生命化建築のための深度データを用いた 居住者の行動の分類と行動認識に関する研究

HUMAN ACTIVITY RECOGNITION BASED ON THE ACTIVITY CATEGORIZATION USING DEPTH DATA FOR “BIOFIED BUILDING”

小川 愛実*, 三田 彰**

Ami OGAWA and Akira MITA

The number of single households is gradually increasing in Japan. The majority of them are consisting of elderly people. The “Biofied Building” that can monitor and control living spaces will help those people in many aspects. One of the key items to be acquired is the information about the resident’s activity. In this paper, we suggest an activity recognition method for two categories of activities, “Single type” and “Multiple type.” We use the consecutive depth data obtained by Kinect. The R transformation, variance, PCA, LDA, and k-means classification are used for classification. Ten activities data sets consisting of six “Single type” and four “Multiple type” conducted by eleven subjects are used for verification.

Keywords: Biofied Building, Depth Data, Human Activity Recognition, Activity Categorization, R Transformation, Variance

生命化建築, 深度データ, 行動, 行動分類, R 変形, 分散

1. はじめに

近年世界に先駆けて超高齢社会となった日本では世帯構成の変化による単身世帯の増加が著しく、今後も増加傾向にあると予想されている。単身世帯においては居住者の年齢層や好みによってそのニーズが多岐にわたり、従来の住宅モデルが必ずしも当てはまらないことは明らかである。さらに近年ますます国際交流が盛んになったことで、人種による体質の違いや宗教による生活形態の違いなど、居住者像と求められる居住空間が多様化している。これらの背景より、居住者の多様なニーズに対応できる居住空間が求められている。

そこで我々は、居住者一人ひとりに合った空間制御を行う「生命化建築」¹⁾を提案している(Fig.1)。「生命化建築」では、生物の適応機能に倣ったシステムを用い、居住者と居住空間をつなぐインタフェースとして小型のセンサエージェントロボットを用いている。ロボットが居住空間内のあらゆる情報をデータベースに蓄積し、その膨大なデータに基づいて連動したデバイスへの指令値を送信することで、居住者のその時々々の状態に合わせた制御が可能となる。例えばある照明器具を制御することを想定した場合、その時の季節、時間、気温などの環境情報や、居住者のその時の好み、行動などの居

住者情報を取得し、過去のデータに基づいた適切な初期値を入力する。さらに居住者の不快情報を随時入力することで微調節を行い、最適な制御へと収束させるというシステムが挙げられる。このとき、ある時点において特定の居住者が求めている制御を行うためには、居住者の行動情報が必要不可欠である。

そこで本研究は「生命化建築」を実現するため、居住空間内における人の行動を認識することを目的としている。人の行動認識技術は店舗での万引きなどの犯罪検知やライフログを用いた消費エネルギーの最適化など幅広いニーズのある分野であり、これまでに多くの研究がなされている^{2)~9)}。ここで著者らは居住空間内に焦点を絞るために居住者の拘束条件およびプライバシー保護を考慮して、センサから対象物までの距離情報である深度データのみを用いた行動認識手法を提案する。人の行動認識手法に関する既往研究の多くは、「立つ」「座る」などの「単一の動作による行動」を対象としている^{2)~9)}。一方で居住空間において頻繁に行われる「食べる」「書く」などの「複数の動作による行動」を対象とした行動認識手法の提案は比較的少ないと言える⁹⁾。そこで著者らは「複数の動作による行動」を対象に含めた行動認識手法を提案することで、より現実的な居住

* 慶應義塾大学大学院理工学研究科 大学院生

** 慶應義塾大学 教授, Ph.D.

Grad. Stud., Keio University

Prof., Keio University, Ph.D.

空間への応用を目指した。本研究では、居住空間内における行動にはそれを構成する動作の単複による分類が可能であるとして、その定義と分類、およびそれらの性質を考慮した行動認識手法の提案により、居住空間における行動認識システムの提案を行う(Fig.2)。はじめに、著者らの定義した行動分類について述べた上で、行動認識における既往研究を概観し、行動分類および「単一の動作による行動」に対する提案手法に続いて「複数の動作による行動」に対する提案手法を示す。最後に検証実験の概要、結果および考察を述べ、結論にまとめる。

Fig.1 The concept of the "Biofied Building"

Fig.2 The flow of activity recognition system

2. 行動と分類の定義

小林(2010)は「動作は『ひとまとまりの型が決まった運動』、行動は『いろいろな動作を組み合わせてつくる、ある目的を達成するための行為』であるとしている¹⁰⁾。例として「食べる」という行動を挙げると、箸を持つ動作、食べ物を箸でつかむ動作、つかんだ食べ物を口へ運ぶ動作などの、複数の動作により構成されると言える(Fig.3)。一方で「立つ」「座る」などの行動は、動作であると同時に行動でもある(Fig.4)。すなわち行動とはある目標を達成するための単一の動作、或は複数の動作群である。

著者らは本研究において居住空間における人の行動を「単一の動作による行動」および「複数の動作による行動」の2種類に定義した。

Fig.3 Activity of "Eating" consisting of multiple movements

Fig.4 Activity of "Standing up" consisting of single movement

3. 先行研究と提案手法の方針

(1) 使用するデータ

人の行動認識に関する先行研究では画像^{2)-8), 11)-14)}や加速度^{6), 9)}が多く用いられている。「生命化建築」では居住者の行動情報を逐次データ化するため、データの取得を終日行う必要がある。加速度計を用いる場合センサを装着することで居住者を拘束するため、長期的な測定を前提とする本研究の目的には適さない。画像を用いた研究ではRGB、深度、二値、熱赤外線画像などが用いられている。このうちRGB、深度、二値画像はデバイスの入手が容易である一方、RGB画像の居住空間での使用はプライバシーの侵害が懸念されるため好まれない。また二値画像よりも深度画像を用いたほうが行動認識の精度が高いことが報告されている³⁾¹¹⁾。

よって本研究では居住者への拘束性およびプライバシー保護の観点から、居住空間に設置したセンサより深度データのみを用いることとした。使用するセンサはマイクロソフト社のKinect for windows(以下Kinect)である。KinectはFig.5に示したようにRGBカメラ、赤外線センサ、チルトモータおよびマイクロフォンアレイを搭載しており、RGBデータ、深度データ、音声データを取得できる。またスケルトントラッキング機能およびフェイストラッキング機能を持つ。

(2) 特徴抽出手法

行動や動作を分解した最小セグメントである姿勢の認識に関する研究の中でも、深度データを用いているものの多くは機械学習を用いてより厳密に身体の部位を認識している¹⁵⁾⁻¹⁹⁾。一方で、行動を認識するには動画の切り出しなどの連続した複数のフレームをひとかたまりとして扱う必要がある。各フレームにおける身体の部位をすべて認識するにはコストがかかるため、行動認識ではよりシンプルに深度データから特徴を抽出することが求められる。

深度データの特徴抽出において、主成分分析(PCA)、独立成分分析(ICA)、線形判別分析(LDA)が多く用いられている^{3)-7), 13), 14)}。中でもPCAとLDAの両方を用いた研究が多く^{3)-6), 13), 14)}、PCAのみを用いた手法よりも識別精度が高いという報告がある⁶⁾。よって本研究ではPCAおよびLDAを用いる。

またPCAおよびLDAを適用する前に、対象となる人の体型や位置などの特徴を排除し、特徴量の次元を削減する手法としてRadon変換が用いられている³⁾⁴⁾¹²⁾。さらにRadon変換とR変換を組み合わせたR変形の精度がより高いと報告されている³⁾。他にも、連続した熱赤外線データの各ピクセルの和をとった歩行エネルギー画像が複数フレームの特徴抽出手法として提案されている⁵⁾。

しかし、これらの先行研究では先に述べた「単一の動作による行

動」を対象として高い識別率が報告されている一方で、「複数の動作による行動」を対象とした結果は報告がない。「複数の動作による行動」においては行動中に同じような動作が継続的に繰り返されるとい特徴があり、深度データ群の中で動きの頻度やその位置が特徴量となり得る。そこで著者らは歩行エネルギー画像⁵⁾より着想を得て、連続した深度データの分散を算出する手法を新たに提案する。

本研究では PCA および LDA 適用前の特徴量抽出手法として、R 変形および分散を用いる。R 変形の「複数の動作による行動」への適用結果を明らかにするとともに、各フレーム群の分散を特徴量として用いる手法を新たに提案することで、両手法の両行動種類に対する適性を明らかにする。

Fig.5 Kinect and depth image of Kinect

4. 提案手法

(1) 行動分類および「単一の動作による行動」の認識

本研究は「生命化建築」を目的とした居住空間内での応用を想定しているため、対象となる居住者に対し非拘束的かつプライバシー保護を考慮した Kinect の深度データを使用する。行動情報は時間に依存した姿勢の変化であるため、1 つの行動につき連続した複数のフレームのデータを用いる。

提案する行動認識手法のフローを Fig.6 に示した。主に前処理、特徴抽出および識別の 3 つのフェーズから成る。前処理は Visual Studio 2010 C++ (Microsoft, USA)、特徴抽出および識別では Matlab (MathWorks, USA) を用いて開発を行った。

まず前処理として深度データの取得および背景を含むノイズ除去を行った。あらかじめ取得しておいた被験者がいない状態の深度データを背景とし、各深度データとの差分を取った。また、実験環境とセンサの取得範囲を考慮した閾値を設置することでノイズ除去を行った。本実験においてはセンサから壁までの距離が 3m であったため 3m 以上の値をノイズと判断した。

次に特徴抽出として、R 変形、PCA および LDA を用い、学習データによって変換行列を求め、テストデータに適用する。

最後に k-means 法を用いて行動を認識した。

R 変形は画像解析に用いられる手法であり、対象物の位置や大きさの変化に依存しないことが特徴である。R 変形でははじめに Radon 変換を、続いて R 変換を行う。

Radon 変換では、1 から 180°までの角度 θ の軸に対して深度データをそれぞれ積分し(式(1)²⁰⁾、 $R_\theta(x_\theta)$ を得る。このとき x_θ および y_θ は式(2)²⁰⁾で表せる。Fig.7 に Radon 変換の概念図を示した。

$$R_\theta(x_\theta) = \int_{-\infty}^{\infty} f(x_\theta \cos \theta - y_\theta \sin \theta, x_\theta \sin \theta + y_\theta \cos \theta) dy_\theta \quad \dots (1)$$

$$\begin{bmatrix} x_\theta \\ y_\theta \end{bmatrix} = \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \quad \dots (2)$$

R 変換では θ 毎に $R_\theta(x_\theta)$ の二乗和を算出する(式(3))。

$$R_T(x_\theta) = \int_{-\infty}^{\infty} R_\theta(x_\theta)^2 dy_\theta \quad \dots (3)$$

R 変形により 1 枚の画像を 1×180 の行列に変換し、連続フレームのデータを 1 行に並べることでシーケンスとして扱う。R 変形のプロセスを Fig.8 に示した。

PCA では R 変形処理後の全データの分散を 1 に標準化した後、固有値の大きさが 1 以上となる主成分のみを採用することで次元を決定した²¹⁾²²⁾。

LDA では深度データのうち 40%を学習データ、60%をテストデータとした。まず学習データを LDA 処理することで、クラス内共分散行列を Σ_W 、クラス間共分散行列を Σ_B とした、 $\Sigma_W^{-1}\Sigma_B$ の固有ベクトルを変換行列として求め、テストデータに用いた。

k-means 法では各プロトタイプの初期値として学習データの各クラスターの平均値を用いた。

Fig.6 Flow of categorization and "Single type" activity recognition

Fig.7 Abstract of Radon transform (文献 15 を参考に作成)

Fig.8 Process of feature extraction by R transformation

(2) 「複数の動作による行動」の認識

「複数の動作による行動」に対しても上記と同様の手順にて行動認識を行うが、前処理に深度データの切り取りを加え、R変形に代わって分散を用いる。(Fig.9)

「複数の動作による行動」は基本的に上半身のみの動作によるため、深度データ 640×480 ピクセルのうち上半身を表すピクセルを切り取り、その他のノイズの影響を除外することで識別精度の向上を図る。

連続するフレームの各ピクセルにおける深度データの分散を算出することで、どの部位がどれだけの頻度で動いているかという特徴を抽出する。分散の算出手順を Fig.10 に示した。まず、n 枚の連続したフレームの各ピクセルにおける深度データを取り出し、その分散を算出する。算出された分散をそのピクセルの位置に当たるセルに配置し、これをすべてのピクセルにおいて行うことで深度データと大きさの等しい分散行列を作る。作成した分散行列の列ごとおよび行ごとに分散を算出し、それぞれ 1×列数および 1×行数の行列とする。最後にそれらを繋ぎ合わせ、1×(列数+行数)の行列とする。

Fig.9 Flow of "Multiple type" activity recognition

Fig.10 Process of feature extraction by variance

5. 検証実験と結果

(1) 行動データの取得

本研究における深度データ取得実験の配置図を Fig.11 に示した。「生命化建築」において提案されている小型ロボットの使用を想定して、センサは地上より 0.2m の高さに設置した。対象とした行動は「単一の動作による行動」として、「立つ」「座る」「しゃがむ」「しゃがんだ状態から立つ」「物を拾う」「物を拾った状態から立つ」の 6 種類、「複数の動作による行動」として、「スマートフォンを操作する」「本を読む」「書く」「食べる」の 4 種類の合計 10 種類であった。このうち「しゃがむ」「しゃがんだ状態から立つ」「物を拾う」「物を拾った状態から立つ」の 4 種類は Fig.11 の左図に示した配置、「スマートフォンを操作する」「本を読む」「書く」「食べる」の 4 種類は右図に示した配置、「立つ」「座る」の 2 種類は右図から机を除いた椅子のみの配置でそれぞれのデータを取得した。被験者は 20 代の男性 6 名、女性 5 名の合計 11 名で、身長は 165.5±9.1cm であった。またシルエットに影響するものとして、服装は 4 名がスカートを、7 名がパンツを穿いており、髪型は短髪が 6 名、長髪が 5 名であった。被験者は 10 種類の行動を 5 回ずつ、延べ 50 回試行した。各試行につき連続した 10 フレームの深度データを CSV 形式で保存した。被験者 1 人につき 10 フレーム×5 試行×10 種類=500 データを取得し、合計 5500 の CSV ファイルを保存した。

Fig.11 Experiments set-up

(2) 平均識別率および標準偏差

入力データの識別結果は以下の式(4)より識別率として得られる。

$$\text{識別率 [\%]} = \frac{\text{全試行数中正しく識別された試行数}}{\text{ある行動の全試行数}} \times 100 \quad \dots (4)$$

また各行動の識別率より、いくつかの行動の平均識別率および標準偏差を以下の式(5)および式(6)のように求めた。このとき識別率は正規分布に従うと仮定した。

$$\text{平均識別率 [\%]} = \frac{\text{各行動の識別率の和}}{\text{行動数}} \quad \dots (5)$$

$$\text{標準偏差 [\%]} = \sqrt{\frac{\sum (\text{ある行動の識別率} - \text{平均識別率})^2}{\text{行動数}}} \quad \dots (6)$$

(3) 行動分類手法の検証

6 種類の「単一の動作による行動」と 4 種類の「複数の動作による行動」に対して Fig.6 に示した提案手法を適用し、2 つの行動群に分類した。また比較のため R 変形の代わりに分散を用いた手法に

よる分類も行った。結果を Table 1 に示した。

R 変形を用いた提案手法を適用した結果、各行動群の平均識別率は 100%となり、すべての行動を正しく分類することに成功した。LDA 適用後のデータのプロットを Fig.12 に示した。

一方で分散を用いた手法を適用した結果、各行動群の平均識別率は 97.8%，標準偏差は 3.21%となった。

Table 1 The Recognition rate of categorization [%]

		Recognition			
		R transformation		Variance	
True	Single	100	0	95.5	4.55
	Multiple	0	100	0	100

Fig.12 LDA data classification of categorization using R transformation

(4) 行動認識の検証

分類された 6 種類の「単一の動作による行動」に対し、再び Fig.6 に示した提案手法を適用し行動認識を行った。また比較のため R 変形の代わりに分散を用いた行動認識も行った。

次に分類された 4 種類の「複数の動作による行動」に対し、Fig.9 に示した提案手法を適用し行動認識を行った。また比較のため分散の代わりに R 変形を用いた行動認識も行った。

結果を Table 2 に示した。

「単一の動作による行動」認識

Table 2 より「単一の動作による行動」認識の平均識別率は R 変形による結果が 91.4%であったのに対し分散による結果は 57.1%と極めて低く、提案手法が適しているといえる。また分散を用いた手法では標準偏差が大きい値をとることから、行動によって識別率にばらつきが生じていることがわかる。

提案手法による LDA 適用後のデータのプロットを Fig.13 に、各行動の識別率を Table 3 に示した。

「複数の動作による行動」認識

深度データ 640×480 ピクセルのうち上半身を表す(86,116)から(385,265)の範囲である 300×150 ピクセルを切り取った。切り取りの前後の深度データの分散データを Fig.14 に示した。

Table 2 より、「複数の動作による行動」認識の平均識別率は R 変

形による結果が 75.8%であったのに対し分散による結果は 77.3%となり、提案手法がより高い識別率を得ているが大きな差異は見られない。しかし標準偏差に着目すると、R 変形による結果は 21.4%であったのに対し分散による結果は 9.09%とより小さい値をとっていることがわかる。すなわち、すべての行動について安定した識別率を得たということであり、「複数の動作による行動」に対して提案手法がより適していることを裏付ける。

提案手法による LDA 適用後のデータのプロットを Fig.15 に、各行動の識別率を Table 4 に示した。

Table 2 Comparison of the mean recognition rate of each activity type applied two methods (mean±S.D.[%])

	R transformation	Variance
Single type	91.4±9.66	57.1±31.6
Multiple type	75.8±21.4	77.3±9.09

Table 3 The Recognition rate of “Single type” activities [%]

		Recognition					
		SU	SD	CR	RC	PU	UP
True	SU	97.0	0	0	3.03	0	0
	SD	0	72.7	21.2	0	6.06	0
	CR	0	6.06	93.9	0	0	0
	RC	0	0	0	90.9	0	9.09
	PU	0	3.03	3.03	0	93.9	0
	UP	0	0	0	0	0	100

SU=standing up; SD=sitting down; CR=crouching; RC=standing up from crouching; PU=picking up; UP=standing up from picking up.

Table 4 The Recognition rate of “Multiple type” activities [%]

		Recognition			
		SP	RB	WR	EA
True	SP	72.7	27.3	0	0
	RB	9.09	84.8	3.03	3.03
	WR	0	15.2	84.8	0
	EA	0	6.06	27.3	66.7

SP=handling a smartphone; RB=reading a book; WR=writing; EA=eating.

Fig.13 LDA data classification of “Single type” activities

Fig.14 Original data and trimmed data of “Multiple type” activities

Fig.15 LDA data classification of “Multiple type” activities

6. 考察

(1) 行動分類の必要性

行動分類の必要性を検討するため、「単一の動作による行動」および「複数の動作による行動」を合わせた全 10 種類の行動に対し、行動分類を行わずに R 変形および分散を用いた手法により行動認識を行った。比較のため行動分類を行った場合と合わせて、各行動の平均識別率を Table 5 に示した。R 変形を用いた結果は 72.7%，分散を用いた結果は 42.4% となり、行動分類の必要性が示された。

Table 5 Comparison of the mean recognition rate of categorized and uncategorized activities (mean±S.D.[%])

	R transformation	Variance
Uncategorized	72.7±25.9	42.4±40.6
Single type	91.4±9.66	57.1±31.6
Multiple type	75.8±21.4	77.3±9.09

(2) 「複数の動作による行動」認識手法におけるデータ切り取りの必要性

「複数の動作による行動」の認識において深度データの上半身領域を切り取る必要性を検討するため、切り取りを行わずに認識を行った。R 変形および分散を用いた両手法による結果を、比較のため切り取りを行った場合と合わせて Table 6 に示した。

最も平均識別率が高かったのは提案手法であり、標準偏差も小さい値をとった。一方で提案手法から切り取りの過程を除いた結果、平均識別率は 30.3% と非常に低い値をとった。深度データは主に物体の境界におけるノイズが多いため値がちらつくことを避けられない。そのようなデータの分散をとると、周辺環境の様々な輪郭にお

いて値をとってしまい、ノイズの影響が大きくなると考えられる。よって、「複数の動作による行動」の認識を行う場合は上半身の領域を切り取る必要があることが示された。

Table 6 Comparison of the mean recognition rate of trimmed and untrimmed “Multiple type” activity data (mean±S.D.[%])

	Trimmed	Untrimmed
R transformation	75.8±21.4	75.8±18.0
Variance	77.3±9.09	30.0±36.9

7. 結論

本研究は居住空間内における人の行動認識を目的として、行動を構成する動作が単一或は複数であるかによって行動を「単一の動作による行動」および「複数の動作による行動」の 2 種類に定義し、その分類手法とそれぞれの行動種類に応じた行動認識手法を提案した。居住空間での使用を目的とするため、非拘束性とプライバシーの保護を考慮して Kinect の深度データのみを用いた。R 変形または分散と PCA および LDA を適用して特徴量を抽出し、k-means 法によって合計 10 種類の行動を識別した。特徴抽出手法として行動分類および「単一の動作による行動」認識には R 変形を、「複数の動作による行動」認識には分散を用いた。結果として行動分類では 100%，「単一の動作による行動」認識では 91.4%，「複数の動作による行動」認識では 77.3% の平均識別率を得た。さらに提案手法の妥当性を検証するため各行動群および 2 つの行動群を合わせた全体に対し 2 通りの提案手法を適用し、比較を行った。結果、行動分類の必要性および「単一の動作による行動」には R 変形、「複数の動作による行動」には分散の適用が好ましく、「複数の動作による行動」は上半身の領域を切る取ることによって精度の向上がみられるという結論を導いた。本研究では年齢に偏りのある少数の被験者のみでの検討であったため、今後はサンプル数を増やしサンプリングバイアスを削減することで、手法の汎用性についての検討も行いたい。

謝辞

本研究の一部は科学研究費補助金基盤(B) (課題番号 15H02890 研究代表者三田彰) および文部科学省研究拠点形成費等補助金(博士課程教育リーディングプログラム)の助成を受けて実施された。

参考文献

- 1) 秋葉達也, 三田彰: 生命化建築のためのセンサエージェントロボットを用いた人追従, 日本建築学会技術報告集, Vol. 18, No. 39, pp.775-779, 2012.6.
- 2) Ji, S., Xu, W., Yang, M., Yu, K.: 3D Convolutional Neural Networks for Human Action Recognition, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 35, No.1, pp.221-231, 2013.
- 3) Jalal, A., Uddin, M. Z., Kim, T.-S.: Depth Video-based Human Activity Recognition System Using Translation and Scaling Invariant Features for Life Logging at Smart Home, IEEE Transaction on Consumer Electronics, Vol. 58, No. 3, pp.863-871, 2012.
- 4) Khan, Z. A., Sohn, W.: Real Time Human Activity Recognition Systems based on Radon Transform, IJCA Special Issue Artificial Intelligence Techniques Novel Approaches & Practical Applications, pp.31-37, 2011.
- 5) Han, J., Bhanu, B.: Human Activity Recognition in Thermal Infrared Imagery, Proc. IEEE Computer Society Conference on Computer Vision

- and Pattern Recognition, pp.17-24, 2005.
- 6) Mojidra, H. S., Borisagar, V. H.: Acceleration and Principle Component Based Human Activity Recognition via Hidden Markov Model, International Journal of Computational Linguistics and Natural Language Processing, Vol 2, Issue 5, pp.350-354, 2013.
 - 7) Uddin, M. Z., Lee, J. J., Kim T. S.: Independent shape component based human activity recognition via Hidden Markov Model, Springer Science + Business Media, 33, pp.193-206, 2010.
 - 8) Wang, L., Suter, D.: Learning and Matching of Dynamic Shape Manifolds for Human Action Recognition, IEEE Transactions on Image Processing, Vol. 16 No. 6, pp.1646-1661, 2007.
 - 9) Gu, T., Wang, L., Wu, Z., Tao, X., Lu, J.: A Pattern Mining Approach to Sensor-Based Human Activity Recognition, IEEE Transaction on Knowledge and Engineering, Vol. 23, No. 9, pp.1359-1372, 2011.
 - 10) 小林朋道: ヒトはなぜ拍手をするのか 動物行動学から見た人間, 株式会社新潮社, 2010.12.20, p.68.
 - 11) Munoz-Salinas, R., Madina-Carnicer, R., Madrid-Cuevas, F. J., Carmona-Poyato, A.: Depth silhouettes for gesture recognition, Pattern Recognition Letters, 29, pp.319-329, 2008.
 - 12) Singh, M., Mandal, M., Basu, A.: Pose Recognition using the Radon Transform, Proc. IEEE 48th Midwest Symposium on Circuits and Systems, pp.1091-1094, 2005.
 - 13) Eleyan, A., Demirel, H.: PCA and LDA Based Face Recognition Using Feedforward Neural Network Classifier, Proc. Multimedia Contents Representation Classification and Security, pp.199-206, 2006.
 - 14) Ji, S., Ye, J.: Generalized Linear Discriminant Analysis: A Unified Framework and Efficient Model Selection, IEEE Transactions on Neural Networks, Vol. 19 No.10, pp.1768-1782, 2008.
 - 15) Baak, A., Müller, M., Bharaj, G., Seidel, H., Theobalt, C.: A data-driven approach for real-time full body pose reconstruction from a depth camera, IEEE international conference on computer vision, pp. 1092-1099, 2011.
 - 16) Helten, T., Baak, A., Müller, M., Theobalt, C., Grzegorzec, M.: Full-Body Human Motion Capture from Monocular Depth Images, Time-of-Flight and Depth Imaging, LNCS 8200, pp. 188-206, 2013.
 - 17) Du, H., Wang, J., Zhong, X., He, Y., Mei, L.: Human Skeleton Extraction of Depth Images Using the Polygon Evolution, Digital Human Modeling. Applications in Health, Safety, Ergonomics and Risk Management, LNCS 8529, pp. 15-23, 2014.
 - 18) Plagemann, C., Ganapathi, V., Koller, D., Thrun, S.: Real-time identification and localization of body parts from depth images, Proceedings of ICRA, pp. 3108-3113, 2010.
 - 19) Schwarz, L.A., Mkhitarian, A., Mateus, D., Navab, N.: Human skeleton tracking from depth data using geodesic distances and optical flow, Image and Vision Computing, Vol. 30, No. 3, pp.217-226, 2012.
 - 20) Mathworks ホームページ,
<<http://jp.mathworks.com/help/images/radon-transform.html>>,
(参照 2015.2.11)
 - 21) Kano, M., Hasebe, S., Hashimoto, I., Ohno, H.: A new multivariate statistical process monitoring method using principal component analysis, Computers and Chemical Engineering, Vol. 25, No. 7-8, pp. 1103-1113, 2001.
 - 22) Kano, M., Hasebe, S., Hashimoto, I., Ohno, H.: Statistical process monitoring based on dissimilarity of process data, AIChE J., Vol. 48, No. 6, pp. 1231-1240, 2002.

HUMAN ACTIVITY RECOGNITION BASED ON THE ACTIVITY CATEGORIZATION USING DEPTH DATA FOR “BIOFIED BUILDING”

Ami OGAWA and Akira MITA***

* Grad. Stud., Keio University

** Prof., Keio University, Ph.D.

The number of single households consisting of elderly people is increasing in Japan. We are conducting series of research on the “Biofied Building” that can control the living spaces suitable for each resident’s needs. This system utilizes the algorithm learnt from living organisms with the help of small sensor agent robots. In this system, every information in living space is collected by the robots for the space control. The information about the resident’s activity is one of the most important data for the “Biofied Building”. When we acquire the information on activities, we have to add tags on the data such as meaning of these activities. For this purpose, we proposed an activity recognition method for residents in living spaces.

An activity is defined as an action done aimed at some target, and consists of several movements. A movement is a motion defined by the unit body form. In this study, we defined two types of activities, activities consisting of single movement and activities consisting of multiple movements. For example, the activity of “eating” consisting of multiple movements, such as; “holding chopsticks” and “catching food” etc., is classified into “Multiple type”. On the other hand, the activity of “standing up” consisting of a single movement of “standing up”, is classified into “Single type”. The activities in living space are thus categorized into two.

In this study, we used Kinect for Windows by Microsoft to acquire the depth data. After the noise rejection, R transformation or variance, PCA, LDA, and k-means classification were applied to recognize the activities. Our proposed method has two phases, the activity categorization phase, and the activity recognition phase.

The ten activities including the six “Single type” activities (“standing up”, “sitting down”, “crunching”, “standing up from crunching”, “picking up something” and “standing up from picking up something”), and the four “Multiple type” activities (“handling a smart phone”, “reading a book”, “writing” and “eating”) were performed by eleven subjects.

We proved that it was effective to categorize the activities into each activity type for activity recognition, because each activity type has own features. We had 100% of mean recognition rate of activity categorization. And we had better mean recognition rates of activity classification with the activity categorization than without it.

After the activity categorization phase, the activity classification using R transformation and variance was applied. These two methods were used for transforming two dimensional depth data into one dimensional parameters. R transformation is often used for image analysis in several conventional methods. However, the R transformation for “Multiple type” activities was rarely referred. We applied both methods to “Multiple type” activities and found that they worked well. In addition, we examined the effect of variance for the transformation of two dimensional matrix. As “Multiple type” activities consist of continuous movements in limited space, we hypothesized that variance of several consecutive depth data may be able to detect the features of “Multiple type” activities.

As the results, we got 91.4% of mean recognition rate of “Single type” activities in case of using R transformation. And we also got 77.3% of mean recognition rate of “Multiple type” activities in case of using variance. Although some improvement is needed, the proposed method showed a certain level of feasibility for practical applications.

(2015 年 8 月 8 日原稿受理, 2016 年 1 月 8 日採用決定)