

# CHLAC 特徴量と部分空間法による複数行動の分類

佐藤 竜太<sup>†</sup> 亀田 能成<sup>†</sup> 大田 友一<sup>†</sup>

<sup>†</sup> 筑波大学大学院システム情報工学研究科 〒305-8573 茨城県つくば市天王台 1-1

E-mail: †{rsato,kameda,ohta}@image.iit.tsukuba.ac.jp

あらまし 本稿では、監視カメラに映る様々な物体の行動を、CHLAC 特徴量をもとに識別する方法を提案する。位置不変で可算可能な CHLAC は、監視目的での映像解析における有力な画像特徴量の一つと考えられる。この CHLAC 特徴量をもとに、部分空間法を用いて複数の行動識別を行う。歩行者、自転車、自動車の 3 種類の物体による合計 9 種類の動作に対して実験を行い、82.9%の識別結果を得た。

キーワード CHLAC, 部分空間法, パターン認識

## Actions Classification Based on CHLAC Features and Subspace Method

Ryuta SATO<sup>†</sup>, Yoshinari KAMEDA<sup>†</sup>, and Yuichi OHTA<sup>†</sup>

<sup>†</sup> Graduate School of Systems and Information Engineering, University of Tsukuba 1-1-1 Tennoudai, Tsukuba, Ibaraki, 305-8573 Japan

E-mail: †{rsato,kameda,ohta}@image.iit.tsukuba.ac.jp

**Abstract** This paper describes an action classification method based on CHLAC features of the movies from a surveillance camera. CHLAC feature is a prominent image feature for general object and action recognition because it has useful property for surveillance video analysis. We exploit CHLAC and form a classifier based on subspace approach. We conducted a preliminary experiment of nine single actions done by pedestrian, bicycle, and vehicle and confirmed that the classification ratio to 82.9 %.

**Key words** CHLAC, Subspace Method, Pattern Recognition

### 1. はじめに

シーン中で様々な物体が起こす行動を分類・識別することには、監視目的を含め様々なアプリケーションが想定されるため、その社会的貢献は大きい。それゆえ、一般行動認識は画像認識における大きな研究分野の一つである。

一般に監視目的でカメラが設置されている場合、映像中で監視対象となる行動は多様であるので、特徴量の選定や認識器の構成方法には慎重な配慮が求められる。Cubic Higher-order Local Auto-Correlation (CHLAC) [1] は、このような監視目的の一般物体認識における有力な画像特徴量の一つである。CHLAC は、単純な時空間特徴量でありながら、画像全体に現れる時空間変動をコンパクトに表現できるという利点を持つ。また、物体の出現位置に不変であることから、シーン中への侵入物体の位置に関わらず同じ記述が得られることや、複数の同時動作に対してはその各々の動作に対応する特徴量の和によって表現できることなど、監視カメラでの動作認識に有用と思われる性質を有する。計算コストも低い部類に入るため、様々な用途にも応用しやすい。

そこで、本稿では、この CHLAC 特徴量をもとに、種類の異なる物体が行う様々な動作を識別する方法を提案する。CHLAC 特徴量の特徴の一つに、計算コストの低さが挙げられるため、識別器にも、識別時の計算コストが低い部分空間法を採用する。

なお、CHLAC 特徴量は、フレーム間差分によって得られた 2 値映像 3 フレーム分から 251 次元のベクトルとして得られる。そこで、本稿では、動作識別の定義を以下のように定める。ある動作は一定時間に渡って行われるため、その動作中に得られた全ての CHLAC 特徴量ベクトルには、全てその動作名が割り当てられる。複数回の同一動作が行われた場合も同様に、そこで得られた全ての CHLAC 特徴量ベクトルには、同一動作名が割り当てられる。本稿における動作識別とは、事前に与えられた複数の異なる動作に対して、入力映像がどの動作に最も近いかをフレーム単位で評価することである。

以降、まず第 2. 章では関連研究について言及する。そのあと、第 3. 章では動作を撮影した映像および CHLAC 特徴量について述べ、第 4. 章では識別器として使用する部分空間法とその構成について述べる。第 5. 章では CHLAC 特徴量と部分空間法を動作識別に使用する方法

について述べ、第 6. 章で実験とその結果について述べる。最後に第 7. 章でまとめを述べる。

st-n

## 2. 関連研究

### 2.1 画像特徴量の検討

人間を対象とした一般行動認識では、対象の形状や属性に依存しない特徴量に基づくアプローチが、その応用性の広さから注目を集めている（例えば [2] [3] [4] 等）。それに対して、CHLAC [1] に基づく行動認識の報告は、その特徴量の性質の良さを考えるとまだそれほど多くはない。

ここで、いくつかの画像特徴量について検討してみる。

SIFT (Scale Invariant Feature Transform) [8] は本研究での求められる条件である、対象に対する非依存性・位置への非依存性に対応できる優れた手法である。しかし、時間方向の微分に反応しないため、そのままでは動作認識に対応しない。

動作へ対応できるものとして、時間方向の畳み込みを行う特徴量である、ST-patch [9] がある。これは動作へ対応しているが、計算コストが高いという難点がある。

また、SIFT や ST-patch のようなインタレストオペレータベースの局所特徴量では、あるフレーム上での 1 つの動作が含まれる画像全体での性質を扱うために、それらの局所特徴量を統合して扱う処理が必要となるという問題点がある。

これに対して、動作全体を表す特徴量の構成が容易な HLAC [10]、CHLAC [1] に注目する。

HLAC [10] は 2 次元の局所領域における相関パターンの出現を全体で数え上げることで特徴の算出を行う。そのため、画像全体からの特徴を記述することが可能である。さらに、画像全体から相関を取るというその性質上、位置の影響を受けない。

2 次元の局所領域を用いる HLAC に対して、3 次元の局所領域を用いる特徴量が CHLAC [1] である。3 次元を画像平面の 2 次元と時間軸の 1 次元として用いることで、時間方向の情報も得ることができる。

CHLAC 特徴量の応用事例として、これまで、動作認識 [7] [11] や異常検出 [5] [6] が行われてきている。

白木ら [7] は CHLAC および Grid コンピューティングを用いた動作認識を提案しているが、研究の主題が高速化であるため、動作認識には並列化した線形認識を用いており、動作の種類が増えた場合、認識精度向上が見込めるかどうか不明である。

一方、森下ら [11] は時間方向への伸縮などを考慮した CHLAC 利用の発展形について提案し、2 種類の動作について実験も行っているが、その主眼は時間的に変動を含む動作への適応であり、複数動作認識が目的ではない。

CHLAC を用いた異常検出のための手法は複数提案されている [5] [6] が、いずれも正常状態のあるクラスとし

てそこからの逸脱を異常と定義しているため、本研究目的のような複数動作認識には対応していない。

そこで、本稿では CHLAC を複数行動認識に用い、実際の実験結果を示すとともに、その結果から考察を行う。

### 2.2 識別器の検討

CHLAC のような高次元データベクトルに対するパターン識別手法としてさまざまな手法が提案されている。

ニューラルネットワークはバックプロパゲーション [12] を利用することで、線型分離不可能な学習データに対しても学習を行うことができるが、汎化性能を保証するものではなく、未知データに対する問題がある。

SVM [13] はマージン最大化することで、ニューラルネットワークにおける入力された未知データに対する挙動の問題を低減することができる。しかしながら、多クラスへの対応は向いておらず、対応したとしても、複雑なアルゴリズムや高い計算コストが必要である。

それに対し、部分空間法 [14] は、特徴ベクトルを高次元である特徴空間の部分空間に射影することで識別する手法であり、高次の特徴が一部の次元に偏って存在する場合に適している。

そこで、本稿では CHLAC 特徴量を部分空間により各動作に識別する。

## 3. CHLAC 特徴量による動作の表現

本節では CHLAC について概要を簡単に説明する。

### 3.1 HLAC 特徴量

(1) 式をある関数  $f(x)$  に対する  $N$  次自己相関という。

$$\int f(x)f(x+a_1)f(x+a_2)\cdots f(x+a_N)dx \quad (1)$$

二値化された画像上のある点  $p$  における高次自己相関について考えるとき、注目画素と近傍の領域は相関が高いと予想できることから、すべての領域で相関をとるのではなく、局所領域内からのみ相関を算出することを考える。これを HLAC (Higher-order Local Auto-Correlation) [10] 特徴量とよぶ。

### 3.2 CHLAC 特徴量

3.1 節で述べた HLAC は 2 次元の局所領域での相関を求めるものであったが、局所領域に時間軸を導入し、 $3 \times 3 \times 3$  の局所領域で相関を求めるものが CHLAC (Cubic Higher-ordered Local Auto-Correlation) [1] である。HLAC と同じく平行移動に対してユニークな相関を 2 値差分画像の連続 3 枚の中の画像上で求めると 251 通りとなる。ゆえに、CHLAC 特徴量  $x_D$  は 251 次元のベクトルである。

## 4. 部分空間法による多クラス分類

本稿では、特徴量ベクトルの多クラス分類に比較的古典的な CLAFIC 法を用いる。また、識別精度向上のため

めに，部分空間の構成に際しては直交化を行う．これらに対する説明と，識別方法についての説明を以下の各節で行う．

#### 4.1 CLAFIC 法

入力ベクトル  $x_q$  を  $N_c$  個のクラスに識別する問題を考える．学習データとしてクラス  $C_i$  に属する  $N_i$  個の CHLAC 特徴ベクトル  $x_{i,j} \in C_i (1 \leq j \leq N_i)$  を使用し，その分散共分散行列  $R_i$  を算出する．

$$\bar{x}_i = \frac{1}{N_i} \sum_{j=1}^{N_i} x_{i,j} \quad (2)$$

$$R_i = \frac{1}{N_i} \sum_{j=1}^{N_i} (x_{i,j} - \bar{x}_i)^t (x_{i,j} - \bar{x}_i) \quad (3)$$

この分散共分散行列  $R_i$  の固有値  $\lambda_{i,1} > \lambda_{i,2} > \dots > \lambda_{i,251}$  とそれに対応する固有ベクトル  $u_{i,1}, \dots, u_{i,251}$  の一部を用いて部分空間  $L_i$  を構成する．

そのために，まず構成する部分空間の次元  $d_i (1 \leq d_i \leq 251)$  を設定する．設定には (4) 式に示す次元  $d_i$  における累積寄与率  $a(d_i)$  を用いる．

$$a(d_i) = \frac{\sum_{j=1}^{d_i} \lambda_{i,j}}{\sum_{j=1}^{251} \lambda_{i,j}} \quad (4)$$

この累積寄与率  $a(d_i)$  がしきい値  $\kappa_d$  に対して条件

$$a(d_i) < \kappa_d < a(d_i + 1) \quad (5)$$

を満たすような  $d_i$  を選択する．これによって決定された次元  $d_i$  により，影響の大きい固有ベクトルから順に  $d_i$  個選択して，部分空間  $L_i$  の基底は  $\langle u_{i,1}, \dots, u_{i,d_i} \rangle$  となり，部分空間  $L_i$  からの射影行列  $P_i$  を (6) 式により定める．

$$P_i = [u_{i,1}, \dots, u_{i,d_i}] \quad (6)$$

#### 4.2 直交化

次に，各クラスを表現する部分空間の違いを明白にするため，直交部分空間法 [14] を使用する．

クラス  $C_i$  の部分空間  $L_i$  の基底を  $u_{i,1}, \dots, u_{i,d_i}$ ，クラス  $C_j$  の部分空間  $L_j$  の基底を  $u_{j,1}, \dots, u_{j,d_j}$  とする．

直交部分空間法では，2つのクラスの部分空間から選んだ任意の基底を直交するように構成する．つまり，部分空間  $L_i$  を直交化した直交部分空間  $M_i$  (基底を  $v_{i,1}, \dots, v_{i,d_i}$  とする)，および部分空間  $L_j$  を直交化した直交部分空間  $M_j$  (基底を  $v_{j,1}, \dots, v_{j,d_j}$  とする) において，任意の  $i, j, k, l$  について

$$v_{i,k}^t v_{j,l} = \delta_{k,l} \delta_{j,l} \quad (7)$$

が成立するように構成するということである．具体的には，4.1 節で求めた部分空間  $L_i$  の射影行列  $P_i$  に対して，直交化行列  $O$  による射影を行うが，実際の直交化においては  $O$  の代わりに擬似直交化行列の  $Q$  を用いる．

#### 4.3 識別

本節では，クエリベクトル  $x_q$  を識別する方法について説明する．

元の特徴空間からクラス  $i$  の直交化部分空間  $M_i$  への変換行列  $Q_i^t$  を用いてクエリベクトル  $x_q$  を射影すると

$$x_{q,i} = Q_i^t (x_q - \bar{x}_i) \quad (8)$$

となる，この直交化部分空間  $M_i$  内で表したベクトルを元の特徴空間で表し直したものと，元の特徴ベクトルとのベクトル差は直交化部分空間に射影したことによる損失にあたり，これを  $err_i(x_q)$  と定義する．

$$err_i(x_q) \triangleq |x_q - \bar{x}_i - Q_i x_{q,i}| \quad (9)$$

ここで得られた情報の損失は直交化部分空間  $M_i$  による情報圧縮による損失である．よって，この損失に対して閾値  $\kappa_{e,i}$  を設定し，それによるクエリベクトルのクラス  $i$  に対する判定を行う．

クエリベクトル  $x_q$  のクラス  $C_i$  での情報の損失  $err_i(x_q)$  が以下の条件を満たすとき，そのクエリベクトル  $x_q$  はクラス  $C_i$  に属していると判定する．

$$x_q \in C_i \quad (\forall i : \kappa_{e,i} > err_i(x_q)) \quad (10)$$

全てのクラスへの判定は平行して行われるので，ひとつのクエリベクトルは複数クラスに分類されうる．また，どのクラスにも分類されないという状況も存在しうる．

#### 5. 多クラス識別器の構成

前節までで説明した手法を実際の動作識別に適用する方法について述べる．

クラス (動作) は数種類とする．一つのクラスはある一定時間の動作試行を複数回含む．また，クラスではよく似ているがまったく同一でもない動作が含まれるものとする．

カメラで撮影された映像はフレーム間差分される．時系列で最新である3枚のフレーム間差分画像をしきい値  $\tau_I$  を用いて二値化したものに CHLAC を適用し，251次元のベクトルを算出する．

さらに，学習データから直交部分空間を構成し，得られた投影行列を用いて 4.3 節のとおりに識別を行う．

このとき，ある動作の映像シーケンスを構成するフレームから算出された複数のクエリベクトルをクエリベクトル集合  $G_q$  とする． $G_q$  に属するすべてのベクトルについて識別を行い，結果とする．

なお，(10) 式において必要とされる，それぞれのクラスに識別する閾値  $\kappa_{x,i}$  は，以下のように決定する．

まず，クラス  $C_i$  の  $N_i$  個の学習データ自体をそのクラスの識別器で識別し，その損失の大きさにより，降順に並べる．

$$err_{i,1} > err_{i,2} > \dots > err_{i,N_i-1} > err_{i,N_i} \quad (11)$$

その上で、クラス  $C_i$  の学習データの個数  $N_i$  に対する割合  $r_i$  により、閾値  $\kappa_{x,i}$  を決定する。

$$\kappa_{x,i} = \text{err}_{i,j}(j = r_i \cdot N_i) \quad (12)$$

この、 $r_i$  は別途定めるものとする。

## 6. 実験

### 6.1 実験方法

実験として実際に動物体の動作を分類する。ここでは、AXIS214 カメラを監視カメラとして使用した。カメラから得られた映像は  $640 \times 480$  の VGA サイズである。

実験で使用した動作は、車の前後移動・自転車での移動・人間の歩行・拾得動作・ドアの開閉である。それぞれの動作は、同じ画角にセットされた同じカメラから撮影された。撮影は学習データとして 10 回、クエリデータとして 1 回の動作について行った。それぞれの動作を図 1 に示す。すべての映像は CHLAC を算出する前処理として、 $1/3$  に縮小した。CHLAC を実際に参照する 3 つのフレーム間差分画像の例を図 2 に示す。

撮影した映像は、グレイスケール 256 階調のフレーム間差分画像を閾値  $\tau_l = 15$  を用いて 2 値化し、CHLAC 特徴量を抽出し、累積寄与率が閾値  $\kappa_d = 0.99$  に対して条件 (式 5) を満たすように次元を設定し、部分空間を構成した。クエリデータは別途撮影したそれぞれの動作を使用した。こちらにも、CHLAC 特徴量を映像から生成した。5. 節で記述した方法によりクエリデータを分類させた。その際に、閾値を決定するパラメタ  $r_i$  は全てのクラス  $i$  について  $r_i = 0.7$  とした。

### 6.2 実験結果

結果を表 1 に示す。上段に並んでいるクエリが左列に示されているクラスに何フレーム分類されたかを示す。クラスは自動車の前進・後進および左方向および右方向への歩行・自転車での移動、拾得動作、ドア開閉の 9 クラスであるが、「A+B」とある項目はクラスへの分類がそれぞれ平行して行われたため、A と B という複数クラスに分類されたことを示す。対角の項目が分類に成功した項目である。

クエリ集合を構成する CHLAC 特徴量ベクトルについて、29 ~ 100% の割合でクエリを正解クラスに分類することができた。成功率のすべての動作について平均をとると、82.9% である。一方、最悪値の 29% という結果になったのは、車の後進である。また、拾得動作もほかの動作に分類されることがあった。前者は正解クラス  $i$  に射影誤差に対する閾値  $\kappa_{e,i}$  が低すぎることで、後者は非正解クラス  $j$  の射影誤差に対する閾値  $\kappa_{e,j}$  が閾値が高すぎることで生じており、 $r_i$  による決定 (式 10) に改善の余地があることを示していると考えられる。

CHLAC は対象の大きさ変化に依存するので、映像中で対象が多少大小することに対しては、スケーリング等

で対応 [7] することが求められる。また、対象が異なる速度の場合、異なる CHLAC 特徴となるため、対応 [11] していく必要があるであろう。

## 7. おわりに

本研究では、監視目的で設置された固定カメラの映像による、多種類の物体の動作を識別する手法を提案した。

本手法では、映像からフレーム間差分し得られた 2 値画像系列から CHLAC 特徴量を算出し、識別に用いる。

方法としては、まず事前に用意した複数物体の複数動作それぞれについての学習データについて、部分空間法の一種である直交部分空間法により部分空間を構成した。その上で、入力 CHLAC 特徴量に対して、クラスごとに情報の損失の閾値を設定し、独立した識別を行える識別器を構成した。

実験では、歩行者、自転車、自動車の 3 種類の物体による合計 9 種類の動作に対して実験を行い、82.9% の識別結果を得た。これは識別器にまだ改良の余地が多く含まれていることを考えると、将来性のある手法であると考えられる。

今後の課題として、スケーリングや速度変化への対応が考えられる。また、岩村ら [15] により考察された部分空間法の原点にまつわる問題も考慮すべきであろう。

また、監視目的での研究評価においては、TRECVID [17] のような取り組みもあるため、将来はこのような枠組みの中での評価も行っていきたい。

## 文 献

- [1] Nobuyuki Otsu, "Towards Flexible and Intelligent Vision Systems - From Thresholding to CHLAC -", IAPR Conference on Machine Vision Applications, pp.430-439, 2005.
- [2] Juan Carlos Nieves, Hongcheng Wang, and Li Fei-Fei, "Unsupervised Learning of Human Action Categories Using Spatial-Temporal Words", International Journal of Computer Vision, pp. 299-318, 2008.
- [3] 村井泰裕, 藤吉弘巨, 金出武雄, "Space-Time Patch を用いた物体の移動方向識別とセグメンテーション" MIRU 講演論文集, pp.389-396, 2007.
- [4] 木谷クリス真実, 岡部孝弘, 佐藤洋一, 杉本晃宏, "視覚的文脈を考慮した人物動作カテゴリの教師無し学習" MIRU 講演論文集, pp.28-33, 2008.
- [5] 南里卓也, 大津展之, "複数人動画像からの異常動作検出", CVIM, vol.2004, no.91, pp.173-186, 2004.
- [6] 村井泰裕, 藤吉弘巨, 数井誠人 "時空間特徴に基づくエスカレーターシーンにおける人の異常行動検知" 電子情報通信学会 パターン認識・メディア理解研究会 (PRMU), pp.247-254, 2008.
- [7] 白木 孝義, 石黒 勝彦, 深野 亮, 鴨志田 良和, 白井 達也, 斎藤 秀雄, 田浦 健次郎, 大武 美保, 佐藤 知正, 大津 展之, "CHLAC 特徴と Grid コンピューティングを併用したリアルタイム動作認識" 信学技報 PRMU vol.105, pp.97-102, 2006.
- [8] D. G. Lowe, "Object recognition from local scale invariant feature", Proc. of IEEE International Conference on Computer Vision (ICCV), pp. 1150-1157, 1999.
- [9] Eli Skechtman, Michal Erani, "Space-time behavior

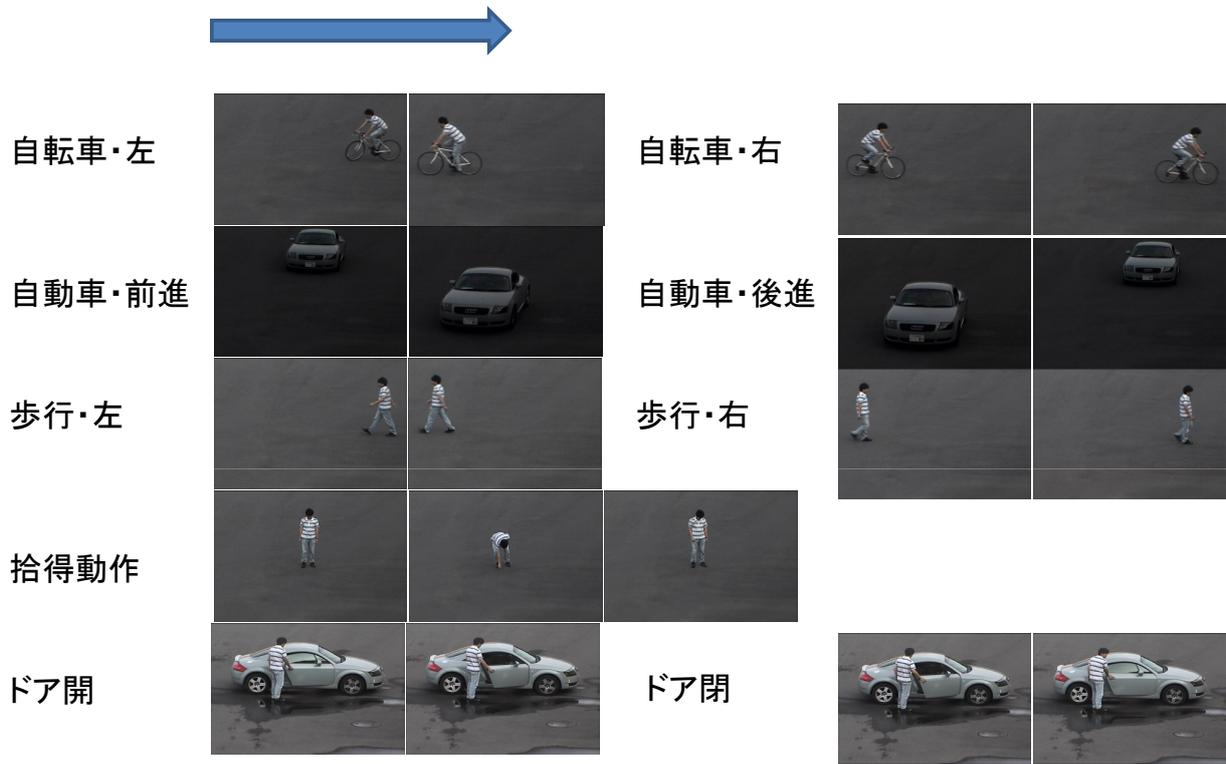


図 1 実験に用いた動作の例

表 1 動作分類結果 (単位:フレーム数・括弧内は全体に対する割合の百分率)

		クエリ								
		自転車左	自転車右	車前進	車後進	ドア開き	ドア閉め	拾得	歩行左	歩行右
	全体	58	60	92	124	15	22	29	92	105
結果	無分類	0	0	0	88 (70.97)	0	9 (40.91)	3 (10.34)	0	0
	自転車左	58 (100.00)	0	0	0	0	0	0	0	0
	自転車右	0	60 (100.00)	0	0	0	0	0	0	0
	車前進	0	0	92 (100.00)	0	0	0	0	0	0
	車後進	0	0	0	36 (29.03)	0	0	0	0	0
	ドア開き	0	0	0	0	15 (100.00)	0	0	0	0
	ドア閉め	0	0	0	0	0	13 (59.09)	0	0	0
	拾得	0	0	0	0	0	0	19 (65.52)	0	0
	歩行左	0	0	0	0	0	0	1 (3.45)	85 (92.39)	0
	歩行右	0	0	0	0	0	0	0	0	105 (100.00)
	拾得+歩行左	0	0	0	0	0	0	6 (20.69)	0	0
	自転車左+歩行左	0	0	0	0	0	0	0	7 (7.61)	0

[10] based correlation”, Proc. CVPR, pp.405-412, 2005.  
N.Otsu and T.Kurita, “A new scheme for practical,

flexible and intelligent vision systems”, Proc. IAPR Workshop on ComputerVision, pp.431-435, 1988.



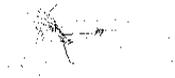
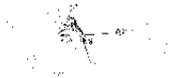
自転車			
自動車			
歩行			
拾得動作			
ドア			

図 2 CHLAC 算出の例

- [11] 森下雄介, 小林匠, 森崎巧一, 大津展之, “時間重みと外的規準を用いた動作評価手法” CVIM, vol.2008, no.27, pp.47-52, 2008.
- [12] Rumerlhar D.E., Hilton G.E. and Williams R.J., “Learning representation by back-propagating errors”, Nature, 323, pp. 533-536, 1986
- [13] C.J.C.Burges, “A tutorial on support vector machines for pattern recognition”, Data Mining and Knowledge Discovery 2, pp.121-167, 1998
- [14] 石井健一郎, 上田修功, 前田英作, 村瀬洋, “わかりやすいパターン認識”, オーム社出版局, 1998.
- [15] 岩村雅一, 大野真一郎, 阿曾弘具, “部分空間法における原点の位置と認識性能に関する考察”, 部分空間研究会, Subspace2006, 2006.
- [16] 小林匠, 大津展之, “パターン識別のための錐制約部分空間法”, 電子情報通信学会論文誌.D, 情報・システム, pp.104-111, 2009.
- [17] Alan F. Smeaton, Paul Over, and Wessel Kraaij, “High-Level Feature Detection from Video in TRECVID: a 5-Year Retrospective of Achievements”, Multimedia Content Analysis, Theory and Applications (Ajay Divakaran Ed.), pp.151-174, Springer, 2009.