に引いた直線上の特徴点数を集計した角度ヒストグラム  $H(\theta)$  に基づき、それぞれの領域の境界の基準となる直線を求める。

一般に画像中の建物群領域は道領域や背景と比べて複雑であり、エッジ点が多数存在する。そのため、角度ヒストグラムを生成すると建物群領域、道、背景それぞれの境界線に対応する角度において明確なエッジが存在する。このエッジ検出のため、次式により角度ヒストグラムの微分値を求め、直線境界線の検出を行う。

$$H_{diff}[\theta] = H[\theta] - H[\theta - 1] \quad (5)$$

次いで、角度ヒストグラムからの境界線検出の手順を示す。各領域の境界線は、図 3 のようにそれぞれ、 $0 \sim \frac{\pi}{2}$ 、 $\frac{\pi}{2} \sim \pi$ 、 $\pi \sim \frac{3\pi}{2}$ 、 $\frac{3\pi}{2} \sim 2\pi$  の範囲にそれぞれ存在している。境界線の検出手順は次の通りである。

- step 1  $H_{diff}[\theta]$  から最大値を探し、max とする。
- step 2  $H_{diff}[\theta] \leq K \times \max(K: \text{定数})$  の条件を満たす  $\theta$  の中で、 $0 \sim \frac{\pi}{2}$ 、 $\pi \sim \frac{3\pi}{2}$  の場合は最も値が大きいものを、 $\frac{\pi}{2} \sim \pi$ 、 $\frac{3\pi}{2} \sim 2\pi$  の場合は最も値が小さいものを  $\theta_{max}$  とする。
- step 3 それぞれの領域において、

$$y = \tan \theta_{max}(x - V_x) + V_y \quad (6)$$

により直線に戻した式を境界線とする。

領域の抽出は、角度ヒストグラムにより算出した直線境界線および K-means 法による領域分割結果に基づき行う。道領域の境界は、直線的になっていることが多いが、道路領域と建物領域の間の濃度差は比較的小さい場合が多く、領域分割の結果繋がってしまうことが多い。そこで、道領域については角度ヒス



(a) テスト画像 1 (Test image No.1)



(b) テスト画像 2 (Test image No.2)

図 4 テスト画像

Fig. 4 Test image

トグラムによって得た直線境界線をそのまま用いることとする。

一方、建物の屋根については複雑な形状であることが多く、背景領域との境界線は複雑な形をしていることが多い。しかし、背景領域と建物領域の間は濃度差が大きいため明確なエッジが存在しており、領域分割により 2 つの領域が繋がってしまうことはほとんどない。そこで建物領域については、背景領域と建物領域双方の含有数を算出し、含有数の多い領域にその部分領域を併合する。建物領域は、背景領域と道領域を除いた領域の内、消失点より左側にあるものを左建物領域とし、右側にあるものを右建物領域とする。

### 3. 評価実験

実際に風景画像に提案手法を実行して評価実験を行った。

#### 3.1 実験条件

図 4 に、今回使用したテスト画像を示す。(a) は長崎市の出島の再現施設の風景画像であり、(b) は長崎大学付属図書館の古写真コレクション [11] 中の風景画像である。

これらのテスト画像に対して、消失点検出、K-Means 法による領域分割および領域併合などを行い、背景領域、道領域、左右の建物領域を抽出する。

##### 3.1.1 結果

図 5, 6 に、テスト画像 1 および 2 についての消失点検出結果、4 領域分割の直線境界線検出結果、K-means 法による領域分割および領域併合結果を示す。図の (a) および (b) については、消失点と消失点を中心とした 4 領域の直線境界線検出結果

を示している。(b) より、道領域の分割が良好に行われており、建物の屋根部分の分割についても、ほぼ良好な結果であることがわかる。図の (c) および (d) は、K-means 法による領域分割結果および領域併合の結果を示している。生じた過分割が領域併合によって統合され、特に建物の屋根領域の付近では領域分割の精度が向上していることがわかる。

図 7, 8 に、テスト画像 1 および 2 についての 4 領域分割結果を示す。両図より、ほぼ良好な結果が得られたことがわかる。特に、図の (c) および (d) より、背景と建物領域の屋根部分との境界の抽出精度が良いことがわかる。

## 4. まとめ

1 枚の風景画像から 3 次元シーンを生成するために、風景画像から背景、道領域、左右の建物群領域への 4 領域抽出を抽出する必要があるが、建物の屋根部分の抽出精度が低いという問題があった。本稿では、その問題解決のため、背景と建物の屋根部分との境界の分割を精度良く行う手法として、消失点を中心とした角度ヒストグラムによる直線境界線による領域分割と K-means 法による領域分割を統合する手法を提案した。評価実験の結果、道沿いの建物群の屋根が不揃いであっても、精度良く建物領域を抽出できることを確認した。

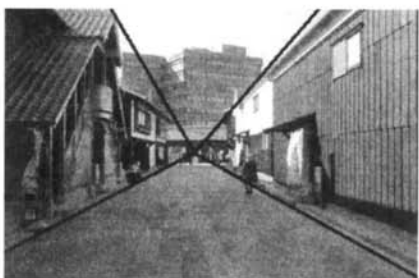
今後の課題は、さらに多くのテスト画像に対する検証、抽出した建物領域から個々の建物をさらに抽出する方法の検討などである。

## 文 献

- [1] Leonard Macmillan and Gray Bishop: "Plenoptic modeling: an image-based rendering system," Proc.SIGGRAPH '97, Los Angeles, California, pp.39-46(Aug. 1997)
- [2] Shenchang Eric Chen: "QuickTime VR: an image-based approach to virtual environment navigation," Proc.SIGGRAPH '95, Los Angeles, California, pp.29-38(1995)
- [3] Youiti Horry, Ken-ich Anjyo and Kiyoshi Arai: "Tour Into the Picture," Proc.SIGGRAPH '97, Los Angeles, California, pp.225-232(Aug. 1997)
- [4] 杉山 慎, 日黒 光彦, 金子 正秀: "少数の透視投影画像に基づいて生成された仮想空間における 3 次元移動表現", 信学技報, MVE 2001-148, pp.19-24(2002.3)
- [5] 杉山 慎, 日黒 光彦, 金子 正秀: "少数の実写画像に基づく擬似 3 次元シーンの自動生成", 映像情報メディア学会誌, vol.56, No.4, pp.643-646(2002.4)
- [6] Makoto Fujimura, Yusuke Kajiwara, Hideo Kuroda and Hiroki Imamura: "3D Computer Graphics World by Stage Setting Model," ACM SIGGRAPH '2005, (Aug-July. 2005)
- [7] 中谷 広正, 大崎 賢治, 阿部 圭一: "複数の領域分割に基づく対象物境界線検出", 電子情報通信学会論文誌, J76-D-II, 4, pp.914-916(1993)
- [8] 境田 慎一, 鹿嶋 善明, 田中 豊, 湯山 一郎: "K 平均アルゴリズムの初期値依存性を利用した統合処理による画像の領域分割", 電子情報通信学会論文誌, J81-D-II, 2, pp.311-322(1998)
- [9] 志治 亜矢子, 浜田 望: "Watershed アルゴリズムと輪郭情報を用いたカラー画像の領域分割", 電子情報通信学会論文誌, J83-D-II, 2, pp.593-600(2000)
- [10] 玉木 徹, 山村 毅, 大西 界: "対象物体に依存しない領域間の情報に基づいた領域併合による物体抽出手法", 映像情報メディア学会誌, vol.55, No.4, pp.571-582(2001)
- [11] 長崎大学付属図書館電子コレクション: <http://oldphoto.lib.nagasaki-u.ac.jp/univ/j/>



(a) 消失点検出結果



(b) 境界線検出結果



(c) 初期領域分割結果



(d) 領域併合結果

図 5 分割結果 (テスト画像 1)

Fig.5 Segmentation Results(Test image No.1)



(a) 消失点検出結果



(b) 境界線検出結果



(c) 初期領域分割結果



(d) 領域併合結果

図 6 分割結果 (テスト画像 2)

Fig.6 Segmentation Results(Test image No.2)

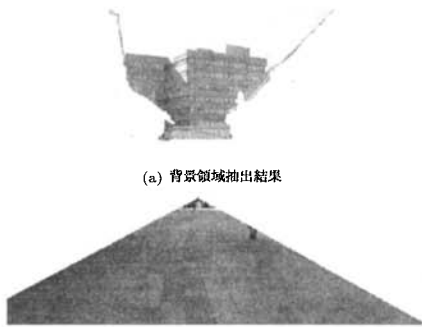


図 7 抽出結果 (テスト画像 1)  
Fig. 7 Extraction Results (Test image No.1)

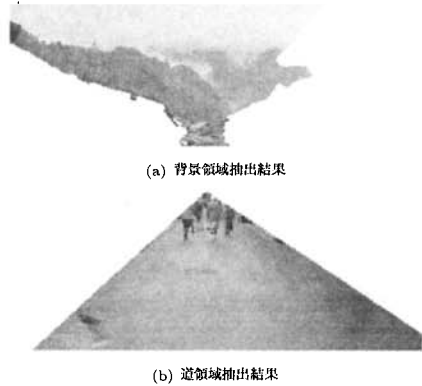


図 8 抽出結果 (テスト画像 2)  
Fig. 8 Extraction Results (Test image No.2)