

ストライプ状背景と動的輪郭モデルを用いた領域抽出

Region Extraction with Chromakey Using Stripe Backgrounds and Active Contour Models

縣 弘樹 山下 淳 金子 透

Hiroki Agata, Atsushi Yamashita, Toru Kaneko

静岡大学工学部機械工学科

Department of Mechanical Engineering, Shizuoka University

E-mail: yamashita@ieee.org

Abstract

本論文では、ストライプ状背景と動的輪郭モデルを用いたクロマキーにより領域抽出を行うことを目的とする。従来の単一色の背景を用いたクロマキー合成では、抽出対象に背景と同一色が含まれているとその部分は抽出することができない。そこで、2色からなるストライプ状背景の前で撮影を行い、色情報に加えてストライプの領域情報を用いることにより、背景と同一色を含む対象物を抽出可能な手法を提案する。本論文ではさらにその抽出結果に動的輪郭モデルを用いることにより抽出精度を向上させる。実験結果より、提案手法の有効性が確認された。

1 序論

アミューズメント関連分野などにおける低コスト化や仮想現実的な画像の実現には画像合成技術が必要である。

画像合成の一般的な手法にクロマキーと呼ばれる手法がある。この手法では青や緑など単一色の背景を用い、その前で合成する対象物を撮影する。次に得られた画像から色情報を用いて背景部分だけ除去し対象物のみを抽出する。そして抽出された対象物を他の背景と合成する。しかし一般的なクロマキーでは、どのような背景色を用いても、対象物の一部の色が背景と同じ色の場合は、その部分が背景と同様に除去されてしまうという問題点がある。

この問題に対して、背景を瞬間的に切り替えて2枚の画像を撮影することで対処する手法[1]、撮影対象の画像を取得すると同時に対象までの距離を計測して画像合成を行う方法[2,3]、画像撮影と同時に温度(体温)を計測して画像合成を行う手法[4]なども提案されている。しかし、これらの装置は高価かつ特殊なものであるため、誰もが気軽に開発・利用することができず、一般に普及するには至っていない。

一方、2色から成るストライプ状の背景を用いて簡

易的に画像合成を行う手法も提案されている[5,6]。この手法により対象物の一部が背景色と同じ色でも、その部分が除去されず、対象物を抽出できるようになる。ところがこの手法では、対象物の周りにノイズとして背景が残るという問題点がある。

よって本論文では2色からなるストライプ状背景を用いた画像合成手法[5,6]を改良し、正確に対象物を抽出することを目的とする。具体的には、ストライプ状背景を用いた対象物抽出画像に対して動的輪郭モデル[7]を適用することで、正確に対象物の輪郭を抽出し、抽出された輪郭の外側、つまり対象物周りの背景を除去する手法を提案する。

2 ストライプ状背景を用いた対象物抽出

画像合成のための対象物抽出手法の概要は以下の通りである。まず背景の2色を除去する。しかしストライプの境界は背景の2色が混ざり合った混合色(Fig. 1)であるため、背景色と認識されず除去されない(Fig. 2)。このストライプの境界を利用することによって、対象物部分で除去された部分を補間する。対象物補間の際には縦 h 画素、横1画素の格子を考える。ここで h は、Fig. 2に示すようにストライプの境界を含むストライプの幅の大きさである。この格子で画像をラスタ走査していき、格子内にストライプの境界が含まれるとき、格子内は背景であると判断できる。逆にストライプの境界が含まれないとき、格子内は対象物であると判断できる[5,6]。

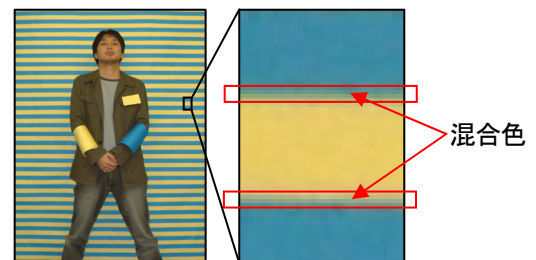


Fig. 1 撮影画像

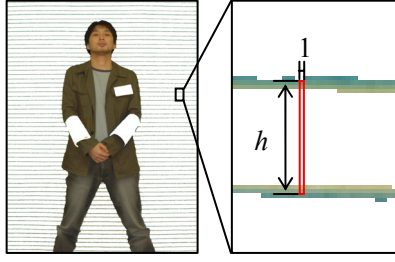


Fig. 2 背景色除去画像

対象物と判断された部分を補間すると Fig. 3 のようになる。しかし、この補間方法では対象物の周りの背景も補間対象となる。つまり、この対象物補間がストライプ状背景を用いた対象物抽出において、背景が対象物の周りにノイズとして残る最大の要因ともなっている。

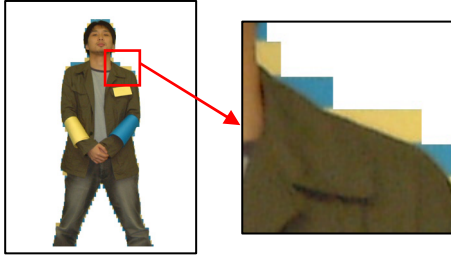


Fig. 3 ストライプ状背景を用いた対象物抽出画像

そこで本論文では、動的輪郭モデルを用いることにより、上記の処理で対象物を抽出した際に周囲に残るノイズを除去し、正確に対象物だけを抽出する手法を提案する。

3 動的輪郭モデル

3.1 原理

動的輪郭モデルでは、まず画像平面 (x, y) 上で対象物の近傍に、閉曲線 $\mathbf{v}_i = (x_i, y_i)$ ($i = 1, 2, \dots, n, \mathbf{v}_{n+1} = \mathbf{v}_1$) を与える。次にあらかじめ定義した輪郭モデルの特徴を表すエネルギーを最小化するようにその閉曲線を変形していく。そして、エネルギーの極小状態として形状が決まる閉曲線を、対象物の輪郭とみなすことにより輪郭抽出を行う。輪郭の特徴を表すエネルギー E_{snake} を本研究では以下の(1)式のように定める。

$$E_{snake} = E_{spline}(\mathbf{v}_i) + E_{image}(\mathbf{v}_i) + E_{area}(\mathbf{v}_i) \quad (1)$$

ただし、

$$E_{spline}(\mathbf{v}_i) = \sum_{i=1}^n (w_{sp1} |\mathbf{v}_i - \mathbf{v}_{i-1}|^2 + w_{sp2} |\mathbf{v}_{i+1} - 2\mathbf{v}_i + \mathbf{v}_{i-1}|^2) \quad (2)$$

$$E_{image}(\mathbf{v}_i) = - \sum_{i=1}^n w_{image} |\nabla I(\mathbf{v}_i)| \quad (3)$$

$$E_{area}(\mathbf{v}_i) = \sum_{i=1}^n w_{area} \{x_i(y_{i+1} - y_i) - (x_{i+1} - x_i)y_i\} \quad (4)$$

である。ここで、 w_{sp1} , w_{sp2} , w_{image} , w_{area} は各エネルギーの重み係数である。また、これらの値を変化させることで様々な形状の輪郭を抽出することができる。

(2)式において、右辺第1項のエネルギーは Fig. 4 の赤線で示したベクトルの大きさの2乗の総和である。このエネルギーは、輪郭モデルの周囲長に依存するため、輪郭モデルが収縮し、凸凹が少なくなれば最小化していく。また第2項のエネルギーは Fig. 4 の青線で示したベクトルの大きさの2乗の総和である。このエネルギーは、輪郭モデルの凸凹が直線的になると最小化していく。

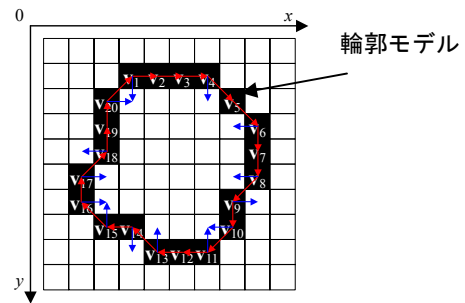


Fig.4 E_{spline} の算出

(3)式において $I(\mathbf{v}_i)$ は輪郭モデルの点 $\mathbf{v}_i$ での画素の輝度値を表す関数である。よって $|\nabla I(\mathbf{v}_i)|$ は輝度勾配の絶対値を表す。(3)式は負の値をとるので、 $|\nabla I(\mathbf{v}_i)|$ が大きいほど、小さな値となる。よってこのエネルギーは輝度勾配の絶対値が大きいところに輪郭モデルが移動すると最小化していく。なお $|\nabla I(\mathbf{v}_i)|$ は(5)式のように計算する。

$$|\nabla I(\mathbf{v}_i)| = |I(x_i + 1, y_i) - I(x_i, y_i)| \quad (5)$$

一般的には輝度勾配は(5)式のように輝度値の x 方向の差分だけではなく、 y 方向の差分との和を用いる場合が多い。しかし輝度値の y 方向の差分をとると、Fig. 1のストライプの境界のような横線部で輝度勾配が大きくなる。そのため輪郭モデルがストライプの境界に誤って移動する可能性がある。よって本研究では(5)式のように輝度勾配を求める。

(4)式において右辺第1項は画像平面 (x, y) 上で y 軸に関して輪郭モデルが囲む面積を表し、第2項は x 軸に関して輪郭モデルが囲む面積を表す。このエネルギーは Fig. 5(a)のように輪郭モデルの点を時計回りに読み込んでいくと正の値をとり、Fig. 5(b)のように反時計回りに読み込んでいくと負の値をとる。よってエネルギーは前者の場合は輪郭モデルが収縮、後者の場合は輪郭モデルが膨張すると最小化していく。

この性質を利用すると、輪郭モデルが凹形状の抽出が可能となる。輪郭モデルの点を時計回りに読み込む

でいくと、凹形状の部分では一時的に読み込みが反時計回りになる。そのため凹形状を膨張させようと輪郭モデルが変形していき、凹形状が抽出できるようになる[8].

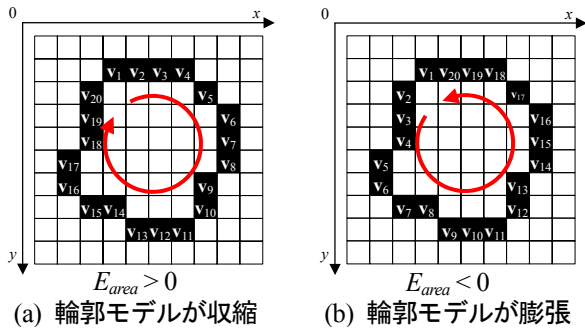


Fig.5 E_{area} の性質

3.2 重み係数の決定

動的輪郭モデルのエネルギーの重み係数は、抽出する対象の形状などにより値を変えなければならない。そのため動画など画像の処理枚数が多い場合、重み係数を自動で決定する必要がある。重み係数決定の際、輪郭が囲む形状の周囲長 l 、面積 S 、円形度 e を参考にする。 e は値が小さいほど形状が複雑なことを示す。 l は輪郭の画素数、 S は輪郭が囲む領域の画素数とし、 e は(6)式で表す。

$$e = \frac{4\pi S}{l^2} \quad (6)$$

各エネルギーの性質として、 l が大きくなると E_{spline} の第1項のエネルギーと E_{image} は大きくなり、 S が大きくなると E_{area} が大きくなる。また形状が複雑、つまり e が小さいほど E_{spline} の第2項のエネルギーは大きくなる。各エネルギーが大きくなるときは重みを小さくすればよいと考えられる。重み係数の自動決定は動画の第2フレーム以降のみで行うとし、第 j フレームのそれぞれの重み係数は(7)式～(10)式のように決定する。

$$w_{sp1}(j) = \frac{a_1}{l(j)} \quad (7)$$

$$w_{sp2}(j) = a_2 e(j) \quad (8)$$

$$w_{image}(j) = \frac{a_3}{l(j)} \quad (9)$$

$$w_{area}(j) = \frac{a_4}{S(j)} \quad (10)$$

ここで、 a_1 、 a_2 、 a_3 、 a_4 は各重み係数決定の際のパラメータで、これらの値はあらかじめ第1フレームにおいて決めておく。また $l(j)$ 、 $e(j)$ 、 $S(j)$ は第 j フレームでの周囲長、円形度、面積である。

以上の処理により、対象物の輪郭を正確に抽出する

ことができる。

具体的には、ストライプ状背景を用いた対象物抽出画像 (Fig. 3) に対して、動的輪郭モデルの初期位置を設定する。初期位置は Fig. 3 における白の背景部分と、抽出された対象物の境とする。そして輪郭モデルを変形していき対象物の輪郭を抽出する。最後に抽出した輪郭の外側を除去することにより、対象物の抽出を行う。

4 実験

提案した手法の有効性を確かめるため、実験を行った。まず、対象物の一部（人物の腕及び胸）に背景と同じ色が含まれた2枚の撮影画像 (Fig. 1) に本手法を用いた。なおこの画像はデジタルカメラを用いて 2400×1800 pixel で撮影したが、対象物抽出に必要な部分を切り取り 1060×1400 pixel にしたものである。

実験結果を Fig. 6 に示す。また抽出された対象物と他の背景を合成した結果を Fig. 7 に示す。

これらの結果を見ると対象物周りのストライプ状背景が除去され、ほぼ正確に対象物を抽出できているのが分かる。

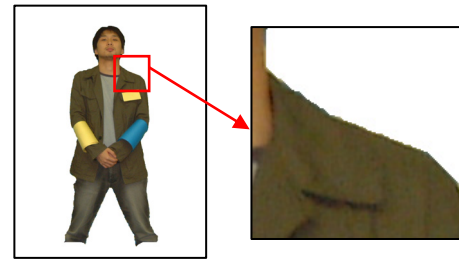


Fig. 6 対象物抽出結果



Fig. 7 合成画像

また、動画画像合成結果を Fig. 8 に示す。(a) が撮影画像、(b) が合成画像である。このときの撮影画像サイズは 640×480 pixel である。なお、動的輪郭モデルの重み係数決定の際、(7)～(10)式におけるパラメータの値は $a_1=1300$ 、 $a_2=1$ 、 $a_3=450$ 、 $a_4=9400$ とした。この結果より、対象物の数や面積、形状が変化しても対象物が抽出できていることが分かる。

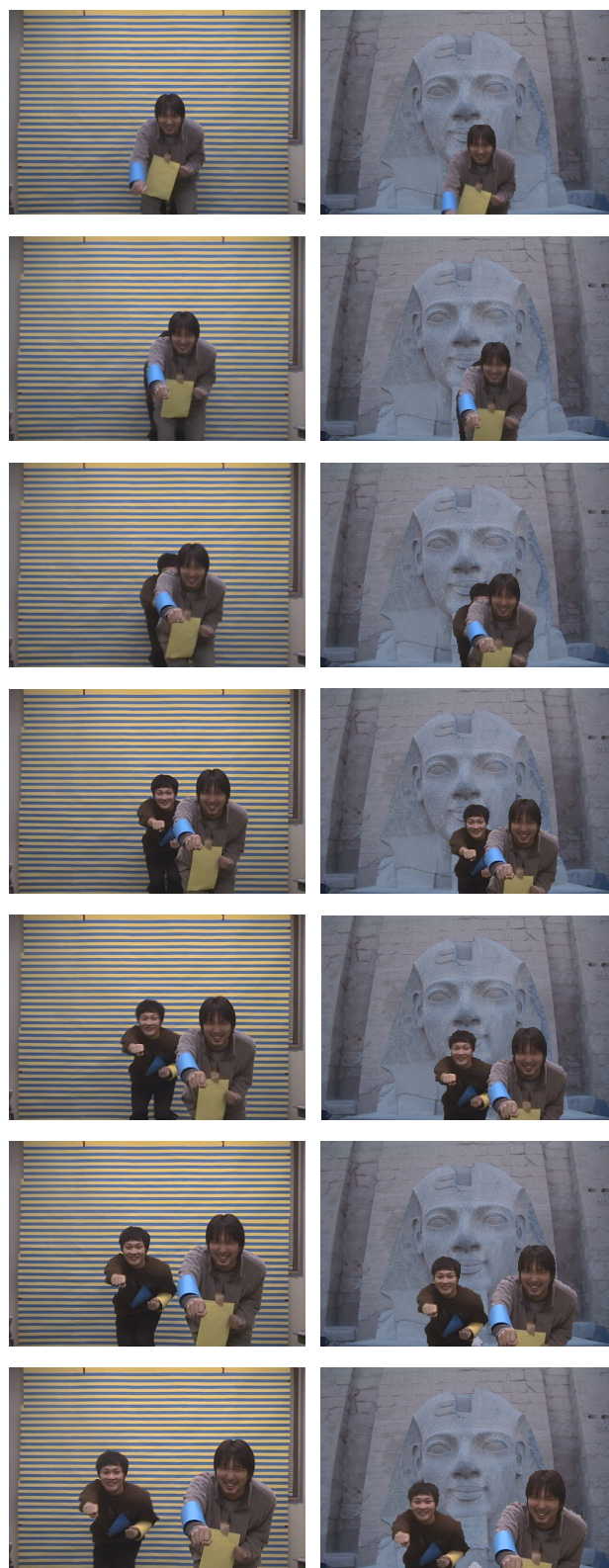
5 結論

本研究ではストライプ状背景を用いた対象物抽出画像に対して動的輪郭モデルを適用することにより、正確に対象物を抽出する手法を提案した。実験結果より、高い精度で対象物の抽出ができていることが確認できた。

今後の課題としては、画像合成の際対象物と背景の境を違和感なく合成していくこと[9]や、位置関係や光の関係を考慮した合成を行っていくことなどが挙げられる。

参考文献

- [1] Alvy Ray Smith and James F. Blinn: “Blue Screen Matting”, *Computer Graphics (SIGGRAPH1996)*, pp.259-268, 1996.
- [2] Takeo Kanade, Atsushi Yoshida, Kazuo Oda, Hiroshi Kano and Masaya Tanaka: “A Stereo Machine for Video-Rate Dense Depth Mapping and its New Applications”, *Proceedings of the 1996 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp.196-202, 1996.
- [3] Masahiro Kawakita, Keigo Iizuka, Tahito Aida, Hiroshi Kikuchi, Hideo Fujikake, Jun Yonai and Kuniharu Takizawa: “Axi-Vision Camera (Real-Time Distance-Mapping Camera)”, *Applied Optics*, Vol.39, No.22, pp.3931-3939, 2000.
- [4] Kazutaka Yasuda, Takeshi Naemura and Hiroshi Harashima: “Thermo-Key: Human Region Segmentation from Video”, *IEEE Computer Graphics and Applications*, Vol.24, No.1, pp.26-30, 2004.
- [5] Atsushi Yamashita, Toru Kaneko, Shinya Matsushita and Kenjiro T. Miura: “Region Extraction with Chromakey Using Stripe Backgrounds”, *IEICE Transactions on Information and Systems*, Vol.87-D, No.1, pp.66-73, 2004.
- [6] 山下 淳, 櫻井 貴之, 金子 透: “ストライプ状背景を用いたクロマキーによる領域抽出”, 映像情報メディア学会技術報告, Vol.28, No.31, pp.5-8, 2004.
- [7] Michael Kass, Andrew Witkin and Demetri Terzopoulos: “Snakes: Active Contour Models”, *International Journal of Computer Vision*, Vol.1, No.4, pp.321-331, 1988.
- [8] 荒木昭一, 横矢直和, 岩佐英彦, 竹村治雄: “複数物体の抽出を目的とした交差判定により分裂する動的輪郭モデル”, 電子情報通信学会論文誌 D-II, Vol.J79-D-II, No.10, pp.1704-1711, 1996.
- [9] Jian Sun, Jiaya Jia, Chi-Keung Tang and Heung-Yeung Shum: “Poisson Matting”, *Computer Graphics (SIGGRAPH2004)*, pp.315-321, 2004.



(a) 撮影画像 (b) 合成画像
Fig.7 動画合成結果