

## 単眼カメラによる舞台上の多人数の追尾と位置検出システムの検討

孫 光宇<sup>†</sup> 田所 嘉昭<sup>†</sup>

<sup>†</sup>豊橋技術科学大学工学部  
〒441-8580 愛知県豊橋市天伯町雲雀ヶ丘 1-1  
E-mail: †{son,tadokoro}@signal.ics.tut.ac.jp

あらまし 本研究では舞台上の移動人物を追尾するために、単眼のUSBカメラにより得られた動画像から、舞台上の多人数の移動人物を認識したり、追尾したりすることを実現する。本研究では、単眼のUSBカメラで撮影された3枚の連続するフレームの各々の差分画像を用い、差分画像に対する閾値処理及び論理積(AND)を求める方法により、舞台上の移動人物の追尾を実現する。舞台のサイズ、カメラの3次元位置、舞台からカメラまでの距離などの既知条件により、カメラの変数と傾き角度を求め、舞台上の移動人物の実際の位置を計算した。また、移動人物が多人数の場合には、個々の移動人物の色情報を利用し、個々の移動人物を認識し、追尾することも検討する。そして、シミュレーションにより、本研究の手法の有効性を確認する。

キーワード 移動人物, 追尾, 位置検出, 差分画像, 色情報

## A Study of Tracking Some People on Stage and Detecting Their Positions Using One Camera

GuangYu SUN<sup>†</sup> and Yoshiaki TADOKORO<sup>†</sup>

<sup>†</sup> Faculty of Engineering, Toyohashi University of Technology  
Hibarigaoka 1-1, Tempaku, Toyohashi, Aichi, 441-8580 Japan  
E-mail: †{son,tadokoro}@signal.ics.tut.ac.jp

**Abstract** A tracking and position detecting system for the people on the stage is developed using one USB camera. This is realized using two different images from three continuous frames. In this system, we evaluate some factors of camera(magnification/focus distance, inclination)using the stage dimension and the camera position. We can track each moving people even when two people overlap using the color information and the position estimation of each people. The experimental results demonstrate that the proposed system can track some people simultaneously and detect their positions.

**Key words** moving people, tracking, position detection, different image, color information

### 1. ま え が き

人間追尾のシステムは、最近いろいろな分野で、応用されている [1]-[7]。例えば、舞台上の歌手位置の追尾、ゲーム中の選手位置の追尾、視覚障害者の歩行支援、自律移動ロボットの人間の追尾などである。

本研究室では、音声のホルマント位相差に基づく追尾システム [1]-[3] を提案した。しかし、このシステムでは、雑音が多い場合、被験者が音声を発しない場合、移動人物が多人数の場合には、各々の移動人物を追尾することが困難である。そのため、新たに画像処理による人間追尾システムを検討した [4][5]。

連続する2枚のフレームの差分画像では、抽出したものが重

なる現象が存在する。そのため、環境の変化が激しい場合には背景差分を用いただけの方法は、移動人物の抽出は不十分である [6][7]。これら为了避免するため、本研究では、連続する3枚フレームの差分画像を用い、その現象を改善した [8]。また、実際には、カメラの傾き角度を測るのが困難であり、さらにカメラの傾き角度が変化する場合も存在する。そのため、本研究では、舞台の上辺と下辺の位置情報とカメラの位置情報を用い、カメラの変数と傾き角度を求めることで、抽出した移動人物の位置を検出する手法を提案した。なお、移動人物は多人数の場合には、個々移動人物を検出、認識、追尾する必要があるため、個々の移動人物の色情報を利用し、それぞれの移動人物を認識する。特に、移動人物が重なった場合には、移動人物の位置予測と移

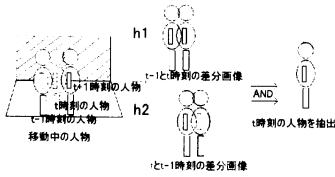


図1 連続する3枚のフレームからの差分画像を用いた人物抽出

動人物の色情報により、重なっている移動人物を検出する。

本研究で使用されるUSBカメラの画素点の数は30万画素であり、入力画像サイズは320×240である。

## 2. システムの概要

### 2.1 連続する3枚のフレームからの差分画像

従来の方法は二つあり、一つは2枚の連続フレームの差分を用い、動画中の移動物体を抽出する手法である(方法1)。もう一つは背景差分により、移動人物を抽出する手法である(方法2)。しかし、方法1では、抽出した移動人物は時刻 $t-1$ のフレームにある移動人物と、時刻 $t$ のフレームにある移動人物が重なったものになる。方法2は、あらかじめ背景画像を撮らなければならない。また、背景の一部が激しく変化すれば、移動人物を抽出することに誤差が生じる。これらの問題を回避するため、3枚の連続フレーム(時刻: $t-1, t, t+1$ )を用い、移動人物を抽出する。図1に示すように、時刻 $t-1$ と $t$ の差分画像 $h1$ と、時刻 $t$ と $t+1$ の差分画像 $h2$ を求める。なお、 $h1$ と $h2$ は、式(1)により2値化を行っている。ここで、 $i(x, y)$ は座標 $(x, y)$ の輝度値である。

$$g(x, y) = \begin{cases} 1 & i(x, y) \geq M \\ 0 & i(x, y) < M. \end{cases} \quad (1)$$

これにより、背景部分の輝度値の変化が、移動人物部分の輝度値の変化より小さいことを利用し、一定な閾値 $M$ を選択し、移動人物を抽出することができる。次に、このようにして得た差分画像に対し、収縮と膨張処理を行うことで、差分画像の雑音を排除する[9]。そして、これらの二つの差分画像 $h1, h2$ の論理積(AND)を求めることで、時刻 $t$ の移動人物を抽出する。

### 2.2 移動人物の位置検出

前節では、動画中の移動人物を抽出する手法を検討した。その手法によって、求めた位置情報は移動人物の実際の空間の位置情報ではなく、カメラによって撮影された画像空間の位置情報であるため、実際の空間の位置に変換する。まず画面上の任意の点を $P'$ を定義し、舞台における $P'$ の対応点を $P$ とする。図2のように、幾何学の計算により、 $P$ 点の $y$ の座標 $P_y$ は式(2)で与えられる。ここで、移動人物の足もとの位置を軌跡点とする。

$$\begin{cases} K = k/f \\ P_y = h \tan(\theta - \tan^{-1}(P'_y \cdot K)) \end{cases} \quad (2)$$

式(2)中の $k$ はカメラの拡大率、 $f$ は焦点距離、 $\theta$ はカメラの傾き角度である。図3を示すように、任意の点 $P$ の $y$ 座標を求めることで、 $P$ の $x$ 座標を求めることができ、式(3)で表される。

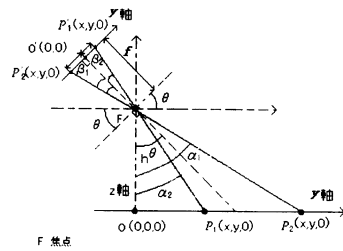


図2 横からみた舞台の様子

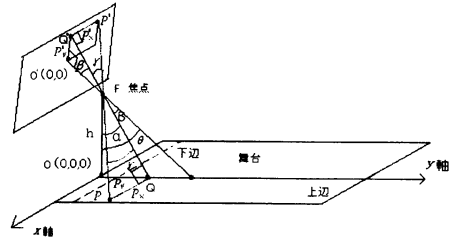


図3 斜めからみた舞台の様子

$$\begin{cases} P_x = h \sin(\alpha - \theta) P'_y / (P'_y \cos \alpha) \\ \alpha = \tan^{-1}(P_y / h) \end{cases} \quad (3)$$

### 2.3 カメラ変数と傾き角度

これまで、移動人物の実際位置を求める方法を提案した。しかし、実際には、カメラの傾き角度 $\theta$ を測定することが困難であり、角度が変化する場合もある。そのうえ、カメラの拡大率 $k$ 、焦点距離 $f$ が既知ではない場合も存在するため、カメラの傾き角度 $\theta$ と変数 $K(k/f)$ を求める必要がある。その手法として、舞台の上辺と下辺の位置情報を既知とし、カメラの角度を検出する方法を提案した。ここでは、舞台の形の特性を用い、移動人物の実際位置を計算する手法を検討する。ここで、次の三つの条件を仮定する。

- (1) カメラは舞台の前に存在し、一定の高さから撮影する。
- (2) 舞台は長方形である。
- (3) 舞台のサイズやカメラの位置やカメラから舞台までの距離を既知とする。

これらを既知条件として、移動人物の実際位置を計算する。図2に示すように、下辺上の点を $P_1$ 、上辺上の点を $P_2$ とし、 $P_{1y}, P_{2y}$ を式(2)に代入すると、式(4)となる。

$$\begin{cases} P_{1y} = h \tan(\theta - \tan^{-1}(P'_{1y} \cdot K)) \\ P_{2y} = h \tan(\theta - \tan^{-1}(P'_{2y} \cdot K)) \end{cases} \quad (4)$$

式(4)から式(5)を導くことができる。

$$\begin{cases} K = \tan(\theta - \tan^{-1}(P_{1y}/h)) / P'_{1y} \\ K = \tan(\theta - \tan^{-1}(P_{2y}/h)) / P'_{2y} \end{cases} \quad (5)$$

式(5)から式(6)を導き出す。

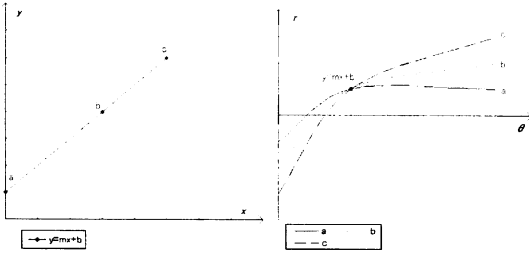


図4 直線関数と直線の Hough 変換

$$\begin{cases} \frac{\tan(\theta - \tan^{-1}(P_{1y}/h))}{\tan(\theta - \tan^{-1}(P_{2y}/h))} = \frac{P'_{1y}}{P'_{2y}} \\ \tan^{-1}(P_{1y}/h) = \alpha_1 \\ \tan^{-1}(P_{2y}/h) = \alpha_2 \end{cases} \quad (6)$$

式(6)から式(7)が導かれる。

$$\frac{\tan(\theta - \alpha_1)}{\tan(\theta - \alpha_2)} = \frac{P'_{1y}}{P'_{2y}} = k' \quad (7)$$

式(7)から式(8)を導き出す。

$$\begin{cases} a = k' \tan \alpha_1 - \tan \alpha_2 \\ b = (1 - k')(\tan \alpha_1 \tan \alpha_2 - 1) \\ c = \tan \alpha_1 - k' \tan \alpha_2 \\ \tan \theta = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} \end{cases} \quad (8)$$

式(8)から、カメラの傾き角度 $\theta$ を求めることが可能である。 $\theta$ を式(5)に代入し、カメラの変数 $K$ を求める。 $\theta$ の値は0より大きいので、正数を選択する。

#### 2.4 直線の Hough 変化

式(8)で、カメラの傾き角度を求めるため、式(4)の画面上の舞台上辺の $P'_{1y}$ と上辺の $P'_{2y}$ を検出しなければならない。舞台が長方形の場合、画像は台形形で表現される。台形の辺は全部直線なので、直線検出によって舞台の辺を検出する。ここでは Hough 変換を用いて直線の検出を行う。 $(x, y)$ 平面の直線の関数は次式となり、

$$y = mx + b \quad (9)$$

式(9)の Hough 変換は次式で表される。

$$r = x \cos \theta + y \sin \theta \quad (10)$$

式(10)中の $r$ は直線から原点までの距離で、 $\theta$ は直線の法線と $x$ 軸の角度である。 $(x, y)$ 平面の直線上の点は図4示すように、 $(r, \theta)$ 空間において、一つの曲線で表れる。これらの曲線の共有点は式(9)の直線に対応し、共有点の個数が直線を構成する点の個数である。この手法を用いて、舞台の辺を抽出する。これにより、舞台上の移動人物の位置を求めることができる。

#### 2.5 移動人物の分離

移動人物が2人以上の場合は、個々の移動人物に対し追尾するために、個々の移動人物を分離しなければならない。今回は、移動人物は最初は分離にしているということを条件とする。図

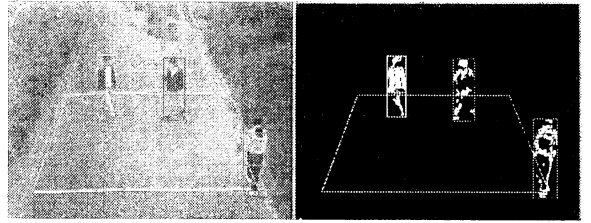


図5 移動人物の分離

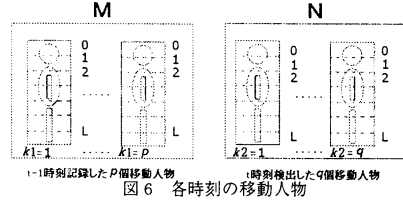


図6 各時刻の移動人物

5に示すように、求めた個々の移動人物の差分画像に対して分析することにより、個々の移動人物の差分画像は $x, y$ 軸において、連続なものである。一つの広い連続の区域は一つの移動人物に対応する。この特徴を利用すれば、各々の移動人物を検出することが可能である。

#### 2.6 移動人物の認識

移動人物が2人以上の場合には、時刻 $t-1$ と $t$ の各々での移動人物を区別することが必要である。図6のように $t-1$ に記録した移動人物 $M_{k1}$ と $t$ での移動人物 $N_{k2}$ を $L$ 等分に分割し、求めた各部分の画素点の平均値 $M_{k1}[i], N_{k2}[i]$ を比較し、二つの値が最も類似する場合には、 $t-1$ の移動人物 $M_{k1}$ に対応する $t$ での移動人物 $N_{k2}$ を同一人物と考える。式(11)で、このアルゴリズムを表す。

$$k1 \leftrightarrow \operatorname{argmin} \left[ \sum_{i=0}^L |M_{k1}[i] - N_{k2}[i]| / (L+1) \right] \quad k2 = 1, 2, \dots, q(11)$$

#### 2.7 位置の予測

移動人物の移動軌跡を予測することで、 $t-1$ 時刻から、 $t$ 時刻の移動人物の運動状況を把握する。ここで、 $t$ 時刻の移動人物の予測位置を $E_t = (x_t, y_t)$ 、 $t-1$ 時刻の移動人物の位置を $E_{t-1} = (x_{t-1}, y_{t-1})$ とする。

$t-1$ 時刻での移動人物の状態が移動である場合には、 $t$ 時刻でのこの移動人物は移動を続ける場合と、急に静止する場合が存在する。ただし、 $t-1$ 時刻での移動人物の状態が静止である場合には、 $t$ 時刻でのこの移動人物は静止状態を保つと考える。そのため、 $t-1$ 時刻の移動人物の状態により、予測位置は式(12)、式(13)で示される。

$$E_t = \begin{cases} (E_{t-1} - E_{t-2}) + E_{t-1} & t-1 \text{ 時刻での移動状態(12)} \\ E_{t-1} & \end{cases}$$

$$E_t = E_{t-1} \quad t-1 \text{ 時刻での静止状態} \quad (13)$$

#### 2.8 移動人物が重なる時の処理及び個々の移動人物追尾

2人の移動人物が重なる時には、抽出した移動人物は1人と

みなされるので、得られた色情報は2人の色情報が混ざったものである。従って、これまでの手法では直接処理できない。本研究では移動人物の移動軌跡を予測することで、移動人物の移動状態を把握することができる。以下の条件を満足すると重なったと判断する。

- (1)  $t-1$  時刻に、移動人物 A, B が合うことを予測する。
- (2)  $t$  時刻に、予測した範囲に移動人物は一つになる。
- (3) この範囲以外に移動人物 A, B と判断できる移動人物は存在しない。移動人物 A, B が急に静止した場合も含む。

特に、条件1の判断として、 $t-1$  時刻の移動人物の範囲を分析し、もし重なる、あるいは重なる可能性が高い場合、移動人物 A, B が重なったと判断する。移動人物 A と移動人物 B が重なった時、後ろの A の体の一部が手前の B に隠されるが、B の体は全部現れることが分かる。ここで、移動人物の色情報を用い、重なる場合の2人の移動人物の位置を抽出することを提案する。 $t-1$  時刻の移動人物の色情報を用い、手前の移動人物を検出し、残った部分は隠された移動人物であると考え。画面上の全ての移動人物を抽出する手順は以下に示す。

- (1)  $t-1$  時刻に、全ての移動人物の  $t$  時刻での状況を予測する。
- (2)  $t$  時刻に、重なった移動人物が存在するかどうかを判断する。
- (3) 重なると、重なった移動人物以外の移動人物を判断し、追尾する。
- (4) 重なった移動人物に対し、 $t-1$  時刻の移動人物の色情報を用い、重なった移動人物を分離し追尾する。

### 3. 追尾実験

#### 3.1 実験条件

ある場所に2本同じ長さの白いテープを平行に貼り、舞台と仮定する。3人の被験者はこの範囲内で移動し、システムは被験者の画面における移動軌跡を記録する。カメラと2本の線の位置に関する変数を表1に示す。

表1 実験の変数

変数	値	説明
$\theta$	未知	カメラの角度
$h$	5.00m	カメラの高さ
$P_1$	-	線1上の点
$P_{1y}$	6.67m	線1(下辺)上の点のy座標値
$P_2$	-	線2上点
$P_{2y}$	11.67m	線2(上辺)上の点のy座標値

#### 3.2 実験の結果と分析

今回は移動人物が3人の場合の実験を行った。図7は各々の移動人物に対する追尾した軌跡を表す。図8、図9、図10は2人の移動人物が重なる前、重なったとき、重なった後の追尾状況を示す。実験結果から、提案した手法により、移動人物が多人数の場合に、個々の移動人物を検出、認識、追尾できることがわかる。また移動人物の移動状況を予測することで、移動人物が重なったことを判断できる。さらに、色情報を使用し、重なった移動人物の分離を実現していることがわかる。検出した

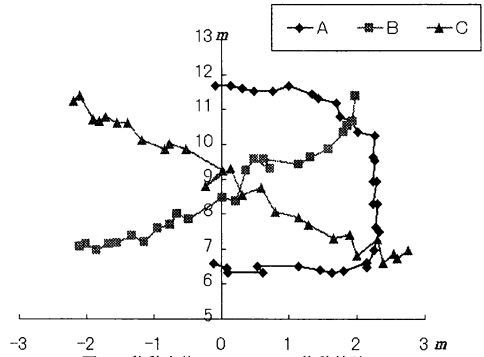


図7 移動人物 A, B, C の移動軌跡

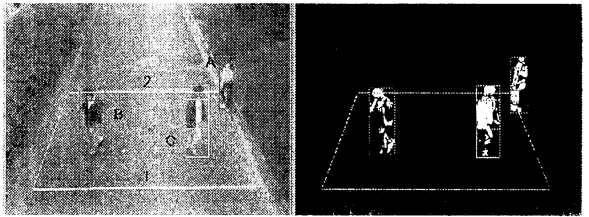


図8 移動人物 A, C が重なる前

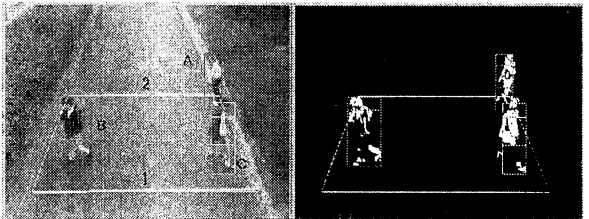


図9 移動人物 A, C が重なった時

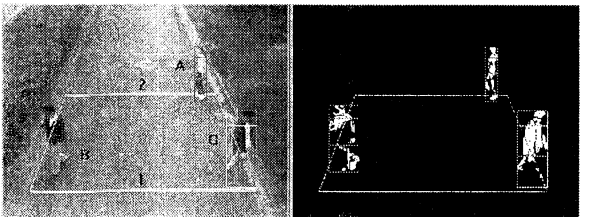


図10 移動人物 A, C が重なった後

表2 移動人物位置の検出誤差

移動者	A	B	C
誤差	2.30cm	3.20cm	3.10cm

移動人物の実際位置の平均誤差を表2に示す。移動人物の実際位置を求める時に、この精度は移動人物の範囲、特に、移動人物の足もとの位置に関係する。移動人物の範囲の抽出が悪ければ、求めた位置も悪くなる。

### 4. 結論

本稿では、連続する3枚のフレームからの差分画像を利用し

て、移動人物を抽出することを検討した。舞台の形(長方形)を用い、カメラの変数  $K$ (拡大率  $k$ /焦点距離  $f$ ) を求め、画像上の移動人物の実際の位置を検出する手法を提案した。また、移動人物が多人数の場合には、移動人物の位置予測手法と色情報から移動人物を認識する手法を組合せ、移動人物の移動状況を把握する。このことより、移動人物が重なることが発生する場合でも、個々の移動人物を分離できることを確かめた。実験の結果により、移動人物の実際位置を検出する精度は、画面上の移動人物の範囲を抽出する精度にかかわることを明らかにした。

今後の課題として、検出した移動人物の実際位置の精度を向上するため、正確に移動人物を抽出する手法を開発する必要がある。また、もっと複雑な環境に対応するためには、提案したシステムのさらなる検討が望まれる。

#### 文 献

- [1] 小幡賢三, 野口健太郎, 田所嘉昭, “ホルマント位相差を利用した音源位置推定”, 電気学会論文誌 C, vol.122-C, no.12, pp.2067-2075, 2002.
- [2] 小幡賢三, 野口健太郎, 田所嘉昭, “音像定位のためのホルマント位相差に基づく移動音源追従”, 計測自動制御学会論文誌, vol.39, no.10, pp.896-903, 2003.
- [3] Kenzo Obata, Kentaro Noguchi and Yoshiaki Tadokoro, “A New Sound Source Location Estimation Algorithm noticing Formant Frequency for Sound Image Localization”, Proceedings of the 20th IEEE Instrumentation and Measurement Technology Conference, IMTC2003, pp.1070-1073, Vail, Coload, USA, 2003.
- [4] 孫光宇, 和賀宗仙, 田所嘉昭, “カメラによる舞台上の移動人物の追尾”, 平成 15 年電気関係学会東海支部連合大会, no.588, 2003.
- [5] 孫光宇, 田所嘉昭, “カメラによる舞台上の多人数の追尾と位置検出システムの検討”, 平成 16 年電気関係学会東海支部連合大会, 0-304, 2004.
- [6] I. Haritaoglu, D. Harwood and L. S. Davis, “W: Who? When? Where? What? A real time system for detecting and tracking people”, FG'98, pp.222-227, 1998.
- [7] 福島健介, 西原明法, “テンプレートマッチングを用いたリアルタイム選手追跡システムの開発とバドミントンシングルのゲーム分析”, 信学技報, pp.329-334, 2003, 06
- [8] 安居院猛, “C 言語による画像処理入門”, 昭見堂, 2002.
- [9] AL BOVIK, “HANDBOOK OF IMAGE AND VIDEO PROCESSING”, AP 出版社, 2000.