

パノラマベースドアノテーション手法

円筒面・平面への再射影と輝度勾配に着目した位置合わせの安定化

Improvement of panorama-based annotation overlay using a fast and robust image registration method with cylindrical and projective transform

興梠 正克 蔵田 武志 坂上 勝彦

Masakatsu Kourogi Takeshi Kurata Katsuhiko Sakaue

産業技術総合研究所 知能システム研究部門

m.kourogi@aist.go.jp

Abstract 本稿では、ビジョンを備えたウェアラブルコンピュータ着用者の位置・方位を実時間取得する手法とそのアプリケーションについて述べる。提案手法は、着用者視点カメラから得られる画像列を、撮影地点が既知であるパノラマ画像群と照合・位置合わせすることで、着用者の位置・方位を同定する。撮像系の異なるカメラ画像とパノラマ画像群を正しく照合・位置合わせするため、提案手法は、カメラ画像をパノラマ画像の投影面(円筒面または平面)上に再射影し、その投影面上でアフィン変換による位置合わせを行う。撮像系と照明条件の相違による写り方(輝度値)の変化に対して影響を受けにくい輝度勾配値を画像間で一致させる位置合わせパラメータを推定する。提案手法を適用して、着用者の位置・方位に応じて、ライブ映像上に実環境に関する注釈情報を提示するアプリケーションを開発し、ウェアラブルビジョンシステム上に実装した。

1 はじめに

近年、計算機、撮像・表示デバイス、各種センサー、無線通信端末機器等の小型化・高性能化が急速に進んでいる。その進歩に伴い、常時着用と常時動作を前提としたウェアラブルコンピュータの研究が活発となっている[1][2][3]。ウェアラブルコンピュータが日常的に活用されるためには、明示的な説明や指示を与えられることなく、着用者の置かれている状況を認識する要素技術が欠かせない。視覚(画像)情報は、計算機による着用者の状況認識に有用な手がかりを含んでおり、同時に人間にとっても理解が容易なメディアである特長を持つ。

このような背景を踏まえ、筆者らは着用者視点のカメラから得られる画像に基づいて、状況認識のうち最も重要な一つである着用者の位置・方位同定と、その応用例として、状況に応じた実物体への注釈提示に関する研究を進めている[13][14][15]。他の従来研究では、Feinerらは屋外での利用を前提として、GPS(Global Positioning System)と磁気コンパス等を用いた位置・方位同定を行い、ライブ映像上に道標等の注釈情報を重畳表示するプロトタイプシステムを試作している[4]。しかしながら、これらの位置・方位センサーは利用可能な環境が制限される問題点がある。

また、多くの画像ベースの従来研究では、人工的なマーカ群(fiducials)を実環境中に配置し、画像中に検出されるマーカの位置に基づいて、カメラの位置・方位を推定する手法を用いている[5][6][7]。しかしながら、マーカを物理的に配置する必要があるため、広範囲な環境をカバーすることは困難である。一方で、青木らは位置センサーやマーカを用いることなく、画像の色ヒストグラムと事前に獲得された学習データに基づいて着用者の位置同定と軌跡追跡を行う手法[10]を示しているが、方位を正確に同定することが難しいという問題点がある。

そこで、筆者らは、実環境中の各地点で撮影されたパノラマ画像群を事前知識として用いて、着用者視点の画像をパノラマ画像群と照合・位置合わせすることで、位置・方位を簡易に取得するアプローチ(図1)を提案してきた[14]。画像間の位置合わせ手法として、画像間の輝度差分を最小化するアフィン変換による位置合わせパラメータを高速推定する手法[11][12]を用いていた。しかしながら、一般的には、円筒面または平面に投影されているパノラマ画像と平面に投影されているカメラ画像をアフィン変換で位置合わせすることはできない。そのため、アフィン変換による近似的な位置合わせが可能な範囲でしか、正しく動作しない問題点があった。また、パ

ノラマ画像と着用者視点の画像では撮像系が異なる上に、撮影時刻等の相違により、照明条件が変化する。このため、画像間の輝度差分を最小化するパラメータが必ずしも良好な位置合わせ結果を与えない問題点があった。

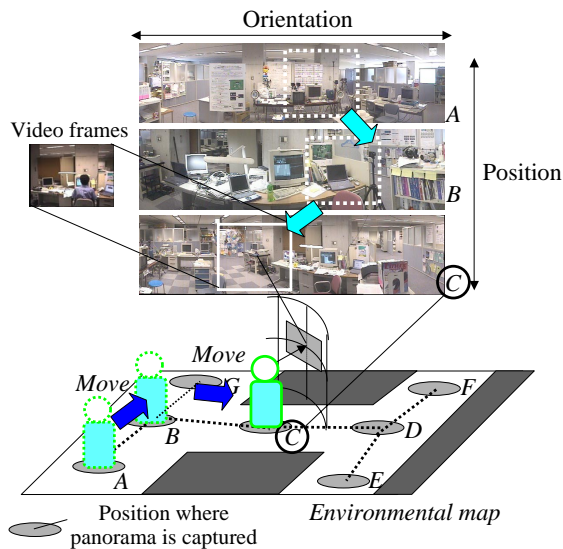


図 1: 画像間の照合・位置合わせに基づく位置・方位同定

そこで、本稿では、カメラ画像をパノラマ画像の投影面上に再射影して、その投影面上でアフィン変換による位置合わせパラメータを推定することで、前者の問題点を解決する。撮像系と照明条件の相違に起因する後者の問題点については、位置合わせパラメータ推定において、最小化すべき評価関数として画像間の輝度差分と輝度勾配値の差分の線形和を用いることでその解決を試みる。

以下、2節では投影面の異なる画像間の位置合わせ手法を述べ、3節では我々が試作しているウェアラブルビジョンシステムと、その上で動作する注釈提示アプリケーションについて述べる。4節では、提案手法を用いた画像間の位置合わせ精度と着用者の位置同定精度を評価する。5節でまとめを行う。

2 位置・方位同定のための画像間の位置合わせ手法

2.1 処理の流れ

本研究は、実環境に関する事前知識として、撮影地点とその方位が既知であるパノラマ画像群を用い

る。本手法は、ウェアラブルカメラから得られた画像をパノラマ画像の投影面に再射影し、その投影面上でパノラマ画像群との照合・位置合わせを行う。処理の流れを図 2 に示す。

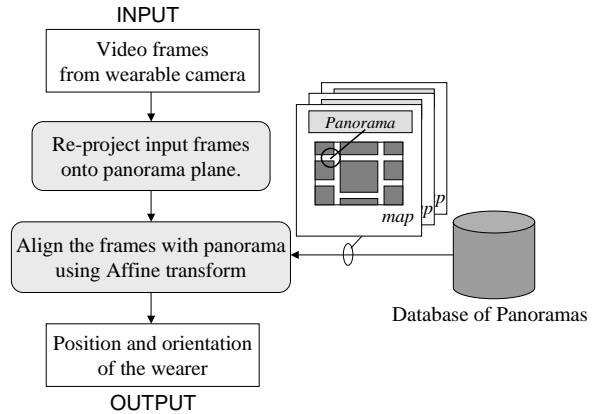


図 2: 提案手法の処理の流れ

以下の節で各処理の詳細について述べる。

2.2 パノラマ画像投影面への再射影

通常のカメラにより撮影される画像(以下、フレーム画像)は、ピンホールカメラモデルに基づいて画像平面上に中心射影されていると見なせる。一方、パノラマ画像は、光軸が一定(水平)方向に固定された画像平面、または円筒面上に射影されている。このため、これらの画像間をアフィン変換を用いて位置合わせすることは一般的にはできない。そこで、本研究では、フレーム画像を各パノラマ画像の投影面に再射影することで、アフィン変換を用いた画像間を位置合わせを可能とする。

パノラマ画像の投影面に再射影するためには、フレーム画像の中心点に対応するパノラマ画像上の位置 (x_P, y_P) が必要であるが(図 3 を参照)、事前にその位置を知ることはできない。そこで、初期状態では、パノラマ画像上を十分に小さな一定の角 $\Delta\theta$ 間隔で水平・垂直方向を網羅する点の集合 S を与える。このような点集合のうち、点 (x_P, y_P) から十分近くにある点が少なくとも一つ存在することが保証される。これらの各点の位置に基づいて再射影されたフレーム画像に対して後述の画像間の位置合わせを行い、その中でもっとも高い評価値を与える結果を最終的な位置合わせ結果とする。

円筒面パノラマ画像では、与えられた点 (x_P, y_P)

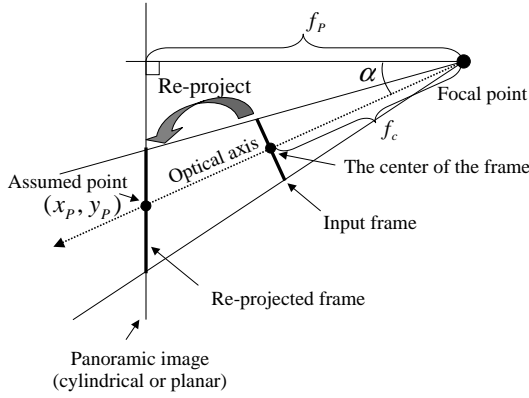


図 3: フレーム画像の再射影

より、焦点からの仰角 α と方位角 β を以下の式に基づいて算出できる。パノラマ画像の焦点距離を f_P とすると、

$$\alpha = \tan^{-1} \left(\frac{y_P}{f_P} \right) \quad (1)$$

$$\beta = \frac{x_P}{f_P} \quad (2)$$

したがって、フレーム画像上の点 (x, y) は、以下の式によりパノラマ画像上への点 (x', y') に再射影することができる。フレーム画像の焦点距離を f_c とすると、

$$x' = f_P \tan^{-1} \left(\frac{x}{y \sin \alpha + f_c \cos \alpha} \right) + f_P \beta$$

$$y' = f_P \beta \left(\frac{y \cos \alpha - f_c \sin \alpha}{y \sin \alpha + f_c \cos \alpha} \right)$$

平面上に投影されたパノラマ画像についても、同様に再射影することができる。

2.3 輝度値とその勾配値に着目した位置合わせ手法

アフィン変換によるグローバルな動き推定手法 [11][12] は、画像間の輝度差分を最小化するパラメータを推定する。しかしながら、入力画像とパノラマ画像の撮像系は異なる上に、撮影時刻の相違等により照明条件は変化する。このため、それぞれの画像上の同一物体が同一の輝度値を持つことは保証されない問題点がある。そこで、本研究では動き推定手法を改良して、輝度値と輝度勾配ベクトルの絶対値の線形和を最小化することで、この問題点を解決する。

提案する画像間の位置合わせパラメータ推定の手順は以下に示す。なお、 $I_f(x, y), I_P(x, y)$ をそれぞれ

入力画像とパノラマ画像の点 (x, y) 上の輝度値とし、 $I'_f(x, y), I'_P(x, y)$ をそれぞれの画像の輝度勾配ベクトルの絶対値（勾配値）とする。

1. 入力画像上で輝度勾配ベクトル $(\frac{\partial I}{\partial x}, \frac{\partial I}{\partial y})$ の絶対値が一定のしきい値以上の画素を選別する。
2. 選別された画素上で以下の式に基づいて仮動きベクトル (u_p, v_p) を生成する。これは輝度値と勾配値による仮動きベクトルの線形和に相当する。ここで、 c は定数、 w は重み係数である。

$$u_p = -(1-w) \frac{\partial I}{\partial t} / \frac{\partial I}{\partial x} - cw \frac{\partial I'}{\partial t} / \frac{\partial I'}{\partial x} \quad (3)$$

$$v_p = -(1-w) \frac{\partial I}{\partial t} / \frac{\partial I}{\partial y} - cw \frac{\partial I'}{\partial t} / \frac{\partial I'}{\partial y} \quad (4)$$

ここで、

$$\frac{\partial I}{\partial t} = I_P(x, y) - I_f(x, y) \quad (5)$$

$$\frac{\partial I'}{\partial t} = I'_P(x, y) - I'_f(x, y) \quad (6)$$

とする。輝度値のみから仮動きベクトルを生成する場合 ($w = 0$ の場合に相当) と比べて、輝度値の変化による同一の画素上で生成される偽の動きベクトル成分が弱められ、照明条件等の変化に対してロバストな仮動きベクトルを生成できると考えられる。

3. 生成された仮動きベクトルによる画素単位マッチングを行う。以下の式、

$$(1-w) |(I_P(x+u_p, y+v_p) - I_f(x, y))| + cw |(I'_P(x+u_p, y+v_p) - I'_f(x, y))| < \theta$$

を満たすとき、この仮動きベクトルとその画素位置をデータ集合に追加する。

4. 収集されたデータ集合を以下のアフィン変換の式への統計的当てはめを行う。

$$x' = a_0x + a_1y + a_2 \quad (7)$$

$$y' = a_3x + a_4y + a_5 \quad (8)$$

ここで、 $x' = x + u_p, y' = y + v_p$ である。

5. 推定されたアフィン変換パラメータを用いて各点における動きベクトル (u_b, v_b) を算出して、 $I_f(x, y), I'_f(x, y)$ の画素値を動きベクトル分だけシフトして、手順 2-5 を反復計算する。推定されたパラメータが収束するか、一定の反復回数を経過したとき、反復計算を終了する。

6. 得られた位置合わせパラメータによる画像間の輝度値と輝度勾配値の正規化相関値を算出して、位置合わせ結果の評価関数とする。

3 ウェアラブルビジョンシステムと注釈提示アプリケーション

筆者らはビジョンを備えたウェアラブルコンピュータによる様々な可能性を追求するため、実世界で稼働するシステムを構築している [16][17]。

3.1 システム

本システムは着用者が装着する計算機(ノートPC)と超小型 CCD カメラ(東芝, IK-SM43H), 着用型ディスプレイ(MicroOptical社, Clip-On Display), センサー(Intersense社, InterTrax²), 無線ネットワーク装置と, 外部の計算処理装置としてのPCクラスタ(CPU: Dual PentiumIII, OS: Linux, 計4台)から成る。着用者視点の画像は, ウェアラブルPCによりキャプチャ, 圧縮され, 無線LANを介してクラスタ側へと伝送される。クラスタ側の処理結果は圧縮され, 再び無線LANを介してウェアラブルPCに伝送される。本システムの構成を図4に示す。

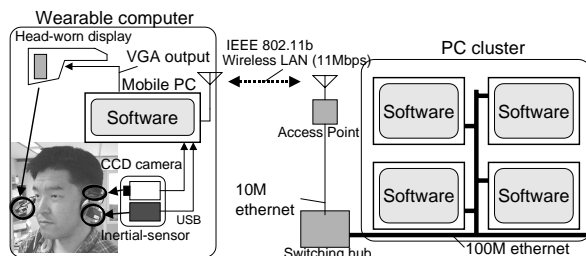


図 4: システムの構成図

3.2 注釈提示アプリケーション

実環境に関する注釈情報や実環境にリンクしたメモ書き(リマインダ)などをライブ映像上に重畳表示する注釈提示アプリケーションは, 拡張現実感とウェアラブルコンピュータにおけるもっとも有力な応用例の一つである [8][9]。

パノラマ画像上に提示すべき注釈情報の位置を事前に与えておくことで, 2節で述べた入力画像とパノラマ画像の位置合わせ手法を用いて, 注釈付き映像

を簡易に生成することができる。その概要を図5に示す。

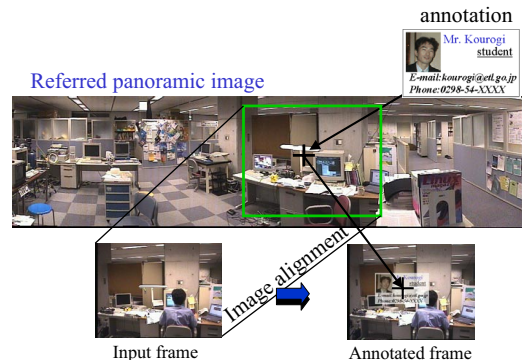


図 5: 注釈提示手法の概要

4 実験

2節で述べた画像間の位置合わせ手法を評価する。最初に, 画像間の正しい位置合わせ結果が既知である入力画像列とパノラマ画像を用いて, 提案手法による画像間の位置合わせ精度を評価する。円筒面への再射影なしのアフィン変換による位置合わせ精度との比較評価する。次に, 本システムの着用者が実環境中を既知の軌跡上を移動するオンライン実験を行い, 本手法による位置の推定精度とその処理のスループット, 遅延時間を評価する。

4.1 位置合わせ精度の評価

着用者視点で撮影されたフレーム画像列100枚(毎秒15フレーム, 画像サイズ: 320×240画素)に対して, 事前知識として1枚のパノラマ画像(画像サイズ: 2000×636画素)を用いて評価実験を行う。パノラマ画像上で特徴が著しい点(人間が視認により確実に識別できるもの)を取り, これらの点が各フレーム画像上で対応する点を人手により6点選び出して正解データとして作成した。パノラマ画像は, 全方位撮像系 HyperOmni Vision (アコウル社, 仰角15度~60度まで撮影可能) [18]で撮影した全方位画像を円筒面に展開して生成した。全方位画像の撮影には, デジタルカメラ(ソニー, DSC-F505V, 334万画素)を用いた。なお, 位置合わせ計算における式(3)(4)の定数 $c = 2.0$, 重み係数 $w = 0.25$ とした。各フレーム画像に対するアフィン変換および提案手法による位置合わせ結果の誤差グラフを図6に示す。この結果より, ほぼ全フレーム画像に対して, 提

案手法による位置合わせの精度は、再投影なしのアフィン変換による位置合わせよりも5-10ピクセル程度高いことが分かる。また、仰角が ± 30 度の範囲を越えると、アフィン変換による位置合わせ結果は失敗することがあったが、提案手法による位置合わせは安定して良好な位置合わせ結果を得ることができた。

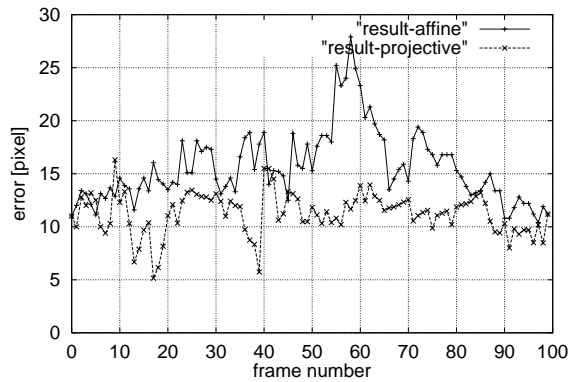


図 6: 位置合わせ結果の誤差

また、撮影時刻が昼夜で異なる照明条件が大きく変化した画像間の位置合わせに対しては、改良前の手法 [11][12] と比べてロバストな結果を与えることを確認した。

4.2 位置同定精度の評価

室内環境を撮影したパノラマ画像群を事前知識として用いて、ウェアラブルビジョンシステムによる位置・方位取得と注釈提示を行うオンライン実験を行った。図 7 に示す十字印上の地点にてパノラマ画像 (計 45 枚) を撮影した。その一部を図 8 に示す。図中の十字印の間隔は 50cm に相当する。各々の十字印に振られた番号はパノラマ画像の ID を表している。

この環境中を着用者は図 7 に示す軌跡上を歩いて移動する。その際に得られた位置同定結果を図 9 に示し、着用者ディスプレイ上に出力された注釈提示結果の一例を図 10 に示す。この結果より、概ね 2 ブロック分 (100cm) 以内の誤差で着用者の位置を取得できることが確認できる。

全処理のスループットは毎秒 9-12 フレーム程度であり、遅延時間は 500-800 ミリ秒であった。これは主としてウェアラブル PC 側での画像取り込み、圧縮、送受信、表示の処理に起因する。クラスタ側での投影面への再射影と画像間の位置合わせ処理は、スループット毎秒 15 フレーム、遅延時間 100-150 ミリ

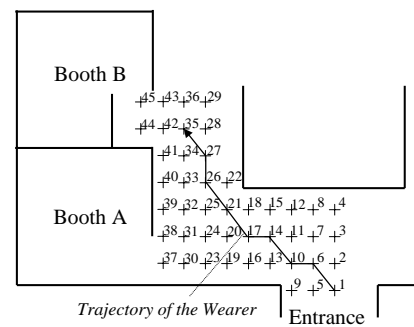


図 7: パノラマ画像群の撮影位置と着用者の移動軌跡



図 8: 実験に用いたパノラマ画像群 (一部)

秒であった。

5 まとめ

本稿では、ビジョンを備えたウェアラブルコンピュータ着用者の位置・方位を実時間取得する手法とそのアプリケーションについて述べた。入力画像をパノラマ画像の投影面に再射影することにより、画像間の位置合わせ精度を向上させることができた。また、輝度値と輝度勾配値を評価基準として考慮した位置合わせ手法により、輝度差分のみを評価基準とした手法と比べて、照明条件の変化に対してロバストとなることを確認した。

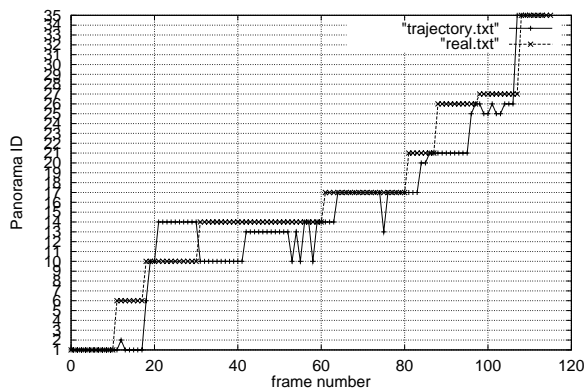


図 9: 位置同定結果

謝辞 本研究は経産省リアルワールドコンピューティング (RWC) プログラムの一環として行われた。

参考文献

- [1] M. Lamming and M. Flynn, "Forget-me-not Intimate Computing in Support of Human Memory," Technical Report EPC-1994-103, RXRC Cambridge Laboratory, 1994.
- [2] S. Mann, "Wearable Computing: A first step toward Personal Imaging," *Computer*, Vol. 30, No. 2, pp. 25–32, 1999.
- [3] T. Starner, J. Weaver and A. Pentland, "A wearable computer based American sign language Recognizer," in *Proc. Int'l Symp. on Wearable Computers (ISWC97)*, pp. 199–202, 1997.
- [4] S. Feiner, B. MacIntyre and T. Höllerer, "Wearing It Out: First Steps Toward Mobile Augmented Reality Systems," in *Mixed Reality – Merging Real and Virtual Worlds*, pp. 363–377, Ohmsha-Springer Verlag, 1999.
- [5] U. Neumann and J. Park, "Extendible object-centric tracking for augmented reality," in *Proc. of VRAIS'98*, pp. 148–155, 1998.
- [6] T. Starner, S. Mann, B. Rhodes, J. Levine, J. Healy, D. Kirsh, R. Picard and A. Pentland, "Augmented reality through wearable computing", in *Presence: Teleoperator and Virtual Environments*, Vol. 6, No. 4, pp. 386–398, 1997.
- [7] D. Koller, G. Klinker, E. Rose, D. Breen, R. Whitaker and M. Tuceryan, "Real-time vision-based camera tracking for augmented reality," in *Proc. of VRST'97*, pp. 87–94, 1997.
- [8] R. Azuma, "A survey of augmented reality," in *Presence: Teleoperators and Virtual Environments*, Vol. 6, No. 4, pp. 355–385, 1997.
- [9] J. Rekimoto and K. Nagao, "The world through the computer: Computer augmented interaction with real world environments", in *Proc. of UIST '95*, pp. 29–36, 1995.
- [10] H. Aoki, B. Chiele and A. Pentland, "Realtime Personal Positioning System for a Wearable Computers," in *Proc. Int'l Symp. on Wearable Computers (ISWC99)*, pp. 37–43, 1999.
- [11] 興梠, 村岡, "グローバルなアフィン動きパラメータの実時間推定手法", 信学論 (D-II), Vol. J82-D-II, No. 7, pp. 1161–1170, 1999.
- [12] M. Kouroggi, T. Kurata, J. Hoshino and Y. Muraoka, "Real-time image mosaicing from a video sequence," in *Proc. Int'l Conf. on Image Processing (ICIP99)*, Vol. 4, pp. 133–137, 1999.
- [13] 興梠, 蔵田, 坂上, 村岡, "パノラマ画像群を位置合わせに用いたライブ映像上への注釈提示とその実時間システム," 第5回知能情報メディアシンポジウム論文集, pp. 17–24, 1999.
- [14] M. Kouroggi, T. Kurata, K. Sakaue and Y. Muraoka, "A panorama-based technique for annotation overlay and its real-time implementation," in *Proc. Int'l Conf. on Multimedia and Expo (ICME2000)*, TA. 2.05, 2000.
- [15] M. Kouroggi, T. Kurata, K. Sakaue, and Y. Muraoka, "Improvement of panorama-based annotation overlay using omnidirectional vision and inertial sensors," in *Proc. Int'l Symp. on Wearable Computers (ISWC2000)*, pp. 183–184, 2000.
- [16] 蔵田, 坂上, 興梠, 村岡, "シリーズ: RWCプロジェクトにおけるコンピュータビジョン研究3: 気が利いたウェアラブルビジョンシステムの実現に向けて - 状況把握に有効な要素機能の実時間実装 -" 画像ラボ, Vol. 11, No. 6, pp. 49–54, 2000.
- [17] 蔵田, 大隈, 興梠, 坂上, "ハンドマウス: 正規混合分布による色モデルの動的生成と平均値シフトを用いた手の検出と追跡", 第7回画像センシングシンポジウム (SSII2001), 2001.
- [18] 山澤, 八木, 谷内田, "移動ロボットのナビゲーションのための全方位視覚系 HyperOmni Vision の提案", 信学論 (D-II), Vol. J79-D-II, pp. 698–707, 1996.

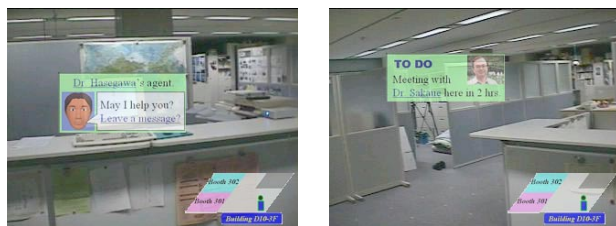


図 10: 注釈付き映像の出力例