ヒューマンインタフェース　95－12
コンピュータビジョンと　129－12
イメージメディア
（２００１．９．１３）

ジェスチャ認識のための多視点カメラによる人物位置推定および手領域抽出手法の提案

冨永 将史† 本郷 仁志† 輿水 大和†‡ 丹羽 義典† 山本 和彦¶

†：(財) ソフトピアジャパン HOIP ／ 科学技術振興事業団
‡：中京大学 情報科学部 ¶：岐阜大学 工学部

あらまし 対象人物の要望を察知し，意図に適したサービスを行う空間（パーセプトルーム）の実現を目的とした人間センシングの研究を行っている．人間の身振り手振りや表情に対応して，機器の操作を可能とする技術が確立できれば，快適な生活環境を提供することができるのではないかと考えられる．

本稿では，室内空間の位置に依存しないジェスチャ認識を目指し，壁に設置した複数のカメラからの映像を基に，視体積交差法を利用することで対象人物の位置を推定する手法を提案する．複数カメラ統合画像に背景差分を行い，人物領域を推定する．更に，フレーム間差分を行い，動作領域の推定を行う．これらの結果から視体積交差法により人物領域および動作領域の同定を行い，情報を統合することで，手サインの提示タイミングの検知と，その際の手領域の抽出を行う．本稿では，提案手法の有効性を実験と共に示す．

Position Estimation of Human and Extraction of Hand from Multiple Cameras for Gesture Recognition

Masafumi Tominaga† Hitoshi Hongo† Hiroyasu Koshimizu†‡ Yoshinori Niwa† Kazuhiko Yamamoto¶

†：HOIP, Softopia Japan ／ JST
‡：SCCS, Chukyo University ¶：Faculty of Engineering, Gifu University

Abstract We have been researching human sensing technologies based on computer vision for the realization of the Percept-room as an intelligent environment. This room offers several services such as the control of electrical appliances by use of humans' gestures. A more caring technology is strongly expected for providing more human-friendly interfaces. The control of electrical appliance by use of a gesture is one of the important contributions to the implementation of the Percept-room.

The recognition of human motion in the Percept-room is important as a first priority. In this paper, we propose an estimation method of the human position in the Percept-room by integrating the silhouettes from multiple cameras. We then propose an extraction method of the hand by integrating the motions detected from frame subtractions during a gesture in the Percept-room. We precisely describe the proposed methods and the experimental results obtained by use of a PC system implemented in the Percept-room.

1. はじめに

近年，情報化社会の進歩が目覚しく，一般の家庭環境にも自然にコンピュータが受け入れられる環境が整ってきた．一般家庭のリビングにおいて家電製品を操作する際通常はリモコンが用いられるが，多機能化に伴いリモコンの操作手順は複雑になり，手間も増えており，高齢者社会が進行する現在，人にやさしい機器制御を望む声もあがりはじめた．そのような中，人間の身振り手振りや表情に対応して，機器の操作を可能とする技術が確立できれば，快適な生活環境を提供することができるのではないかと考えられる．そこで，我々のプロジェクトでは，対象者の意図や要望に合ったサービスを提供する空間（パーセプトルーム[1][2]）を提案し，人間の識別をはじめとして，視線認識や動作認識に係わる研究を行ってきた[2][3][4][5]．

このような応用を目的とした研究には，介護医療を考慮した TV コントロールの研究[6]や，物を移動させるような動作を観察する研究[7]の他，音楽指揮を行う研究[8]などがある．更にこのようなシステムの実現へ向けたジェスチャ認識の研究として，形状モデルや動作モデルを用いた手法[9][10][11]や，動き情報を用いる手法[8][12]などが研究されている．これらの研究と並行して，本プロジェクトにおいても，ジェスチャ認識のための研究[2][3][4][5]も行っており，現在，パーセプトルームの実用化へ向けた研究を行っている．

パーセプトルームの構築へ向けた研究の重要な課題の一つに，ジェスチャ認識の前段階である人物追跡と動作タイミングの検知が上げられる．複数のカメラを用いた人物トラッキングの研究には，歩行者道路での流れを観測する研究[13]や，サッカーシーンでの選手の動きを観測する研究[14]，実時間処理を考慮した PC クラスタ化などによる研究[15]などが行われている．更に，室内空間を想定した研究[16] [17]として，仮想空間とのインタラクションやロボット制御を目指したものがあるが，現在の一般家庭の生活環境に対応するものではない．そのような中，現実的な室内空間を想定した研究[18]もあるが，人物検出にトリクロプスカメラによる距離情報を使用しており，画像からのジェスチャ認識による機器制御でなく，人物位置・操作対象に応じた様々な計測機器を用いることで機器制御に対応している．

このような背景から，本稿では，家電製品の操作を目的としたジェスチャ認識を考慮した，人物トラッキングの手法と，その際の手領域の抽出法について述べる．

本稿で提案するシステムでは図 1 に示す環境を想定している．リビングを想定した室内空間において，複数台の固定カメラを設置し，カメラ PC で個別の高解像度映像を取得する．メインとなる PC では，低解像度の全カメラ映像を統合した映像を取得し，室内の人物位置や操作対象機器の推定を行う．対象人物が機器を操作するジェスチャを行う際に，メイン PC は人物認識やジェスチャ認識に最適なカメラを判断し，対応するカメラ PC に対して処理を要求する．認識要求を受けたカメラ PC は取得した高解像度映像の対象領域を認識サーバに処理させ，認識結果をメイン PC に返す．この結果に対してマルチリモコンを用いて室内の対象機器を操作・制御する．なお，本研究におけるジェスチャとは身体全体の動作ではなく，手で行うサインする．

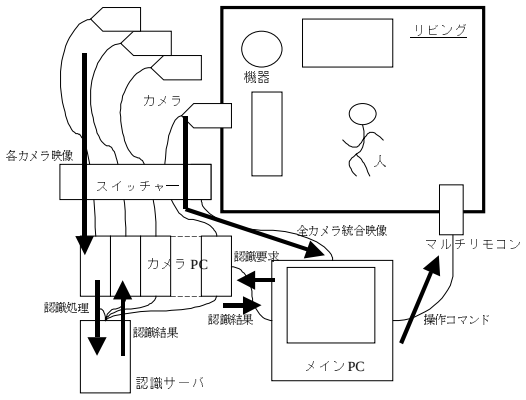


図 1. システムの環境

本稿では，室内空間に配置した複数カメラからの情報を基に，視体積交差法[15][16][17]を利用した人物位置の推定を行い，家電製品の制御を目的とした手サイン提示タイミングの検知を行う手法について述べる．更に，複数カメラのうち，手サインの識別に適したカメラの選択を行う．なお，人物認識や手サインの認識に関しては既存の研究成果[1][2][3][4][5]を利用することを想定している．

2. パーセプトルームにおけるカメラ配置と画像取得の構成

図 2 に室内空間を想定し，データ取得用として構築したパーセプトルームのカメラ配置を示す. 同期の取れたカメラ 16 台を，1 辺 460cm の空間内に設置した．各カメラは 45 度間隔で配置し，8 台を高さ 90cm の位置に水平に，残る 8 台を高さ 220cm の位置に下方 22 度を向くように設置した.

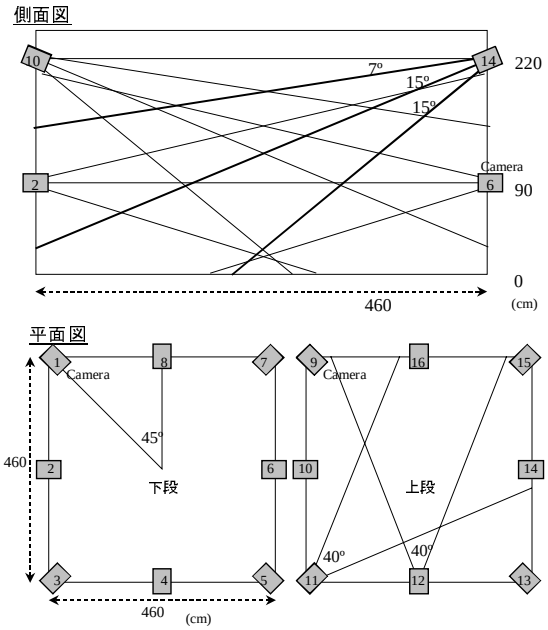


図 2. カメラ配置

16 台のカメラはそれぞれ 1 台の PC（以降カメラ PC）に接続されており，ビデオレートのカラー画像（640×480）を取得できる．更に，各カメラ画像を画面合成スイッチャを通して，1 台の PC（以降メイン PC）に接続することにより図 3 に示す統合されたカラー画像（640×480）として取得できる．図 3 に示す画像領域の番号は, 図 2 におけるカメラ番号に対応している. 上 8 画像領域が下段カメラ画像に相当し，下 8 画像領域が上段カメラ画像に相当する. 各カメラ画像領域のサイズは同一（160×120）である.

メイン PC は得られた 16 カメラ統合画像から人物の位置推定・動作推定を行い，この結果に応じて各カメラ PC に個人認識や手サインの認識要求を出す．カメラ PC は顔領域や手領域の認識要求を認識サーバに処理させ, 得られた認識結果をメイン PC に返す．本稿ではこのメイン PC で処理される人物位置推定と手サインの提示タイミングの検知について述べる.

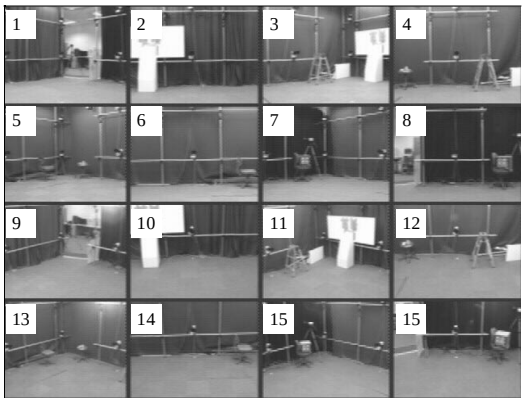


図 3. 16 カメラ統合画像

3. 人物位置の推定とカメラ間での対応

3.1. 人物位置の推定

16 カメラ統合画像から次に述べる方法で人物位置を推定する.

16 カメラ統合画像に対して，事前に取得しておいた背景画像との差分，2 値化を行い，各カメラ画像の領域毎に x 軸方向，y 軸方向への射影処理を行う．射影ヒストグラムを走査し，x 軸方向，y 軸方向共に画素数が閾値を越えた矩形領域を人物候補領域とする．x 軸方向の候補領域において視体積交差法[15][19][20]を適用して人物領域の平面位置を推定する. 視体積交差法やシルエット法は，主に 3 次元形状復元[19][20][21]に用いられる手法であるが，本研究の目的では, 人物の位置推定と動作推定に用いるため，詳細な形状を得る必要はない. そのため分解能を荒く（総ボクセル数を少なく）設定でき高速化が期待できる．なお，視体積交差法を用いる際，全てのカメラパラメータ（カメラ位置，パン・チルト角，焦点距離等）は固定で既知とする.

視体積交差法ではカメラからの 3 次元的なボクセル投影（投票）を行うが，ここではまず x 軸方向の射影ヒストグラムから得られた人物候補領域において平面的な投影を行い, 投票を行う範囲を限定する.

投影平面 F（平面サイズ $N \times M$）に対して，図 4 に下段 8 カメラから投影可能な領域 C_l (l=1,2,…,8：カメラ番号)を示す．投影は各カメラ位置から，カメラに相対する壁側へ錐状の領域に行われる．各カメラからの投影可能領域 C_l は

$$C_l \subseteq F \qquad (1)$$

であり，投影平面 F は 8 つのカメラのうち少なくとも 1 つにより投影，

$$C_1 \cup C_2 \cup \cdots \cup C_8 = F \qquad (2)$$

が成り立つ．図 5 に投影した様子を示す．ここで濃淡値は，カメラからの投影領域の重なりを示す．

通常視体積交差法による形状復元では，精度を要求するため，投影可能領域の積

$$C_1 \cap C_2 \cap \cdots \cap C_8 \qquad (3)$$

即ち全カメラから見える範囲である中央のみを有効とし，実際に投影される領域 T_l (l=1,2,…,8：カメラ番号)の積

$$T_1 \cap T_2 \cap \cdots \cap T_8 \qquad (4)$$

により物体形状を求める．これに対し本手法では，室内空間（投影平面）における位置の推定を目的とすることから，平面全体を網羅する式(2)の範囲を全て有効として投票を行う．

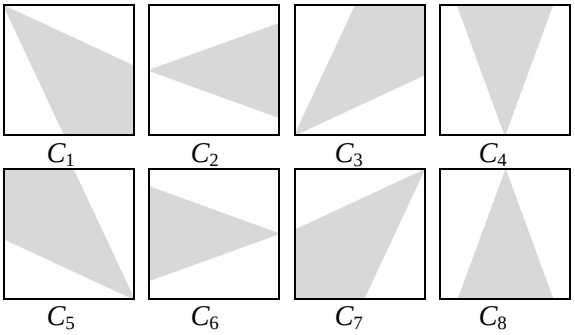


図 4. 下段 8 カメラの投影可能領域

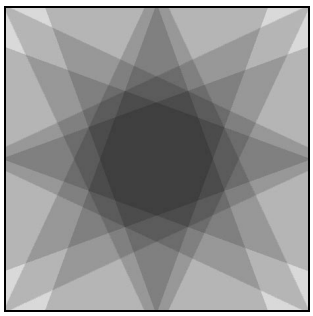


図 5. 8 カメラからの投影可能領域

さらに投影結果から領域を絞り込むため，射影領域範囲外への投影部分になる領域を削除する．各カメラからの投影可能領域 C_l のうち，投影領域 T_l を除いた非投影領域（例：図 6）

$$D_l = C_l - T_l \qquad (l=1,2,\cdots,8) \qquad (5)$$

はすなわち，抽出目的である人物が存在しない背景領域の範囲であるため，削除が可能である．

このようにして，限定した領域の空間にのみ 3 次元的な高さを考慮した視体積交差法による投票を行い，人物位置を求める．

図 7(a)に 8 カメラからの投影領域 T_l の例を，図 7(b)に非投影領域 D_l を削除した投票限定領域 G を示す．

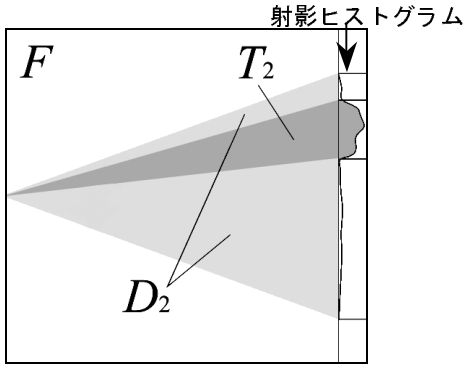
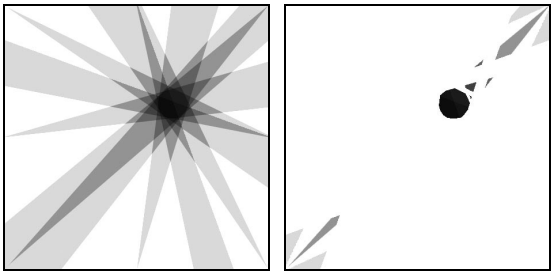


図 6. C_2 における削除可能領域 D_2 の例



(a)投影領域 T_l　　(b)投票限定領域 G

図 7. 8 カメラからの投影

投票によって得られた複数のボクセルに対し，ボクセル数 1 の孤立ボクセルを除去し，ラベリングを行うことで，固まり（ブロック）毎に分ける，更に，体積比較を行い，ブロックの体積が最大であり，前フレームでの人物重心位置に最も近い重心位置のブロックを抽出目的の人物とし，このブロックの重心位置を現在の人物位置とみなす．

3.2. カメラ間の抽出領域の対応

16 カメラ統合画像において，射影により抽出された複数の矩形領域の同一人物対応を，視体積交差法による投票結果を基に推定する．各領域の中心(x_o,y_o)（もしくは領域内の数点）に対し，カメラ位置(x_k, y_k) (k=1,2…8：カメラ番号)から伸ばした直線

$$y=\left(\frac{y_k-y_o}{x_k-x_o}\right)\times(x-x_o)+y_o \qquad (6)$$

が最初に接触したブロックの座標$(x_a,\ y_a)$のラベルを領域に対応付ける．

$$region = Label\left[\min\left((x_k-x_a)^2+(y_k-y_a)^2\right)\right] \qquad (7)$$

全ての領域に対して処理を行い，最終的に人物と見なされているブロックと同一のラベルに対応付けられた全ての領域が，各カメラ画像での同一人物領域となる．

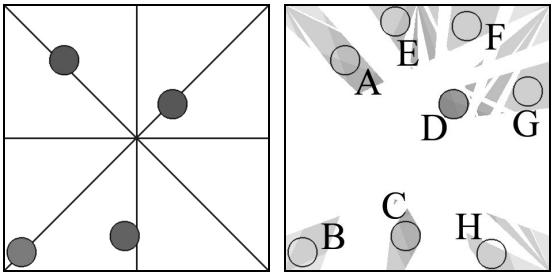
3.3. 複数人への対応

一般家庭のリビングを想定するため，複数人への対応も考慮する必要がある．例として図8(a)に示すような室内に 4 人の人間が存在する場合を考える．その際，下 8 カメラによる射影領域範囲の投影結果（非投影領域削除済み）は図 8(b)となる．この場合，人物同士のオクルージョンにより図 8(b)の○印の領域が人物候補と考えられ，非投影領域の削除だけでは，絞込みはできない．

16 カメラ統合画像のうちカメラ 1～8 に当る画像領域は，図 9 のように想定できる．各ブロックのラベル付けの際，節 3.2 に従いカメラ位置から各領域 a～m へ直線を延ばし，最初に接したブロックのラベルへ対応付ける．対応結果を表 1 に示す．この結果から，

ブロック A は領域 a, g
ブロック B は領域 d
ブロック C は領域 b, e, j, m
ブロック D は領域 c, f ,h ,i ,k ,l

に対応付けられる，ブロック E～H はどの領域にも対応付けられない．このようにカメラから見えない位置でのジェスチャは認識対象とせず，カメラに見えている人物，即ちカメラから最も近い位置にいる人物を対象とし，複数人の場合のブロック対応を行う．



(a) 4 人配置　　(b) 投影領域

図 8. 複数人への対応例

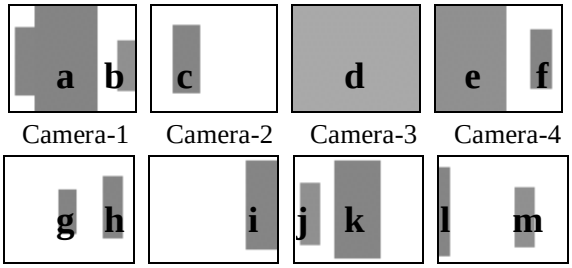


図 9. 複数人時の 1～8 カメラ画像領域例

表 1. 領域とブロックの対応付け

	A	B	C	D	E	F	G	H
Camera-1	a		b					
Camera-2				c				
Camera-3		d						
Camera-4			e	f				
Camera-5	g			h				
Camera-6				i				
Camera-7			j	k				
Camera-8			m	l				

4. 動きと手サイン提示の検知

4.1. 手の動き検出

背景差分では検出できない動き領域の位置推定を行うために，フレーム間差分を用いて動き検出を行い，同一動領域の推定のため視体積交差法を使用する．このように動領域の室内空間位置を推定することで，背景差分のみでは得られない手領域の動きといった身体の部分的な動作が推定できると考えられる．

まず，同一人物として対応付けられた矩形領域の面積比により移動状態か停止状態かを判別する．時間推移における人物対応は，ブロック重心座標の変化が最も少ないブロックを移動後の同一人物として採用することとした．次に，矩形領域の重心位置を比較し，人物が停止状態の時に，高い座標位置に動領域の重心が現

われた場合, 上半身のみを動かしているとみなして手サイン提示の前段階とし, この先頭フレームを手サイン提示開始タイミングとする. この前段階中に得られた動領域の位置から若干広い範囲を, 手サイン提示中の手領域とみなす.

4.2. 手サイン提示時のカメラ選択

手サインの認識要求を出す際, 正面に近いカメラ画像を取得するカメラ PC を選択し, このカメラ PC にのみ認識要求を出す.

視体積によって得た人物ブロックの, 現フレームと 5 フレーム前での重心位置をもとに, 移動方向を推定する. 移動方向を人物正面とみなすことで, 背後からのカメラ PC を選択対象から外す. また, 動領域が検出されないカメラ画像領域に対応するカメラ PC も, 手が映っていないとみなし, 選択対象から外す. 更に, 人物重心位置と動作部分の重心位置の相対位置関係から手の提示方向を推察することも可能であると考えられる.

複数のカメラ PC に対して認識要求を出すことになるが, 各カメラ PC の認識結果から多数決や重要度にもとづき情報を統合することで, 精度の高い最終結果を導く.

5. 実験

パーセプトルームにおいて人物位置推定の実験を行った. 図 10 に背景差分による候補領域抽出結果を示す. 射影からの矩形領域抽出のみでは, 同一人物領域が複数に分割される場合がある. 図 11(a)にカメラ位置から射影領域に対する投影領域 T_l を示す. このとき空間サイズは $(X, Y, Z) = (46, 46, 22)$ とし分解能は 10cm に設定した. なお図中には, 理解し易いよう 2, 4, 6 カメラに正対する壁の位置に, カメラから得られた画像を表示した. 図 11(b)に非投影領域 D_l を削除した投票限定領域 G を示す. このときの有効空間への投票結果を図 12 に示す. 空間中に縦長のブロックが確認できる. 図 13 にブロックに対して領域を対応付けた結果を示す. カメラ 8 およびカメラ 14 で 2 つに分かれていた人物候補領域が同じ領域として対応付けられ, 結合されているのが確認できる.

図 14 に図 10 と同時刻におけるフレーム間差分による候補領域抽出結果を示す. また, 図 15(a), (b)に投影領域 T_l と投票限定領域 G を示す. 更に図 16 に有効空間に投票した結果を示す. このとき手の部分のみを動かしているため, ブロックは上方に浮いた形で抽出できた. 図 17 にブロックに対して人物候補領域を対応付けた結果を示す. 人物候補領域がうまく対応付けられていることが確認できた.

図 18 に手サイン提示開始タイミングとして抽出されたフレーム画像を, 図 19 に手サイン提示中として抽出されたフレームを示す. 図中白線で囲われた領域が最終的に手領域として抽出された領域である. 太い白線で囲まれたカメラ画像領域は最終的に正面に近いカメラとして選ばれたものであり, これに対応するカメラ PC にのみ, 抽出した手領域の座標を送り, 認識要求することになる.

本実験では, 立った姿勢でのサイン提示は良好に抽出できたが, 座った姿勢でのサイン提示では動領域が小さくなることから, 若干不安定であった.

なお, 今回の実験ではメイン PC として Pentium4 1700MHz 256MB Windows2000 SP2 を使用した. 1 フレーム当りの処理時間は平均 150msec であった.

6. 考察とむすび

本稿では, 壁に設置した 16 台のカメラから得られる画像に対して, 視体積交差法を利用することで対象人物の位置を推定する手法を提案し, 実験により有効性を示した. 更に, 動領域を監視することにより, 手サインの提示タイミングの検知と, 手サインの認識に最適なカメラ画像の選択手法を提案し実験により有効性を示した.

本稿では, メイン PC における処理に関して述べた. パーセプトルームの実現には, 認識処理を実際に行うカメラ PC や機器制御部を考慮する必要がある. そのため, 現在ソケット通信を用いたカメラ PC との連携や家電製品の制御を実時間で行うシステムを構築している. 通信速度や制御のディレイを考慮すると, 実時間で全ての処理を行うためにはメイン PC での処理を 100msec 以下にすることが望ましいと考えられる.

図 10. 背景差分による候補領域抽出結果

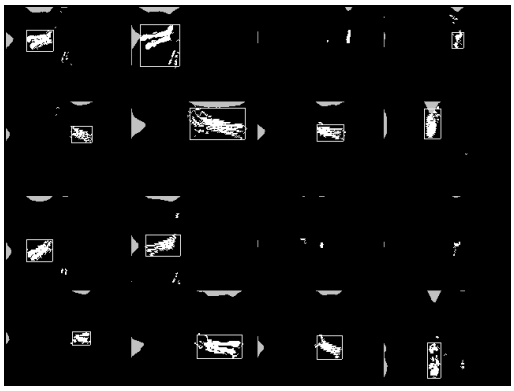
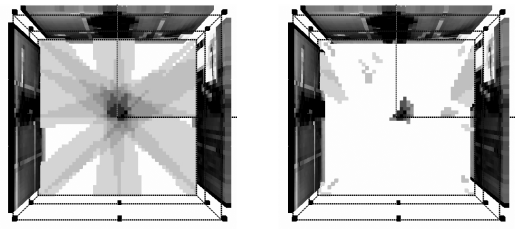
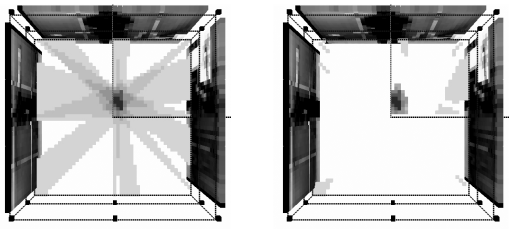


図 14. フレーム間差分による候補領域抽出結果



(a) 投影領域 T_l　　(b) 投票限定領域 G

図 11. 背景差分からの投影



(a) 投影領域 T_l　　(b) 投票限定領域 G

図 15. フレーム間差分からの投影

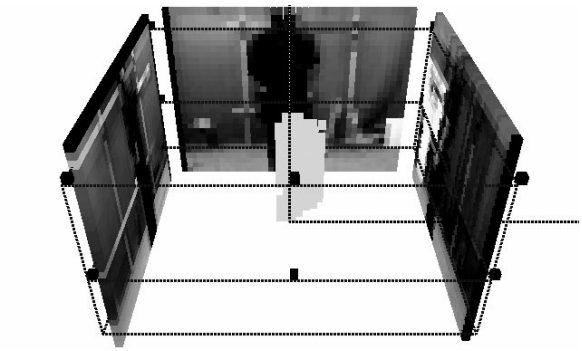


図 12. 背景差分からの投票結果

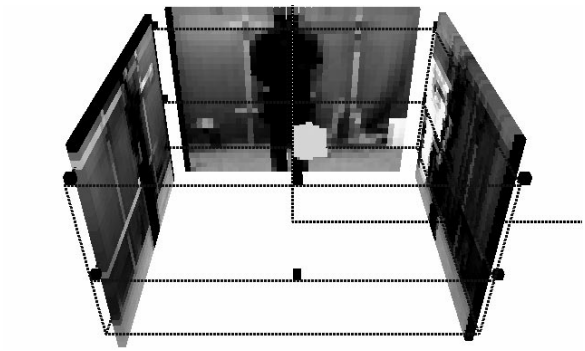


図 16. フレーム間差分からの投票結果

図 13. 背景差分による人物領域抽出結果

図 17.フレーム間差分による人物領域抽出結果

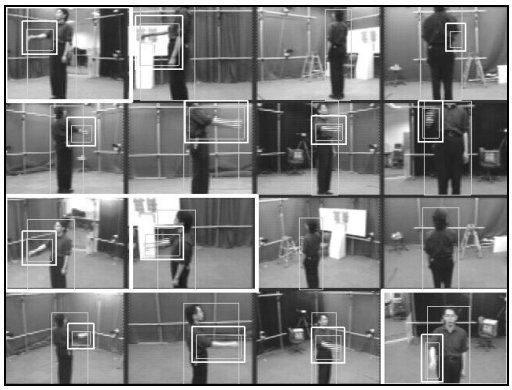


図 18. 手サインの提示開始フレーム

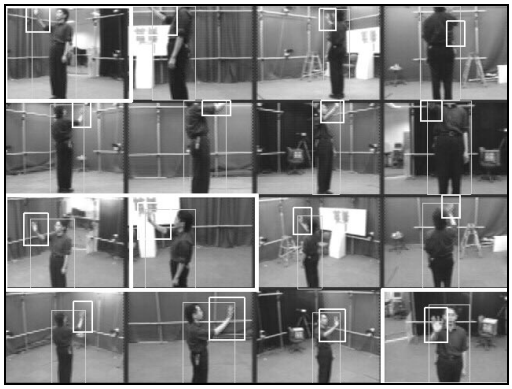


図 19. 手サイン提示中の手領域

謝辞

本稿の執筆にあたり，御助言頂いた岐阜県地域結集型共同研究推進室主任専門研究員安本護氏，同主任専門研究員ジャン・クリストフ・テリヨン氏に深く感謝致します.

参考文献

[1] 本郷仁志，安本護，渡辺博己，ジャン・クリストフ・テリヨン，山本和彦：“パーセプトルーム構築のための多方向顔画像データベース開発”，電子情報通信学会技術研究報告, PRMU2000-108, pp.7-12, (Nov.2000).

[2] Hitoshi Hongo, Hiroki Watanabe, Mamoru Yasumoto and Kazuhiko Yamamoto: “Detection of Facial Parts Occluded by Hands”, 7th Korea-Japan Joint Workshop on Computer Vision – Frontiers of Computer Vision (FCV2001), pp.140-145 (Feb.2001).

[3] Hitoshi Hongo, Mitsunori Ohya, Mamoru Yasumoto and Kazuhiko Yamamoto: “Face and hand gesture recognition for human-computer interaction”, 15th International Conference on Pattern Recognition (ICPR2000), Vol.2, pp.925-928 (Sep.2000).

[4] 安本護，本郷仁志，渡辺博己，山本和彦，輿水大和：“マルチカメラ統合を用いた人物識別と顔方向推定”，電子情報通信学会誌，Vol.J84-DII, No.8, pp.1772-1780, (Aug.2001).

[5] 本郷仁志，山本和彦：“動領域内の肌色推定による顔領域および顔部品抽出”，映像メディア学会誌，Vol.52, No.12, pp.1840-1847 (Dec.1998).

[6] 川野卓也，山本和彦，加藤邦人，本郷仁志，丹羽義典：“高次局所自己相関特徴の相対位置関係を用いた TV コントロールのためのポーズ認識”，第7回画像センシングシンポジウム (SSII2001) 講演論文集，pp.341-346 (Jun.2001).

[7] 川崎広一，久野義徳：“行動認識を用いた記憶支援システム”，第 7 回画像センシングシンポジウム(SSII2001)講演論文集, pp.351-356 (Jun.2001).

[8] 渡辺孝弘，李七雨，谷内田正彦：“インタラクティブシステム構成のための動画像からの実時間ジェスチャ認識手法－仮想指揮システムへの応用”，電子情報通信学会誌, Vol.J80-DII, No.6, pp.1571-1580, (Jun.1997).

[9] 萬上圭太，岩井儀雄，谷内田正彦：“統計的動きモデルを用いたジェスチャ姿勢推定”，画像の認識・理解シンポジウム(MIRU2000)講演論文集 II, pp.103-108, (Sep.2000).

[10] 大垣健一，岩井儀雄，谷内田正彦：“動きと形状モデルによる人物の姿勢推定”，電子情報通信学会誌，Vol.J82-DII, No.10, pp.1739-1749, (Oct.1999).

[11] 佐藤明知，河田聡，大崎喜彦，山本正信：“多視点動画像からの人間動作の追跡と再構成”，電子情報通信学会誌, Vol.J80-DII, No.6, pp.1581-1589, (Jun.1997).

[12] 畠直志，岩井儀雄，谷内田正彦：“動き情報と情報圧縮を用いたロバストなジェスチャ認識手法”，電子情報通信学会誌, Vol.J81-DII, No.9, pp.1983-1992, (Sep.1998).

[13] 長谷川為春，馬原徳行，全炳東：“複数視点映像による歩行者天国の観測”，第 7 回画像センシングシンポジウム(SSII2001)講演論文集, pp.417-422 (Jun.2001).

[14] 三須俊彦，苗村昌秀，境田慎一，鄭文涛，金次保明：“複数情報の融合によるサッカー選手のロバストな追跡法”，電子情報通信学会技術研究報告, PRMU2001-67, pp.23-30, (Jul.2001).

[15] 松山隆司：“分散協調視覚：プロジェクトの成果と今後の展望”，第 7 回画像センシングシンポジウム(SSII2001)講演論文集, pp.187-198 (Jun.2001).

[16] Aaron Bobick, Stephen Intille, Jim Davis, Freedom Baird, Claudio Pinhanez, Lee Campbell, Yuri Ivanov, Arjan Schütte, Andy Wilson. “The KidsRoom: A Perceptually-Based Interactive and Immersive Story Environment”, November 1996. (Appears in PRESENCE: Teleoperators and Virtual Environments, 8(4), August 1999. pp. 367-391.)

[17] Brooks, R. A. with contributions from M. Coen, D. Dang, J. DeBonet, J. Kramer, T. Lozano-Perez, J. Mellor, P. Pook, C. Stauffer, L. Stein, M. Torrance and M. Wessler: “The Intelligent Room Project”, Proceedings of the Second International Cognitive Technology Conference (CT'97), (Aug. 1997).

[18] Brumitt, B., Meyers, B., Krumm, J., Kern, A., and Shafer, S: “EasyLiving: Technologies for Intelligent Environments”, Handheld and Ubiquitous Computing, (Sep. 2000).

[19] ウ小軍，和田俊和，東海彰吾，松山隆司：“平面間透視投影を用いた並列視体積交差法”，情報処理学会論文誌：コンピュータビジョンとイメージメディア,Vol.42, No.SIG6(CVIM2), pp.33-43 (Jun.2001).

[20] 濱崎省吾，吉田裕之，重永信一：“多視点シルエット画像からの高速な3次元形状復元手法”，第7回画像センシングシンポジウム (SSII2001) 講演論文集，pp.59-64 (Jun.2001).

[21] 藤原孝幸，輿水大和，藤村恒太，藤田悟朗，野口孔明，石川猶也：“顔の 3D モデル化と 3D 似顔絵生成”，第7回画像センシングシンポジウム(SSII2001)講演論文集, pp.323-328 (Jun.2001).