

ウェアラブル物探し支援システム “I’ m H e r e !” の試作

上岡 隆宏[†] 河村 竜幸[†] 河野 恭之[†] 木戸出 正継[†]奈良先端科学技術大学院大学 情報科学研究科[†]

1 はじめに

本研究では、日常生活を背景とした物探しタスクを支援するウェアラブルシステムを提案する。日常生活における様々なタスクを遂行する過程において、我々は様々な身の回りの物体（オブジェクト）を利用する。オブジェクトは、使用されていないときには身の回りの生活空間内に置かれており、必要に応じてユーザーの手によって把持され、タスク遂行のために用いられる。オブジェクトの生活空間内の移動について考えるとき、オブジェクト側の立場に立ってみると、彼らは自ら動くことができないので、ユーザーに把持されることではじめて移動する機会に恵まれ、ユーザーが意識する・しないに関わらず、ユーザーが把持を解く位置に置かれる。

日常生活において、我々はタスク遂行のために必要なオブジェクトを探索する「物探しタスク」を遂行しなくてはならない場合がある。たとえば、あるとき書き物をするために筆箱を手取る必要が出てきたが、筆箱をどこかに置き忘れてしまった場合などである。自分が筆箱を置いたときの記憶を呼び起こせない限り、筆箱を見つけ出すためには、自分が置いた可能性のある場所について探して回る必要があり、手間と時間にかかるコストが大きい。

本研究で提案する物探し支援システム “I’ m H e r e !” では、オブジェクトが把持されたときのユーザー視点映像を頭部装着型カメラデバイスによって撮影し、ウェアラブルPCに記録する。同時に、あらかじめシステムに登録してあるオブジェクト群とのマッチングを行い、把持されたオブジェクトの登録名によって記録映像にインデックスを作成する。対象オブジェクトの置かれた場所についてユーザーから問い合わせがあったとき、記録映像のインデックスから対象オブジェクトの名前を検索することで、ユーザーが求めているオブジェクトが最後にユーザーによって把持されているときの映像を呼び出すことができる。システムを利用するユーザーにとっては、登録オブジェクト自身が把持されたときにシステムに自分の位置を伝え、検索時には「ここにいるよ」とシステムを通して訴えているかのように思われる[図1]。

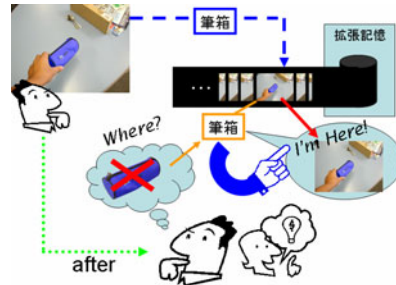


図1：“I’ m H e r e !”のコンセプト

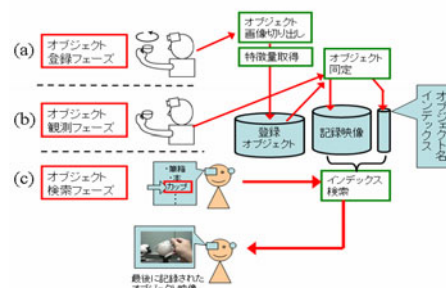


図2：“I’ m H e r e !”の機能

2 “I’ m H e r e !”

“I’ m H e r e !” は、日常生活を背景として、あらかじめ登録してあるオブジェクトの中からユーザーが探したいオブジェクトを選択してシステムに入力することで、対象のオブジェクトが最後に観測されたときのユーザー視点からの記録映像をユーザーに提示し、それによって物探しタスクを支援するシステムである。

ユーザーは頭部に撮像デバイスとHMDを装着する。撮像デバイスは、カラー・赤外反射画像が同時に光軸一致で取得できる複合撮像デバイス[1]が用いられる。この複合撮像デバイスによって、ユーザー視点の映像がウェアラブルPCに常時記録されるとともに、ユーザーが把持したオブジェクトを認識するためのオブジェクト画像が取得される。HMDには、システムからユーザーに提示される情報が表示される。オブジェクト検索時には、登録されているオブジェクトの一覧がHMDに表示され、ユーザーはその一覧を操作して検索対象のオブジェクトを指定する。また、HMDには、検索結果のオブジェクト映像も表示される。ユーザーは表示されるオブジェクト映像を見ることによって、そのオブジェクトがある場所についての情報を得ることができ、物探しタスクを効率よく実行することができる。

ユーザーは図2に示した3つのフェーズを通して “I’ m H e r e !” システムを利用する。ユーザーが新しいオブジェクトをシステムに登録したいとき、システムはオブジェクト登録フェーズに移行する。ユーザーからの指

“I’m Here!” – A Wearable Object Remembering Support System

[†]Takahiro Ueoka, [†]Tatsuyuki Kawamura, [†]Yasuyuki Kono, and [†]Masatsugu Kidode

[†]Graduate School of Information Science, Nara Institute of Science and Technology

示がない場合には、システムは登録オブジェクト情報をもとにして、ユーザ視点映像を分析してインデックス付けを行うオブジェクト観測フェーズを実行する。ユーザが登録したオブジェクトを探したいとき、システムはオブジェクト検索フェーズに移行する。

“I’m Here!”をオンラインで動作させ、物探しタスクの支援を行わせるために必要な技術要素としては、次のものがある。

- ・ ユーザ視点映像の常時記録
- ・ 映像中のオブジェクト画像のオンライン認識
- ・ 記録映像へのオブジェクト名インデックス付加
- ・ HMDへのオブジェクト一覧表示とユーザからの入力インタフェース実装
- ・ HMDへの検索結果映像提示

現在は、登録済みのオブジェクト特徴量を用いて入力画像中のオブジェクトを認識するオフライン処理が実装されている。

2.1 装置

“I’m Here!”は、頭部装着型カメラデバイス、HMD、ウェアラブルPC、そしてユーザからの命令の入力デバイスで構成されている。頭部装着型カメラデバイスとして用いられる複合撮像デバイス[1]は、カラー・赤外反射画像を光軸一致状態で取得することができる。日常生活における把持オブジェクトを撮影したユーザ視点映像は、背景が多様かつ複雑であることが多く、オブジェクトだけの情報を取り出すことは難しい。本システムで用いるオブジェクト認識では、複合撮像デバイスが取得する赤外反射画像が、照射した赤外光の環境からの反射光を観測したものであり、その輝度値は一般的に距離の自乗に反比例することを利用した。赤外反射画像に対して閾値処理を行い、背景と近接領域を分割し、さらにカラー画像中の肌色領域を手領域として除去することで、オブジェクトだけの画像を取得した。

2.2 システムの機能

図2に示したように、本システムはユーザに対して3つの機能を提供する。オブジェクト登録フェーズとオブジェクト検索フェーズはユーザが望むときに起動し、オブジェクト観測フェーズはユーザがシステムに対して何も要求していないときに動作する。

ユーザが新しいオブジェクトを登録したいとき、システムはオブジェクト登録フェーズに移行する。日常生活においてユーザに把持されるオブジェクトは、その姿勢によって見た目が変化する。その変化に対応するため、“I’m Here!”のオブジェクト登録フェーズでは、ユーザは登録したいオブジェクトを把持し、目の前で様々な姿勢をオブジェクトにとらせ、それを眺める動作をする[図2(a)]。このとき、ユーザの頭部に装着された複合撮像デバイスによって撮影された時系列のオブジェクト画像を基にして、図3のようにオブジェクトの様々な姿勢を代表する複数枚の画像を選出し、それぞれの特徴量を抽出する。入力画像の特徴量をオブジェクトのそれぞれの代表画像特徴量とマッチングしていくこと

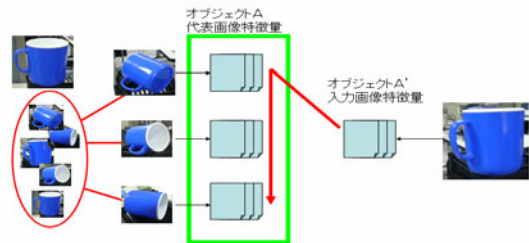


図3: 様々なオブジェクト姿勢に対応可能な画像マッチング手法

で、様々な姿勢に対応可能な画像マッチングが実現する。

ユーザからの問い合わせがない状態のとき、本システムはユーザ視点画像を常時記録しながら画像中の把持オブジェクトを観測し、登録オブジェクト特徴量を用いて認識する。認識結果は、ユーザ視点映像に付随してオブジェクト名のインデックスとして保存される[図2(b)]。

ユーザが物探しタスクを行うとき、システムに登録されているオブジェクト一覧から探索対象オブジェクトを選択する[図2(c)]。対象オブジェクト名に基づいて記録映像に付随するインデックスを検索し、最後に記録された対象オブジェクトの画像を基点としてその前後の映像をユーザに提示する。ユーザは、提示された映像を見ることでそのオブジェクトを探すための手がかりを得ることができる。

3 オブジェクト認識

ユーザが把持するオブジェクトは、ユーザ視点画像において様々な姿勢で観測される。似通った姿勢画像についてそれぞれの特徴量を比較する必要があるため、ひとつのオブジェクトについての特徴量が多くなる。比較する特徴次元数を削減し、ユーザ視点画像の視線軸回転方向の変化にロバストな認識をするために、本システムではオブジェクト画像のヒストグラム特徴量を用いてオブジェクト認識を行う。すなわち、入力画像のヒストグラムと最も類似したヒストグラム特徴量が属するオブジェクトを認識結果とする。

本節では、物探し支援システムに適したオブジェクト認識手法について説明する。入力画像から抽出したオブジェクト画像から、カラー・赤外反射輝度からなる画素値に基づくヒストグラム特徴量を抽出する。ヒストグラム特徴量は、ユーザ視点映像の視線軸回転方向の変化にロバストな特徴量である。提案するヒストグラム特徴量は、カラー情報のみを用いるヒストグラムとは異なり、緩やかな立体形状情報を持つ赤外反射輝度値を用いることで、画面奥行き方向の形状が異なるオブジェクトを弁別することができる。また、オブジェクト重心位置からの各画素の距離を考慮することで、ユーザ視点画像の視線軸回転方向のロバスト性を保ったまま、平面上のゆるやかな形状情報の特徴量に含ませる。

本研究では、ヒストグラム特徴量の軸について前述の色情報・赤外反射輝度値・オブジェクト画像重心からの距離の3つに着目し、それぞれを軸とする多次元ヒストグラムについて、オブジェクト認識率の観点から、認識手法への適用可能性を評価した。

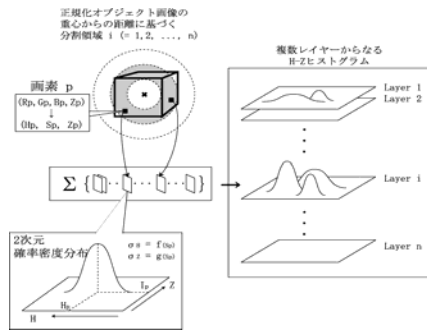


図 4 : オブジェクト画像からのヒストグラム抽出

3.1 オブジェクト特徴量の抽出

ひとつのオブジェクト画像からヒストグラム特徴量を抽出する手法について述べる. ヒストグラム特徴量の欠点として, 形状情報が欠落するために, 類似した色を持ちながら形状の異なるオブジェクト同士を誤認識する可能性がある. そこで, 提案するオブジェクト認識手法では, オブジェクト画像の画素ごとにカラー・赤外反射輝度値によるヒストグラム特徴量を抽出する. 画面奥行き方向の緩やかな形状情報を持つ赤外反射輝度値と, 画像平面上の緩やかな形状情報を持つオブジェクト画像重心からの距離を用いて, 図 4 のようなヒストグラム配列を作成する. ヒストグラム特徴量の抽出時において, あらかじめオブジェクト画像を規定画素数へ圧縮し, 正規化しておく. これにより, すべてのオブジェクト画像について画素数が同じになり, ヒストグラム差分値によってオブジェクト画像間の類似度を求めることができる.

オブジェクト画像 x 内の画素 j が, x の重心位置から同心円状に分割した領域 i に属するとき, $\{H-Z\}$ 平面上の 2 次元ヒストグラム h_{xi} を次のように定義する.

$$h_{xi}(H, Z) = \left\{ \sum_j N(H - H_j, \sigma(s_j)) \times N(Z - Z_j, \sigma_{z0}) \mid j \in \text{Layer}_i \right\} \quad (1 \leq i \leq n)$$

ここで $N(\mu - \mu_0, \sigma)$ は, 平均値 μ_0 , 分散 σ の正規分布関数である. (H_j, S_j, Z_j) は画素 j のカラー+赤外反射輝度値 (R_j, G_j, B_j, Z_j) のカラー成分を HSV 表色系[2]に変換し, 色相 H と彩度 S , 赤外反射輝度 Z で再構成した画素値である. $\sigma_H(S_j)$ は図 5 右に示すような指数関数である. 画素 j が無彩色の場合, 彩度 S_j が 0 となり, H_j が不定値をとる. 単純に H_j を足し合わせた場合, ヒストグラム h_{xi} に雑音要素として含まれてしまう. σ_H を S_j の指数関数として定義することで, 無彩色に近い画素によるヒストグラムへの雑音要素を抑える[図 5].

同心円状の分割領域それぞれについて h_i を求めることで, オブジェクト画像中心からの距離 C を 3 目目の軸として, $\{H-Z-C\}$ 空間の 3 次元ヒストグラム $H(x)$ を次のように定義する.

$$H(x) = \{h_1, h_2, \dots, h_n\} \quad (2)$$

ヒストグラム $H(x)$ を用いて, 画像 x と画像 y の類似性を表すヒストグラム差分値 $F(x, y)$ を次のように定義する.

$$F(x, y) = H(x) - H(y) \quad (3)$$

$$= \sum_{i=1}^n |h_{xi} - h_{yi}|$$

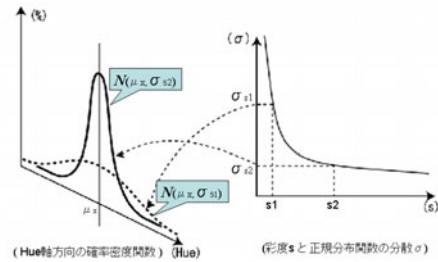


図 5:H 軸方向の確率密度関数

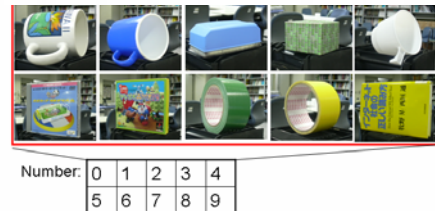


図 6 : 実験に用いるオブジェクト群

3.2 オブジェクト認識手法の検証実験

提案するオブジェクト認識手法に対して, その性能を評価するためにオフライン実験を行う. オブジェクト画像内のある画素について, 色相値 H , 赤外反射輝度値 Z , オブジェクト重心からの距離 C としたとき, $\{H\}$, $\{H-C\}$, $\{H-Z\}$, $\{H-Z-C\}$ といった組み合わせの次元 (すなわち 1 ~ 3 次元) をもつヒストグラムを用いて, 式 (3) で導入したヒストグラム差分値を求め, それぞれから導く認識結果を比較して, $\{H-Z-C\}$ の 3 次元ヒストグラムを用いる提案手法の性能を検証する.

入力画像の取得は, 日常生活環境の 1 つである屋内の蛍光灯照明下で行った. 認識対象のオブジェクトは回転台に設置し, 1 軸の自由度を与えて見た目の姿勢を変化させた. フリッカーによるノイズを抑えるために, 同じ姿勢について連続 20 枚撮影した画像を平均し, 1 枚の入力画像とした.

図 6 に示す 10 個のオブジェクトは, 形状が類似した組み合わせ・色が類似した組み合わせをとることができる内容になっている. 10 個のオブジェクトに対して, 1 軸自由度で 4 つの姿勢をそれぞれ定義し, 全 40 組の姿勢グループごとに 3 つの入力画像を取得した.

オブジェクト認識手法の単純な認識率を, 次のようにして算出する. オブジェクト o , 姿勢グループ i に属する画像 x と o 以外のオブジェクトに属する画像との最小差分値を d_{xomin} , i 以外の姿勢グループに属する画像との最小差分値を d_{ximin} , 姿勢グループ i 内の画像同士による最大差分値を d_{imax} とすると, $d_{xomin} > d_{imax}$ となった場合にはオブジェクト単位の認識が成功, $d_{ximin} > d_{imax}$ となった場合には姿勢グループ単位の認識が成功するとみなす. 全ての画像に対してオブジェクト単位の認識判定と姿勢グループ単位の認識判定を試み, 認識が成功した画像数の全体画像数に対する割合を認識率とする. 求めた認識率を表 1 に示す.

姿勢グループ i について, i に含まれる画像と全姿勢グループに含まれる画像との差分値集合 P_i を次のように定義する.

$$P_i = \{F(x, y) | x \in G_i, y \in G_{all}\} \quad (4)$$

G_i : 姿勢グループ i の画像集合

G_{all} : 全姿勢グループの画像集合

$F(x, y)$: 画像 x と画像 y のヒストグラム差分値

P_i について、全平均差分値 p_{ia} , 分散 σ_i とする。姿勢グループ i に含まれる画像と姿勢グループ j に含まれる画像との最大差分値 p_{ijmax} が、差分値集合 P_i に対してどの程度分散して存在するかを表すための指標として、最大差分値分散比率 α_{ijmax} を次のように定義する。

$$\alpha_{ijmax} = \frac{p_{ia} - p_{ijmax}}{\sigma_i} \quad (5)$$

同様に、姿勢グループ i と姿勢グループ j との最小差分値を p_{ijmin} とし、最小差分値分散比率 α_{ijmin} を次のように定義する。

$$\alpha_{ijmin} = \frac{p_{ia} - p_{ijmin}}{\sigma_i} \quad (6)$$

オブジェクトの形状や色が類似した姿勢グループの組を $\{a, b\}$ ($a \neq b$) とするとき、認識手法のロバスト性を示す評価値 ϵ_{ab} を次のように定義する。

$$\epsilon_{ab} = |\alpha_{aa\max} - \alpha_{ab\min}| \quad (7)$$

ϵ_{ab} は、姿勢グループ a についての差分値分布における姿勢グループ a と b との分離状態を表す。類似形状オブジェクトの組として図 6 にあるオブジェクト $\{0, 1\}$ を、類似色オブジェクトの組としてオブジェクト $\{3, 7\}$ を挙げ、各オブジェクトに属する姿勢グループの中から類似条件にあう姿勢グループ $\{a, b\}$ を 8 組抽出し、それぞれについて ϵ_{ab} を求める。各次元のヒストグラムごとに求めた ϵ_{ab} を表 2 に示す。

表 1 を見ると、オブジェクト単位の認識においては、H だけの 1 次元ヒストグラムを用いた差分値で十分であるかのように思える。また、表 2 左には、同じ形状で異なる色情報を持つオブジェクト同士の識別に対しては、H を用いた 1 次元ヒストグラムと他の多次元ヒストグラムが同程度の性能を持つことが示されている。しかし、表 2 右の結果から、H だけを用いた 1 次元ヒストグラムの差分による認識手法では、あるオブジェクトが自分自身を認識するときの差分値が、他のオブジェクトとの差分値ときわめて近い分布であることがわかり、多くのオブジェクトを登録する可能性がある日常生活でのオブジェクト認識に用いるには危うい手法であると言える。表 1 において $\{H-C\}$ 2 次元ヒストグラムの列が示すように、オブジェクトの緩やかな形状情報をもつ重心からの距離 C をヒストグラムの軸として導入することで、姿勢グループ単位でのオブジェクト認識率が向上する。また、表 2 右の $\{H-Z\}$ 2 次元ヒストグラムの列が示すように、画面奥行き方向の緩やかな形状情報をもつ赤外反射輝度値 Z をヒストグラムの軸として導入することで、類似した色をもつ異なる姿勢グループに対する差分値分布を、同じ姿勢グループ内での差分値の分布からはっきりと弁別しやすくなる。表 1 では、それらの提案手法の特徴によって、 $\{H-Z-C\}$ ヒストグラムによる認識手法が姿勢グループ単位の認識率を最も高く記録している。以上のことから、 $\{H-Z-C\}$ の 3 次元ヒストグラムによる提案オブジェクト認識手法は、色だけの類似性や形状だけの類似性によって影響を受けず、ロバストにオブジェクト認識を行う能力を持つことが示された。

表 1: 各次元ヒストグラムを用いたオブジェクト認識率

	ヒストグラムの次元			
	H	H-C	H-Z	H-Z-C
オブジェクト単位の認識率	99.2	99.2	91.7	96.7
姿勢グループ単位の認識率	32.5	77.5	70.0	78.3

表 2: 類似形状・類似色オブジェクト認識における各次元ヒストグラムの性能評価

類似形状オブジェクト[0,1]に対する ϵ_{ab}				類似色オブジェクト[3,7]に対する ϵ_{ab}			
姿勢グループの組	ヒストグラムの次元			姿勢グループの組	ヒストグラムの次元		
	H	H-C	H-Z-C		H	H-C	H-Z
A	3.546	3.682	3.245	2.572			
B	1.118	1.125	0.956	1.386	0.058	0.255	0.559
C	1.608	1.725	2.312	2.696	0.018	1.101	1.606
D	3.441	3.701	2.578	3.673	0.027	0.394	1.207
A'	2.461	2.047	2.913	2.461	0.034	1.438	0.926
B'	1.269	1.171	1.272	1.454	0.050	0.125	0.941
C'	3.485	3.588	3.119	3.575	0.019	1.236	2.282
D'	3.676	3.818	4.124	3.425	0.025	0.323	1.726
平均 ϵ_{ab}	2.543	2.607	2.565	2.655	0.032	0.772	1.242

4 おわりに

日常生活における物探し支援のためのウェアラブルシステム “I’ m Here!” を提案した。ユーザによって把持されるオブジェクトを頭部装着型カメラで撮影し、画像情報に基づいて認識するための手法として、カラー・赤外反射画像から 3 次元ヒストグラム抽出処理を行い、オブジェクト特徴量を構築した。実際の日常生活で用いられるオブジェクトを用いて、提案するオブジェクト認識手法についてオフラインの評価実験を行った。結果として、提案するオブジェクト認識手法は、色特徴と形状特徴に基づいて日常生活でのユーザ把持オブジェクトを弁別することができる能力をもつことが示された。

3 章で述べた 3 次元ヒストグラムによるオブジェクト認識は、単純なヒストグラムに比べて次元数が大きく処理時間コストが高い。処理速度と認識率のトレードオフを見極めながら、“I’ m Here!” の実装を進める予定である。同時に、オブジェクト認識処理の高速化に向けて、特徴量圧縮のための代表オブジェクト画像の選定、データ構造改良によるヒストグラム差分計算の高速化、そして粗同定段階導入によるオブジェクト認識時の登録オブジェクト候補数削減を行っていく。また、“I’ m Here!” のオンライン化に向けて、オブジェクト登録・検索インタフェースの実装をすすめ、日常生活環境内での物探しタスクに対する提案システムを用いた支援の有効性を検証する。

謝辞

本研究の一部は、科学技術振興事業団 (JST) の戦略的基礎研究推進事業 (CREST) 「高度メディア社会の生活情報技術」プログラムによる。

参考文献

- [1] 上岡 隆宏, 河村 竜幸, 浮田 宗伯, 河野 恭之, 木戸 正継, “ウェアラブル装置を用いたオブジェクト登録・検索システムの基礎実験”, ITE Technical Report Vol. 26, No. 7, PP. 25-30, IPU2002-6 (Jan. 2002)
- [2] D. Yagi, K. Abe, and H. Nakatani, “Segmentation of color aerial photographs using HSV color models”, Proc. MVA’ 92, pp. 367-370, 1992.