社団法人　電子情報通信学会
THE INSTITUTE OF ELECTRONICS,
INFORMATION AND COMMUNICATION ENGINEERS

信学技報
TECHNICAL REPORT OF IEICE.

ジェスチャインタフェースのための画像認識とフィードバックの構成論

中野　克己† 　吉本　廣雅†† 　近藤　一晃†† 　小泉　敬寛† 　中村　裕一††

† 京都大学大学院 工学研究科
〒 606-8501 京都市左京区吉田本町
†† 京都大学 学術情報メディアセンター
〒 606-8501 京都市左京区吉田本町
E-mail: †{nakano,yoshimoto,kondo,koizumi}@ccm.media.kyoto-u.ac.jp, ††yuichi@media.kyoto-u.ac.jp

あらまし　ジェスチャインタフェースの課題として，種々の状況や環境に適応し，常に良好な認識状態を保つことがあげられる．そのために本研究はユーザからの支援を上手く引き出す手法を提案する．これは，画像認識状態をユーザへフィードバックすることで，ユーザに動作や周囲の環境を変更する必要性を認識させ，状況を改善するための支援行動へと誘導するものである．本稿ではこのような手法の構成論とその試作について報告を行う．
キーワード　ヒューマン・コンピュータ・インタラクション，画像認識，ジェスチャ認識

Design of Image Recogintion and Feedback for Gesture Interface

Katsumi NAKANO†, Hiromasa YOSHIMOTO††, Kazuaki KONDO††, Takahiro KOIZUMI†, and Yuichi NAKAMURA††

† Graduate School of Engineering, Kyoto University Yoshidahonmachi, Sakyo, Kyoto, 606-8501 Japan
†† Academic Center for Computing and Media Studies, Kyoto University
Yoshidahonmachi, Sakyo, Kyoto, 606-8501 Japan
E-mail: †{nakano,yoshimoto,kondo,koizumi}@ccm.media.kyoto-u.ac.jp, ††yuichi@media.kyoto-u.ac.jp

Abstract　One of the most important points in gesture interface is to keep good performance against condition or environmental changes. This paper introduces a novel framework that prompts the user's assistance for keeping good performance. For this purpose, the system gives feedback that intuitively shows the necessity of changing his/her movements or changing of environments. This paper introduces the system design for this framework and a prototype system build in our experiments.
Key words　Human Computer Interaction, Iamge Recogintion and Gesture Recognition

1. はじめに

ジェスチャを利用したユーザインタフェースの実現方式として，カメラで撮影した画像のみからユーザのジェスチャを認識し動作する方式がある．その実用化を考えると頑健な画像認識処理をいかに実現するかが大きな課題の一つになる．何故ならこれらの画像認識手法は設定された前提条件が成り立つ場合にしか頑健に動作しないためである．画像認識の頑健化に関する研究は幅広くなされており様々な手法 [1], [2] が提案されているが，ユーザの体格，服装，動作パターンの差異，時間経過による周囲環境の変動等，実世界で起こりうる状況の複雑さ多様さを考慮すると，現状ではジェスチャの誤認識や認識精度低下の問題がすべて解決できているとは言えない．

そこで我々は現実的なシステム構築方針として，環境が前提条件を満足するように，ユーザからの支援を誘発し，ユーザに自然な形で環境を維持させる枠組みを提案している [3], [4]．またこの枠組みによりジェスチャ認識が頑健になることでユーザビリティの向上も期待できる．以下本稿ではこの枠組みを構成論として整理し，ユーザに負担を掛けることなく効果的な支援を引き出す方法として，前提条件が満たされている度合いをユーザへフィードバックする方法と，そのフィードバックを起点としたユーザの支援行動のシナリオについてその設計指針を述べる．さらにその実証例として試作したジェスチャ認識システムについて，その実装と実験結果について報告する．

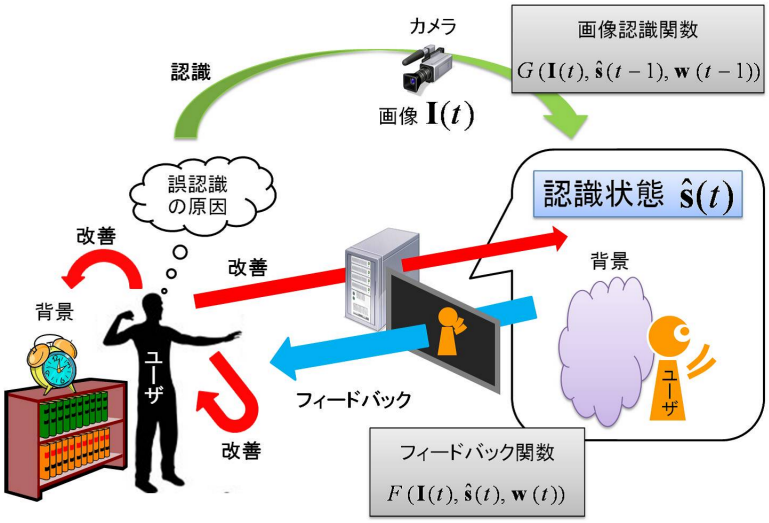

図 1　ジェスチャインタフェースのための画像認識とフィードバックの構成

2. ジェスチャインタフェースのための画像認識とフィードバックの構成論

まず本研究が想定するジェスチャインタフェースはカメラで撮影した画像からユーザのジェスチャを認識し稼働するシステムとする．一般に画像認識の問題は不良設定問題となり，様々な制約条件を設けなければ解を求めることができない．そこで多くの場合，システム設計者は，予めいくつかの**前提条件**を設け，それら前提条件が充足された場合でのみ動作する画像認識アルゴリズムを用いることになる．

本研究はこのようなジェスチャインタフェースについて，誤認識や認識精度低下の問題とその改善方法を議論する．そのためにジェスチャインタフェースのシステム構成を図 1 のように定式化する．

まずシステムはカメラによりシーンを撮影し，画像 $\boldsymbol{I}(t)$ を得る．ここでシーンがシステム設計者の想定した状況内にあれば，そのシーンとその撮影画像 $\boldsymbol{I}(t)$ は状態ベクトル $\boldsymbol{s}(t)$ とその状態方程式 $\boldsymbol{s}(t+1) = f(\boldsymbol{s}(t))$ からなる**状態モデル**で説明できるものとする．具体例を挙げると，ユーザが一人だけカメラの前に存在する状況を想定すると，$\boldsymbol{s}(t)$ はユーザの姿勢，体格や肌，服装の色等に対応するパラメータ群と，周囲環境たとえば背景にある物体の配置や色情報等に対応するパラメータ群とで構成されるベクトルとなる．また状態方程式はユーザの運動や背景の変動を記述したものになる．

この状態モデルを用いると，画像認識処理ではカメラによる観測結果である画像 $\boldsymbol{I}(t)$ から状態モデルのパラメータ $\boldsymbol{s}(t)$ を計算すれば良いことになる．しかし前述のようにこの計算は不良設定問題となる．そこでさらに様々な前提条件を設けることで，画像認識処理は $\boldsymbol{I}(t)$ の時系列から $\boldsymbol{s}(t)$ の推定値 $\hat{\boldsymbol{s}}(t)$ を計算する処理として実現されることになる．良く採用される手法としては，初期値 $\boldsymbol{s}(0)$ を既知とするものや，$\hat{\boldsymbol{s}}(t-1)$ の推定結果が正しいと仮定することで，$\boldsymbol{I}(t)$ から $\hat{\boldsymbol{s}}(t)$ を計算するものがある．このような処理を本研究では関数

$$\begin{bmatrix} \hat{\boldsymbol{s}}(t) \\ \boldsymbol{w}(t) \end{bmatrix} = G(\boldsymbol{I}(t), \begin{bmatrix} \hat{\boldsymbol{s}}(t-1) \\ \boldsymbol{w}(t-1) \end{bmatrix})$$

として定義する．ここで $\boldsymbol{w}(t)$ は，例えばパーティクルフィルタアルゴリズムにおけるパーティクル群のような，画像処理アルゴリズムが計算過程で用いる**内部表現**である．

この定式化に従うと，ジェスチャインタフェースには様々な前提条件が潜在しており，それら前提条件が満たされない状況で誤認識，すなわちシーンと $\hat{\boldsymbol{s}}(t)$ の間に齟齬が生じる．そこでこれら前提条件を**外的前提**と**内的前提**の二つに大別する．

- 外的前提：シーンが想定内の状況にあること
- 内的前提：計算機内部の認識状態 $\hat{\boldsymbol{s}}(t-1)$ が実状態 $\boldsymbol{s}(t-1)$ と合致していること

これら二つの条件が同時に満たされないと，画像認識関数 G はシステム設計者の想定通りに動作することができず，結果誤認識が生じる．そして，実世界の複雑性や多様性を考えると，計算機がこれら 2 条件の充足判定を自動で行うことは原理的に不可能である．

そこで本研究は，ユーザからの支援を誘発させることで既存の関数 G が正しく動作するように状況を改善させるシステム構成法を提案する．これは，前提条件が満たされている度合いが判断できるようにシステムがユーザにフィードバックを行い，ユーザに以下の支援行動を促す仕組みである．

- 直接型支援：外的前提を充足するように，シーンを改善する行動
- 間接型支援：内的前提を充足するように，計算機内部の認識状態 $\hat{\boldsymbol{s}}(t)$ と実状態 $\boldsymbol{s}(t)$ とを合致させる行動

このような支援行動をユーザに負担をかけることなく自然な形で引き出す事ができれば，最新の画像認識手法と相補的に組み合わせることで，誤認識や認識精度低下の問題を回避できる実用的なジェスチャ認識システムが実現できると考えられる．

以下，このようなシステムの構成方法の詳細を議論するため

に，ユーザの支援行動を誘発するためのフィードバック処理を関数 $F(\hat{\boldsymbol{s}}(t), \boldsymbol{I}(t), \boldsymbol{w}(t))$ と定め，F が満たすべき要件を述べる．そして，フィードバックを起点とした良好な認識状態への復帰シナリオについて詳述することで，我々が提案するシステム構成法をまとめる．

2.1 フィードバック関数

上述のように，フィードバック関数 F は， 外的前提と内的前提が満たされている度合いを適宜ユーザに提示することで，ユーザから自然な形で直接型支援や間接型支援を引き出すように設計されなければならない．

ここで F がユーザに提示すべき情報は，

- 主に $\hat{\boldsymbol{s}}(t)$ から生成するジェスチャの認識結果
- 主に $\boldsymbol{w}(t)$ から生成する認識処理の中間段階

の二つに大別できる．認識結果に加え，処理の中間段階を見せることによって，ユーザが誤認識の原因が何かを推定することが可能となり，良好な認識状態へ復帰させるための適切な支援行動を選択できる．また，中間段階を提示することで，誤認識が起こりそうな動作等をユーザが理解することもでき，以後そのような動作を避けさせる効果も期待できる．

またユーザの負担低減という観点では，F は過度の情報提示を避けなければならない．具体的には，状況に応じた情報提示レベルの自動制御や，支援行動候補の自動提示等が考えられる．前者は，良好な認識状態の時は，認識結果のみを提示し，認識状態が悪化，つまり認識状態とシーンとのずれに応じて，他の情報も徐々に提示されていく仕組みである．後者はシステムだけでシーンと認識状態の齟齬の有無やその原因を可能な限り検出し，良好な認識状態への復帰法を提案する仕組みである．

ただしここで忘れてはならないのは，前述のように計算機だけで前提条件の充足判定を行うことは非常に困難な点である．この点を踏まえつつ，F は最終的な誤認識の原因の推定，支援行動の決定のみをユーザに促すよう設計されなければならない．

2.2 良好な認識状態への復帰シナリオ

フィードバックを起点とした良好な認識状態への復帰シナリオは大きく分けて，直接型支援による復帰と，間接型支援による復帰の二つに大別できる．前者はシステムからのフィードバックを介してユーザが外的前提が充足されていない状況を把握した場合のシナリオである．この場合，ユーザは実世界側のユーザの手の位置や背景の配置を変更する等，シーンを直接操作する行動を行うことになる．一方後者の間接型支援では，ユーザは内的前提を充足するように計算機内部のデータ構造を操作する必要がある．そのため間接型支援専用のユーザインタフェースを新たに追加しなければならない．本研究ではこの追加インタフェースもジェスチャインタフェースで実現するものとし，**改善インタフェース**と呼ぶ．

ここで留意すべき点は，改善インタフェース自身にもジェスチャ誤認識の問題が存在する点である．改善インタフェース自身が誤認識，誤動作をしてしまうと提案手法は破綻する．そこで改善インタフェースは例えば以下のように設計する．まず改善インタフェースが用いるジェスチャ認識処理は，通常の画像認識関数 G よりもより多くの前提条件を用いるものとする．その結果，改善インタフェースはより制約の強い限定されたシーンでしか動作しなくなるが，そのような限定された状況下で間接型支援を確実に遂行できるように間接型支援行動のインタラクションを設計する．

3. ジェスチャ認識の具体例

本章では，画像認識処理の実例として，ジェスチャインタフェースを想定した手領域の追跡処理を取り上げる．そして提案する枠組みに従って設計した簡単な追跡処理とフィードバック，復帰シナリオについて詳述する．

まずシーンや画像 $\boldsymbol{I}(t)$ に対応する状態モデルとして，手領域の色を表す肌色モデルを用意し、

- 手の位置、速度
- 肌色モデルのパラメータ

からなるベクトルを $\boldsymbol{s}(t)$ と定義する。同時に $\boldsymbol{s}(t)$ に関する状態方程式としては，手領域が画像上では等速直線運動し，肌色モデルのパラメータは変化しないものとして，その式を定義した．この状態モデルで表現できるシーンとして下記のものを考える．

- ユーザは 1 人である
- 背景に肌色物体は存在しない
- ユーザの左右の手は近接していない
- ユーザの両手部はカメラの視野内にある
- ユーザの服は肌色らしくない
- ユーザは顔，両手部以外の肌を露出しない
- ユーザの顔は常にカメラに正対している

これらのシーンでは，手領域の追跡アルゴリズム G は次のような簡便な方針で実現できる．まず上記の想定より $\boldsymbol{I}(t)$ 中の肌色領域は左右の手領域または顔領域のいずれかになる．そこで追跡アルゴリズムでは肌色モデルを用いて撮影画像 $\boldsymbol{I}(t)$ の各画素が肌色であるか否かを判別し，画像から手と顔に対応する領域を絞り込めるようにする．さらに，この想定の元では顔領域はカメラに正対しているため既存の顔認識手法 [2] を使うことで肌色領域から顔領域を頑健に除外できる．

具体的アルゴリズムは次の通りである。まず入力画像 $\boldsymbol{I}(t)$ に対して，既存の顔検出プログラムにより顔領域を除去する．次に肌色モデルを用いた肌色領域の追跡処理を行う．これはパーティクルフィルタ [1] を用いた．つまり $\boldsymbol{w}(t)$ の実体はパーティクル群を含むことになる．次に肌色領域から左右の手の識別はパーティクルフィルタによる追跡結果をクラスタリングすることで行う．ここでは，検出される手領域は二つであると仮定しているため，2 クラスへ分離する．同一クラスのパーティクルの重心を左右の手位置としており，これらが認識結果に当たる．

次に，この手領域追跡システムにおけるユーザへのフィードバックを起点とした良好な認識状態への復帰シナリオについてその詳細を述べる．まず，システムがユーザに提示するフィードバックの内容を説明する． 2.1 節では，フィードバックの内容は大別すると，ジェスチャの認識結果，認識処理の中間段階，の二つであると述べた．ここでは，認識結果，及び，認識処理の中間段階として以下の情報を入力画像に重畳し，フィードバッ

表 1　誤認識の原因の具体例

誤認識の原因	原因の推定	支援行動
C-1. 背景に肌色物体がある	手ではなく，背景物体にパーティクルが乗っており，肌色モデルが適切な時，その背景物体が誤認識の原因だと推定できる．	背景からその物体をどかす，カメラに映らないようにする等，シーンを改善する直接型支援を行う．
C-2. 手の動作が速い	手の速さが増すにつれ，その手に乗っているパーティクルが徐々に消滅する様子から，手が速いことが誤認識の原因だと推定できる．	手の速さをパーティクルが十分追跡可能な速さに調整する，つまりシーンを改善する直接型支援を行う．
C-3. 肌色モデルが不適切	明らかに肌色ではない物体が肌色であると判定されていたり，手が肌色でないと判定されている時，肌色モデルが不適切なことが誤認識の原因だと推定できる．	肌色モデル再設定用スイッチで適切な肌色モデルを再設定する，つまり認識状態を変更する間接型支援を行う．

クとして図 2 の様に提示する．

F-1.　認識された左右の手の位置

F-2.　入力画像で肌色と判定された領域

F-3.　生存しているパーティクル群

F-4.　前時刻のパーティクル群

F-1 は $\hat{\boldsymbol{s}}(t)$ から生成した認識結果，F-2，F-3，F-4 は $\boldsymbol{w}(t)$ から生成した認識処理の途中段階，をフィードバックしている．図 2 の Display1 に表示されている手のアイコンが F-1 による認識結果の提示である．同様に F-2 は Display2 で緑色に塗られている領域，F-3，F-4 は Display1 に表示されている粒子であり，所属するクラスや時刻の違いにより色が異なる．

ここで，この手領域追跡システムにおいて，シーンと認識状態 $\hat{\boldsymbol{s}}(t)$ の間に齟齬が生じる原因，つまり誤認識の原因の具体例をいくつか表 1 に挙げる．誤認識の原因は，外的前提を充足していない，内的前提を充足していない，の二つに大別できると述べたが，それらはさらに数種類に分類されることに注意されたい．

表 1 の例を用いて，誤認識が発生した時に，ユーザがフィードバックを受けてから原因を推定し，支援行動を行うまでの復帰シナリオを説明する．まず，表 1 における誤認識の原因が「C-1. 背景に肌色物体がある」の場合を説明する．原因の推定では，F-2 により肌色モデルが適切であることがわかり，F-3 により手ではなく，背景物体にパーティクルが乗っていることがわかる．また，支援行動は，直接型支援による復帰，に当たる．ここでの直接型支援ではシーン内の背景を改善している．

次に，表 1 における誤認識の原因が「C-2. 手の動作が速い」の場合を説明する．原因の推定では，F-3，F-4 により手の速さが増すにつれ，その手に乗っているパーティクルが徐々に消滅する様子が観測できる．また，支援行動は，直接型支援による復帰，に当たる．ここでの直接型支援ではシーン内のユーザの手の速さを改善している．

最後に，表 1 における誤認識の原因が「C-3. 肌色モデルが不適切」の場合を説明する．原因の推定では，F-2 により明らかに肌色ではない物体が肌色であると判定されていたり，手が肌色でないと判定されていることがわかる．また，支援行動は，間接型支援による復帰，に当たる．間接型支援では，計算機内部の情報を操作するため，改善インタフェースが必要であると前述したが，この手領域追跡システムでは，改善インタフェースとして，以下の二つをユーザに提供する．

図 2　手領域追跡システムのフィードバック

図 3　原因の推定 (原因：背景に肌色物体がある)

I-1.　肌色モデル再設定用スイッチ

I-2.　左右手位置のリセットスイッチ

ここでは，I-1 を用いた間接型支援を行っている． I-1 は，実状態 $\boldsymbol{s}(t)$ と認識状態 $\hat{\boldsymbol{s}}(t)$ の肌色モデルのパラメータにずれが生じた場合，$\hat{\boldsymbol{s}}(t)$ 側の肌色モデルのパラメータを再設定するためのスイッチであり，I-2 は，実状態 $\boldsymbol{s}(t)$ と認識状態 $\hat{\boldsymbol{s}}(t)$ の間で，ユーザの手の位置や左右に不一致が生じた場合，正しい認識へとリセットするスイッチである．

また，表 1 における誤認識の原因が「C-1. 背景に肌色物体がある」における原因の推定から良好な認識状態への復帰までのシナリオを，図 3 から図 6 に示す．図 4 では，肌色物体をカメラの視野外に移動するという直接型支援を行った後も誤認識が起きている．これは，計算機内部の認識状態 $\hat{\boldsymbol{s}}(t-1)$ は実状態 $\boldsymbol{s}(t-1)$ と合致してなければならない，という内的前提を充足していないからであり，I-2. 左右手位置のリセットスイッチを用いることで，良好な認識状態へ復帰させようと考える．図 5 に示すように，バケツアイコンの上に両手を持っていくことでリセットスイッチが起動し，パーティクルが両手に再設置される．その結果，図 6 に示すように良好な認識状態へ復帰できる．

図 4　肌色物体をカメラ視野外に移動した後

図 5　左右手位置のリセットスイッチを使用

図 6　良好な認識状態への復帰

4.　実験・考察

4.1　実験内容

提案するジェスチャ認識の枠組みの有効性を検証するために，3 章で述べた手領域追跡システム用い，実験を行った．フィードバックの質の良し悪し，改善インタフェースの有無を比較することで，ユーザからの支援行動により認識状況がどの程度改善できるか，またその支援行動がどの程度ユーザの負担となるか，を測定した．実験では提示されるフィードバックの意味や改善インタフェースの使用法について十分に理解している被験者 3 名で行った．実験では以下の二つの手領域追跡システムを用いる．共に手領域追跡部分は，3 章の冒頭で述べたように設計してあり，提示されるフィードバックの質，改善インタフェースの有無の部分で実装が異なる．

- システム 1 ：ユーザへのフィードバックは，手として認識されている位置 (各クラスのパーティクルの重心の位置) のみを入力画像に重畳した画像だけである．つまり，左手と右手の判別結果はユーザには分からない状態である．また，改善インタフェースは実装されていない．
- システム 2 ：ユーザへのフィードバック，改善インタフェース共に 3 章で述べた通りのものが実装されている．

この二つのシステムを用いて，以下の実験を行う．まず実験は，シーンが前提条件を充足しており，手領域の追跡，左手と右手の判別が正しく行われている想定内の状況にある時点から開始する．次に，誤認識の原因を与えることで前提条件を充足していない状況を再現する．これによりシーンと認識状態の間に齟齬が発生した時，被験者が良好な認識状態に復帰させられるかを評価する．なおシステム 1 では，左右の判別結果は提示されないので，誤認識が生じたかどうか気づけない場合がある．よって，被験者以外の人間が別ディスプレイでシステムの認識状態を確認し，被験者に誤認識の発生を知らせることとする．以下に本実験で与えた誤認識の原因を示す．

（1）　カメラ視野内に他の人物が写りこむ (背景に肌色物体がある)

（2）　片手を伸ばしてカメラ視野外に出す (手がフレームアウトした)

（3）　片手をすばやく動かす (手の動作が速い)

（4）　作業場の照明を暗くする (肌色モデルが不適切)

（5）　手を交差させたり，左右の手を近づける (左右の手が近接した)

それぞれの誤認識の原因において，以下の 2 項目を調べ，システム 1 とシステム 2 での結果を比較し，考察する．

(A) 復帰できたかどうか

(B) 復帰できた場合，失敗発生から復帰までの時間

また，30 秒以内に復帰できない場合は，復帰できなかったとする．

4.2　実験結果と考察

図 7 はシステム 1 とシステム 2 における復帰回数を示す．このようにシステム 2 の方が復帰できる回数が多いことから，フィードバックの提示と改善インタフェースを与えることにより，誤認識が生じてしまった時に，良好な認識状態に復帰しやすいことを確認できた．

図 8 は復帰に要する時間の頻度分布を示す．このように現状では復帰までに最大 30 秒程度の時間がかかる．本研究では，誤認識が生じた時，ユーザの支援行動を誘発することにより，良好な認識状態への復帰を可能にしようというものだが，その支援行動はユーザが煩わしいと感じずに，負担なく行えるものである必要がある．図 8 からわかるように，現時点ではユーザに負担をかけることなく，自然な形でユーザの支援行動を誘発するインタラクションの設計ができているとは言い難い．システム 2 において，復帰までに時間がかかる原因は，いくつか考えられる．まず，今回用意したフィードバックの質が不十分であること．今回のフィードバックでは，誤認識の原因が何であるのかの判断が難しいという声が見られ，フィードバックの内容やその見せ方をさらに考え込む必要性を感じた．また，今回用意した改善インタフェースの設計も改善の余地が大いにある．例えば，左右手位置のリセットスイッチにおいて，その操作法が被験者にとって意外と難しく，スイッチ起動までに時間がかかるという意見があった．これらの問題を考慮しつつ，今後はユーザに負担をかけることなく，自然な形でユーザの支援行動を誘発するインタラクションの設計を目指す必要がある．

次に，誤認識の原因ごとに検証すると，カメラ視野内に他の人物が写りこむ，作業場の照明を暗くする，の二つでは，システム 1 で良識な認識状態に復帰できた被験者が 1 人もいないという結果が得られた．あらゆる環境下に対応し，誤認識を起こさないシステムを設計することは，ジェスチャ認識においては困難であることは既に述べており，この結果からも，実利用を志向す

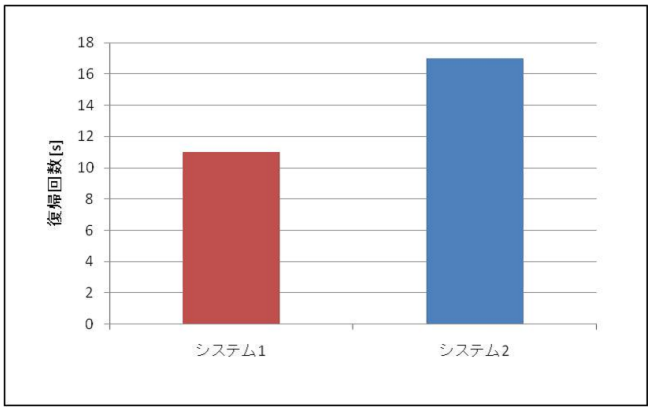

図 7　実験結果 1

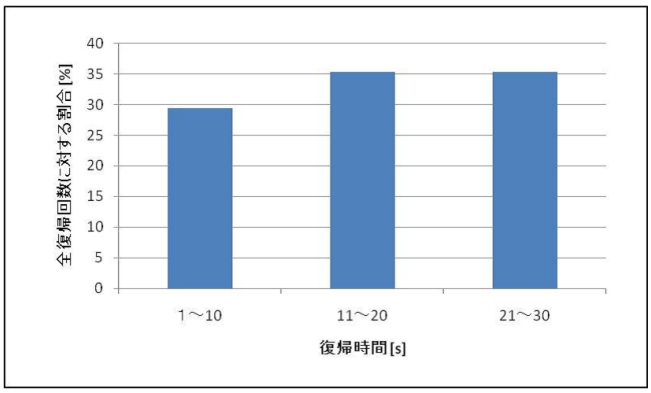

図 8　実験結果 2

るジェスチャインタフェースの実現には, ユーザの多少の支援, それに役立つ適切なフィードバックの提示と改善インタフェースの必要性が示唆される.

5. おわりに

本稿では，画像ベースのジェスチャ認識方式の現実的なシステム構築方針として，環境が前提条件を満足するように，ユーザからの支援を誘発しユーザに自然な形で環境を維持させる仕組みを提案し，その構成論について議論した．そしてその構成論に基づいた画像認識処理の実例として手領域追跡処理を実装し，本枠組みは概ね良好に機能することを示した.

今後は，認識状態のフィードバック内容の質を高め，ユーザの負担を抑えつつ適切な支援を引き出すジェスチャインタフェースを用いた実アプリケーションを実現していく予定である.

謝　　辞

本研究の一部は，独立行政法人科学技術振興機構（JST）戦略的創造研究推進事業 (CREST) 「マルチモーダルな場の認識に基づくセミナー・会議の多層的支援環境」の助成を受けて行った.

文　　献

[1] M. Isard and A. Blake: “Condensation—conditional density propagation forvisual tracking”, Int. J. Comput. Vision, **29**, 1, pp. 5–28 (1998).

[2] P. Viola and M. J. Jones: “Robust real-time face detection”, International Journal of Computer Vision, **57**, pp. 137–154 (2004). 10.1023/B:VISI.0000013087.49260.fb.

[3] 中野, 近藤, 小泉, 中村：“ジェスチャーインターフェースのためのインタラクション設計”, 信学技報, 第 110 巻, 電子情報通信学会, pp. 27–28 (2010).

[4] 近藤, 西谷, 中村：“協調的物体認識のためのインタラクション設計”, 第 13 回画像の認識・理解シンポジウム論文集 (MIRU2010) (2010).