

## 遠隔講義のための2台のカメラを用いた指差し位置計測

鈴木浩平 森谷友昭 高橋時市郎

東京電機大学 工学部

### あらまし

情報通信基盤の進歩に伴い大容量の情報伝送が可能になり、教育現場において遠隔講義の形態がとられるようになった。しかしこの遠隔講義において講義資料を提示している画面の一箇所を講師が指し示す際に指示している位置は遠隔地の学生には伝わりにくい。これは、講師を映し出す画面と講義資料を提示する画面とが別々であるためである。この問題を解決するには、講師が画面のどこを指示しているのかをコンピュータで自動的に計測し、遠隔地の受講者に伝えればよい。

本研究では、上記の問題解決のための手段として、2台のカメラを用いて撮影した画像から人差し指の指先の指し示している方向を検出し、ステレオ画像処理により、講師が講義資料提示画面のどの位置を指しているかを計測する方法について検討した。実験の結果、指差し位置計測率98%という、良好な結果を得たので、報告する。

## A Finger Indexing Interface for Distance Education based on Stereo Cameras

Kohei SUZUKI Tomoaki MORIYA and Tokiichiro TAKAHASHI

School of Engineering, Tokyo Denki University

### Abstract

We have proposed a finger indexing interface for distance education lecturers, and developed an algorithm to measure which position of computer monitor or screen is pointed at. The outline of the algorithm is as follows:

1. Detecting the pixels of lecturer's index finger using image processing techniques.
2. Estimating the equation of the straight line tracing along the index finger pixels.
3. Calculating the position on the monitor/screen by solving two simultaneous equations of the lines.

As an experimental result, 98% of finger indexing positions are correctly measured. It verifies the effectiveness of the proposed finger indexing interface.

### 1. まえがき

近年、情報通信基盤の整備や発展に伴って大容量の映像データ転送が可能となり、大学や予備校などにおいて遠隔講義が行われている。

遠隔講義において、講義資料を提示している画

面の一箇所を講師が指し示す際、指示している位置が遠隔地の学生に伝わりづらい。これは、講師を映し出す画面と講義資料を提示する画面とが別々であるためである。この問題を解決するには、講師が画面のどこを指示しているのかをコンピ

ュータで自動的に計測し、遠隔地の受講者に伝えればよい。

こうしたアイデアは古くは 1980 年に MIT Media Lab.の R. A. Bolt が提案したコンセプトモデル “Put that there” [1]に端を発する。Bolt はそのコンセプトの有用性を示すために、データグローブや音声認識装置等を組み合わせた大規模で複雑なシステムを構成した。

その後、操作者が特殊な装置を着装することなく、指差し動作を解釈するインタフェース構築に関する研究が進められた。

例えば、福本ら[2]はステレオカメラ画像処理により指差し位置の計測を行い、インタフェースの改善をはかった。操作者が移動しない場合、操作者の肩の位置の変動は少ない点に着目して、指先と肩を結ぶ直線とプロジェクタ画面との交点を求めている。この手法は精度よく指差し位置を計測できる。しかし、この手法は、(1)予め操作者の肩の位置をキャリブレーション動作により求めておく必要があること、(2)操作者は移動しないことの2点が前提となっている。このため、教卓の前を歩きながら講義をする講師が使うことが十分起こり得る遠隔講義への適用は難しいと言えよう。

本報告では、PC モニタ画面を指差している人差し指の傾きを計測することにより、福本らの手法の制約を緩和することを目指すものである。

そのための前段の処理として、講師が指差し動作を行っているかどうかを画像処理により検出する方法を提案する。次に、2台のカメラを用いて撮影した画像から人差し指を含む手を検出し、ステレオ画像処理により指差し位置を計測する方法を提案する。両手法とも、実際の映像を使ってその有効性を検証した。その結果も併せて報告する。

## 2. 指差し状態の検出

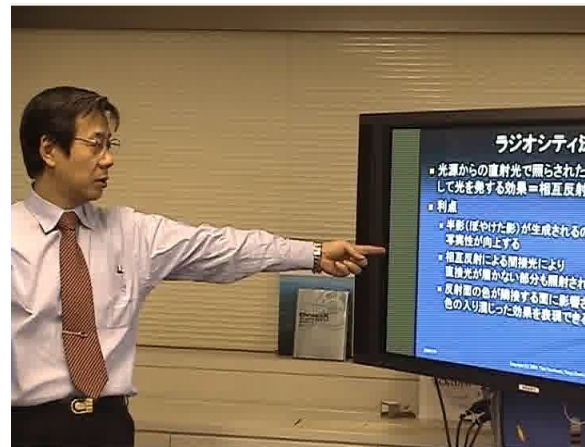
指差し位置を計測する前に、まず被撮影者（講師）が指差ししている状態かどうかを検出する必要がある。指差ししている状態では、指差してい

る手が撮影しているカメラ画像中に現れ、ある程度の時間、静止している。そこで、動画の中から肌色の画素を検出する。前のフレームと比較し、肌色の領域で重なっている画素が多ければ、手は静止していると判断できる。その画素数が閾値を超えて続く状態を指差している状態とすることにした。

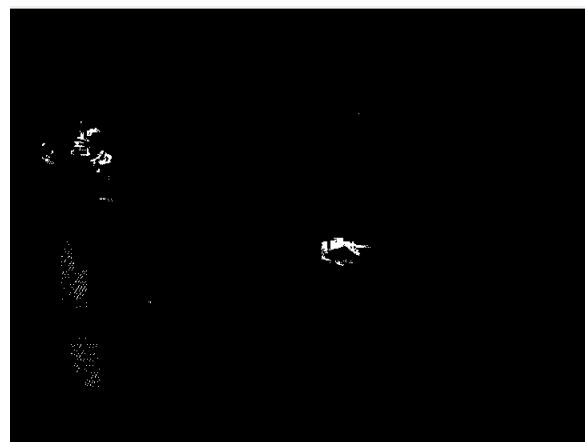
以下、指差し状態の検出処理について説明する。

### 2.1 指差し状態検出処理の流れ

まず、指差し動作をする人の人差し指をビデオカメラで撮影する。ビデオカメラで撮影された画像の中から、指の映っている画素を数画素、人手で指定し、その画素の HSV 色相系での値の平均値を求め、登録しておく。この平均値を肌色値と呼ぶ。



(a) 原画像



(b) 検出結果

図1 指差し状態検出例

次に、実際の講義をビデオカメラで撮影する(図1(a))。撮影された画像中から、予め登録しておいた肌色値と比べ、その差が許容範囲内である画素(肌色画素)を検出する。検出した次のフレーム画像に対しても同様の処理を行い、肌色画素を検出する。

前後2枚のフレームともに肌色画素として検出された画素数を計数する。その値が一定値以上であれば、ある位置で手が静止していたとする。つまり指差している状態であると判定して、検出する。

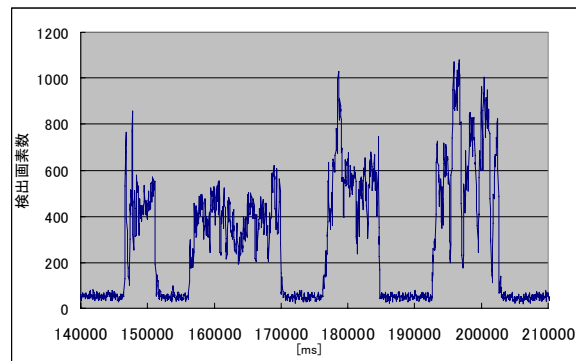
## 2.2 指差し状態検出実験結果

実際の講義において指差し動作を検出できるかどうかを確認するため、実際の講義映像を使って、指差し状態の検出実験を行った。今回は、約4分の実験の講義映像を使った。

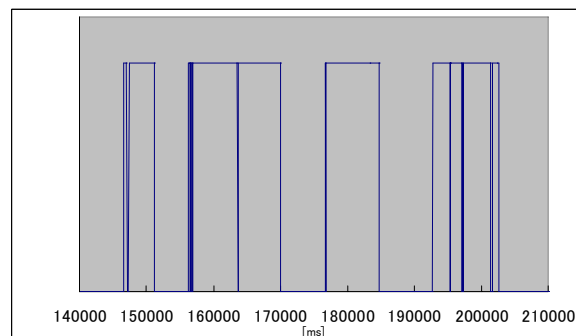
なお、本実験にはCPU: Athlon XP 2600+、メモリ: 512MB、ビデオカード: nVIDIA GeForce FX 5600のPCを用いた。検出処理はリアルタイム(30フレーム/秒)で行われる。画面サイズは640×480画素とした。全フレームに対して検出処理を行った。各フレームとその前のフレームの両方で肌色画素として検出された画素数を求めた。

実験結果を図2に示す。図2(a)はフレーム間で肌色画素と判定された画素数のグラフである。同図(b)は閾値を200画素として二値化した結果である。同図(c)は(b)をもとにして、肌色画素が200画素以上のフレームが続く時間を2.0秒以上とした場合の検出結果である。この実験結果により、指差し状態を自動的に検出できることがわかった。

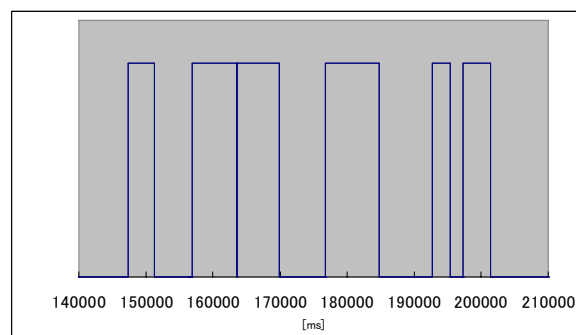
実験結果から、図2(a)に示したように、肌色画素として検出される画素数は常に変動している。このことから、人が画面を指差す精度はそう高くないと推測されるので、システムとして使う場合は注意が必要である。



(a) フレーム間で肌色画素と判定された画素数



(b) 閾値=200画素の場合の検出結果



(c) 継続時間=2.0秒の場合の検出結果

図2 指差し状態検出実験結果

## 3. 指差し位置計測

本章では、指差し位置を画像で計測する処理の概要を述べる。図3に示すように、モニタ上部と側部にカメラを2台設置する。これら2台のカメラ画像を走査して、走査線上に肌色画素が連続して存在した場合、その領域を指のある位置として検出する。数画素おきに画像を走査して、指差している指の傾きを示す直線を検出する。以後、この直線を指差し線と呼ぶ。2台のカメラ画像から求めた指差し線を元に、指差し位置を求める。

### 3.1 指差し位置計算処理

指差しされる対象物（モニタ画面）はカメラに透視投影変換されて映る（図4(a)）。ここで、モニタ画面を指差した指の傾き、つまり指差し線は同図(a)のカメラ画像中の直線①になったと仮定する。画像上に投影されているモニタ画面の四隅の点のy座標が左上隅から反時計廻りにy1, y2, y3, y4、モニタ画面の左辺および右辺と指差し線の交点のy座標をy5, y6とする。モニタ画面の左辺および右辺と指差し線との交点の内分比s, tは次式で与えられる。

$$s = \frac{y_5 - y_1}{y_2 - y_1} \quad t = \frac{y_6 - y_4}{y_3 - y_4}$$

透視投影変換はアフィン変換であるので、この内分比はカメラ画像座標系でもモニタ画面座標系上でも不変である(同図(b))。



図3 2台のカメラの配置例

同様に、上部カメラ画像の場合、yと同様に画像上に投影されているモニタ画面の四隅の点のx座標をx1, x2, x3, x4、モニタ画面の左辺および右辺と指差し線と交点のx座標をx5, x6とする。モニタ画面の左辺および右辺と指差し線との交点の内分比u, vは次式となる。

$$u = \frac{x_5 - x_1}{x_4 - x_1} \quad v = \frac{x_6 - x_3}{x_2 - x_3}$$

よって、点(0, s)と(1, t)、点(u, 1)と(v, 0)を通る2本の直線の連立方程式の解が指差した位置の座標となる(図5)。

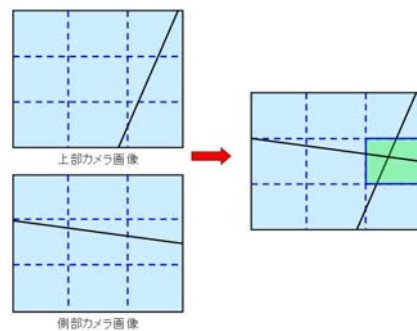


図5 指差し位置計測により求められた指差された点

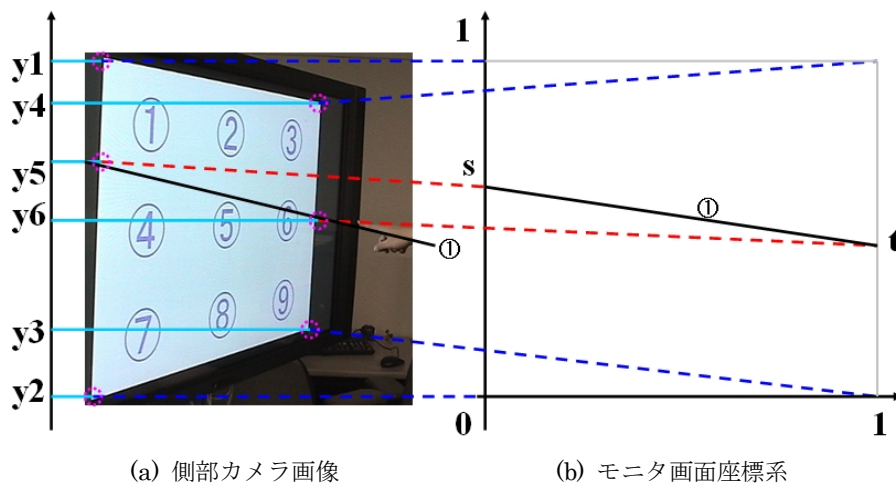


図4 指差し線検出原理



### 3.2 指差し線の検出

指差し線の検出には、2.1 節で予め求めておいた肌色値を用いる。2 台のカメラで撮影された画像中から、予め取得しておいた肌色値と、指定した許容範囲内である画素、肌色画素を検出する。ノイズによる誤差を低減するために、画面を端から順に走査したとき、一定数以上肌色画素が連続する領域を指が存在する領域として検出する。この領域を肌領域と呼ぶ。肌領域の開始画素から指差し線を求める。

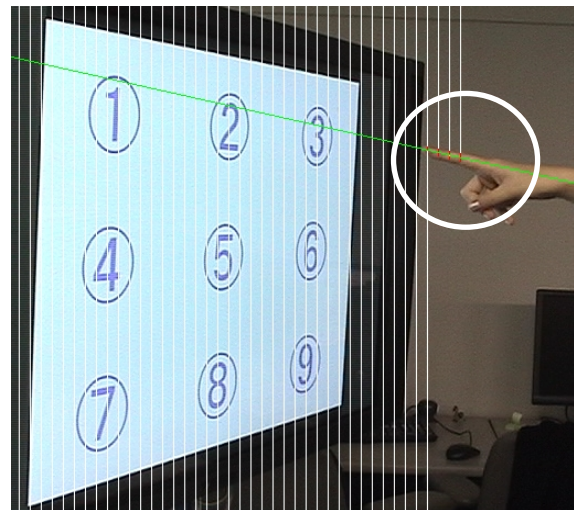
図 6 (a) に示すように、側部カメラ画像の場合、10 画素おきの等間隔に縦に画像を走査し、走査した直線上の画素でのみ肌領域検出処理を行っている。同図(a)では一番右の白線から右から 4 本目の白線が肌領域として検出されている。同図 (b) は上部カメラ画像からの指差し線の検出例である。この例では、一番下の白線から下から 3 本目の白線が肌領域として検出されている。

### 3.3 指差し位置計測実験

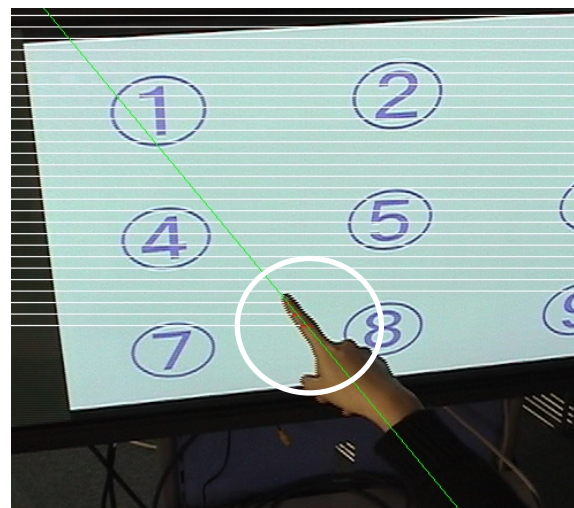
前章の実験結果から、人が画面を指差す精度はそれほど高くない。そこで、画面を縦 3 × 横 3 に粗く分割し、それぞれの小さい領域の中央を指差す動作を繰り返し、必要な精度を確保するインタフェースを考える。

そのようなインタフェースを実現可能かどうかを調べるために、指差し位置がどれくらいの精度で計測できるか、実験を行った。実験は 5 名の被験者（成人男子 3 名、女子 2 名）で行った。50 インチのプラズマディスプレイ（PDP）画面を縦横 3 × 3 の小領域に分割し、9 つの小領域の中央を指で差すように指示した。ただし、画面に向かって正面に立ち、画面から 30cm 離して指差すように指示した。その様子を PDP 画面の上部と左側部に設置されたビデオカメラで撮影し、その映像から指差し位置を計測した。

指差し位置の計測結果を図 7 に示す。肌色値は被験者それぞれの手で求めた値を用いた。その結果、5 名全体での位置計測率 98% を得た。非常に高い精度で指差し位置を計測出来た。



(a) 側部カメラ画像



(b) 上部カメラ画像

図 6 指差し線検出例

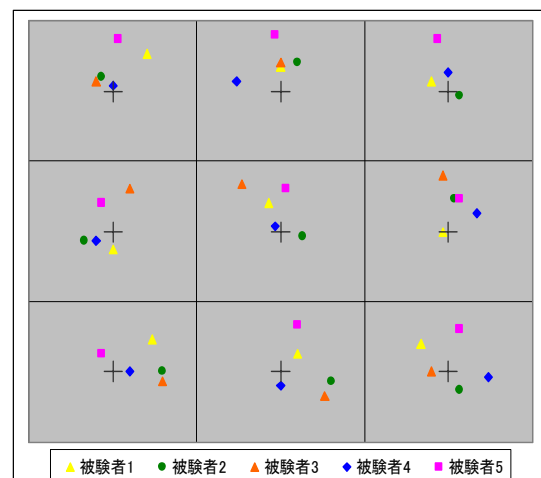


図 7 指差し位置計測結果

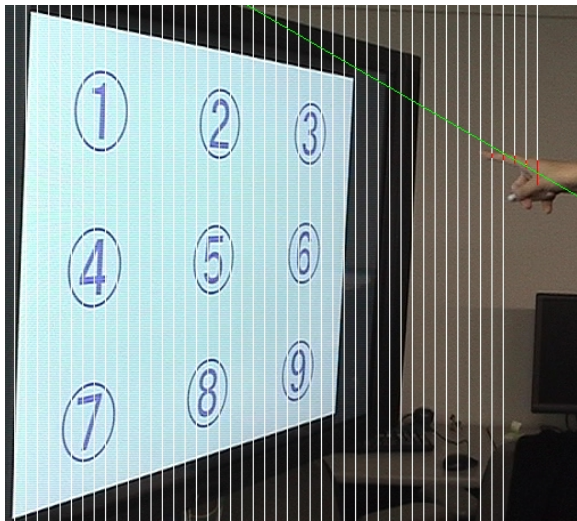


図8 指差し線検出失敗例

誤った一例を図8に示す。図8に示したように、手の甲が肌領域として検出されなかったため、実際の指差し線より上向きに検出されてしまったことによる。肌色領域の能力を向上することによって対応したいと考えている。

#### 4. むすび

遠隔講義の際に、遠隔地にいる講師が指示したモニタ画面の箇所を学生に伝わり易くするために、モニタ画面の指差し位置の計測方法を提案した。2台のカメラからの画像を用いて計測することで、高い精度で実際に指差した位置を計測できた。被験者5名で実験を行ったところ、位置検出率98%を得た。

また、ビデオ映像から指差し状態であるかどうかを検出可能であることも示した。

今後は、肌領域の検出精度の向上をはかるとともに、指とモニタの距離が30cm以上離れた場合や、モニタ画面の正面以外から指差した場合、プロジェクタのような大画面での場合、複雑な背景の場合等、実際の遠隔講義に近い環境において実験を行い、実用面での検討を進めていく。

#### 謝辞

本研究はNTTサイバーソリューション研究所との共同研究の一部として行われた。加藤泰久グループリーダー、田中明通主任研究員、藤本強主任研究員、NTTアドバンステクノロジー 宇井修氏はじめ、関係各位にご指導賜った。記して感謝する。

#### 文献

- [1] Bolt, Richard A. “Put-That-There: Voice and Gesture at the Graphics Interface”, Proc. ACM SIGGRAPH, pp.262-270, 1980.
- [2] M. Fukumoto, K. Mase and Y. Suenaga, “Finger-pointer: Pointing interface by image processing”, Computer & Graphics, Vol.18, No.5, pp.633-642, 1994.
- [3] 尾関基行, 伊藤雅嗣, 里 雄二, 中村裕一, 大田友一, “複合コミュニティ空間における注目の共有～注目誘導行動による物体への注釈付け～”, 日本バーチャルリアリティ学会論文誌, Vol.8, No.4, pp.369-377, 2003.
- [4] 佐藤信, 坂根茂幸, “実空間にマークを投影するインタラクティブハンドポインタの構成”, 電気学会論文誌, Vol.121-C, No.9, pp.1464-1470, 2001.