機械学習を用いたATM異常行動検知技術の検討

ATM Anomaly Behavior Detection with Machine Learning: A Study

ファンチョンフィ[*1] Phan Trong Huy　岸礼子[*1] Reiko Kishi　山本一真[*1] Kazuma Yamamoto　増田誠[*1] Makoto Masuda

沖電気工業株式会社 経営基盤本部 研究開発センター
Corporate Research and Development Center, Oki Electric Industry Co., Ltd.

In recent years, a steady increase in ATM (Automatic Teller Machine)-related crimes has been reported overseas. One of which is ATM skimming; the act of installing skimming devices (a.k.a. skimmers) to ATM to illegally copy information from the magnetic stripes of cash cards, credit cards, etc. As skimmers grow smaller and more sophisticated, detecting such devices with conventional sensors is facing great difficulties. With the purpose of strengthening ATM security, we are developing image sensing technologies that detect anomaly behaviors including ATM skimming acts using video feeds capturing the ATM operational area. Machine learning is employed to represent normal behaviors; the degrees of separation from such representation can be used as an indicator for abnormality level. In this article, we discuss the application of well-known methods (Subspace Representation of [Nanri 2004] and Gaussian Mixture Model of [Yu 2006]) to modelling ATM normal behaviors in order to detect ATM anomaly behaviors. Additionally, we also brief several considerations to realize high anomaly detection accuracy in real practice.

Keywords: ATM, Anomaly Detection, Machine Learning

1. はじめに

近年、海外ではATM（Automatic Teller Machine）に対する犯罪の増加が報告されている。犯罪として、キャッシュカードやクレジットカードの磁気記録情報を「スキミングカードリーダー」と呼ばれるスキミング装置を用いて不正に読み取って偽造カードを作成するという行為がある。スキミング装置は小型化する傾向があるため、従来のセンサーによる検知は難しくなってきている。そこで、我々はATMセキュリティの強化を目的に、ATM操作部の映像からスキミング装置取り付け動作などの異常行動の検知を実現するための技術開発を行っている。

ATM 異常行動検知においては、機械学習を用いて正常動作を集約するモデルを生成し、そのモデルからの分離度を基に異常検知を行う。本稿では、代表的な正常モデルの生成方法（[南里 2004]の部分空間法と[Yu 2016]の混合正規分布モデル）を ATM 操作部の映像に適用し、異常の検知ができることについて紹介する。今後、実運用を想定した取り組みとして ATM 取引情報の活用及び追加学習の実現を行っていく。

2. 異常検知技術の概要

機械学習を用いた異常検知は基本的にデータの特徴抽出処理と識別処理の 2 つの工程から成り立つ。

コンピュータビジョン分野では、映像中の動きを解析するための特徴抽出方法が多く存在する。例として、CHLAC 特徴[Kobayashi 2004]・HOOF 特徴[Chaudhry 2009]・MBH 特徴[Dalal 2006]などがある。本稿では、ATM 操作でみられる細かな手の動作を表現するために、ST-Patch（Space-Time Patch）特徴[Murai 2007]を用いる。式(1)に表すように、ST-Patch 特徴は時系列画像の勾配情報を用い、動作の“見え”及び“動き”の変化を表現する特徴である。

$$F = \{ \sum P_x^2, \sum P_x P_y, \sum P_x P_t, \sum P_y^2, \sum P_y P_t, \sum P_t^2 \} \quad (1)$$

ここで、P_u は u 軸での勾配である。（x、y、t）は時系列画像の（横、縦、時間）軸を表す。

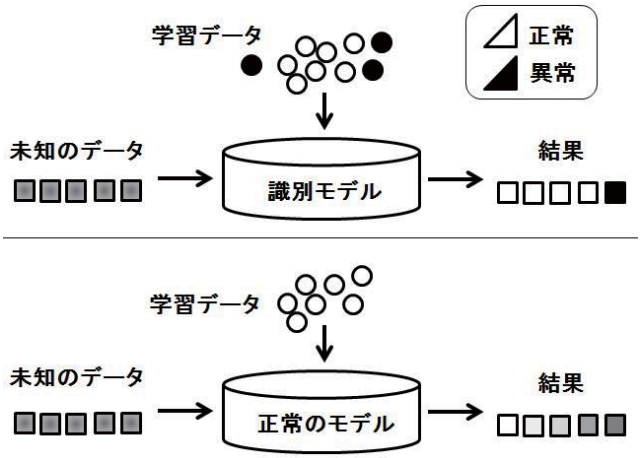

図 1：異常検知のアプローチ
（上）教師あり学習；（下）教師なし学習

識別処理として、教師あり学習（Supervised）・教師なし学習（Unsupervised）といった 2 つのアプローチが存在する[Omar 2013]（図 1）。教師あり学習では、正常データと異常データを別のクラスとして学習し、未知のデータを正常か異常かへクラス分類を行う。教師なし学習では、正常データを用いて正常を表現するモデルを学習する。正常モデルからの分離度を異常度として扱う。

連絡先：ファンチョンフィ、沖電気工業株式会社
経営基盤本部 研究開発センター、埼玉県蕨市中央 1−16−8、
メール：phan586@oki.com

本稿では、教師なし学習のアプローチを選定した。その理由は、ATM 異常行動検知において、スキミング行為以外の未知の異常行動も検知できるためである。我々は、動画での異常検知に実績のある正常モデルの学習手法として、[南里 2004]の部分空間法と[Yu 2016]の混合正規分布を ATM 異常行動検知という問題に適用することを提案する。

2.1 部分空間法

部分空間法はクラス分類手法の一つである。部分空間法を用い正常動作の部分空間を構成し、その部分空間からの距離を異常度として扱うことができる(図 2)。[南里 2004]のように、まず、学習対象の正常データ(*N*次元)を用い主成分分析(PCA)を行う。その結果である固有ベクトル群(主成分) $U=[u_1,..,u_N]$のうち、上位の k 固有ベクトルを使って、正常動作の部分空間 $U_k=[u_1,..,u_k]$を構成することができる(※k はハイパーパラメータである)。正常部分空間からの未知データ xの距離 dは式(2)のように求めることができる。

$$d^2 = x^Tx - x^TU_kU_k^Tx \quad (2)$$

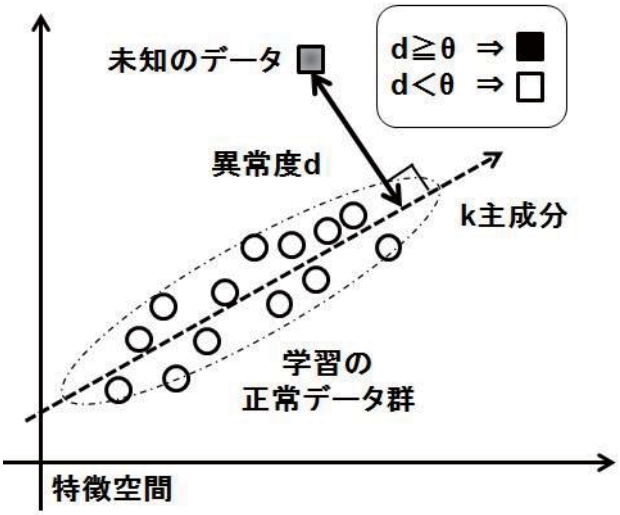

図 2:部分空間法による異常検知

2.2 混合正規分布モデル

混合正規分布モデル(Gaussian Mixture Model)は複数の正規分布の重みづけ和を用いたデータを確率密度分布として表現する手法である。[Yu 2016]のように EM 法(Expectation Maximization)を用いることで、正常データを表す混合正規分布を求めることができる。未知データ xの正常としての確率 $P(x)$は式(3)のように算出することができ、この値を用いて異常判定を行うことができる。

$$P(x) = \sum_{i=1}^{m} w_i p_i(x|\mu_i,\sum_i) \quad (3)$$

ここで、$p_i(x|\mu_i,\sum_i)$は平均 μ_i 分散$\sum_i$を持つ正規分布を表す。w_iは正規分布 iの重みである。ハイパーパラメータ m は正規分布の数を示す。

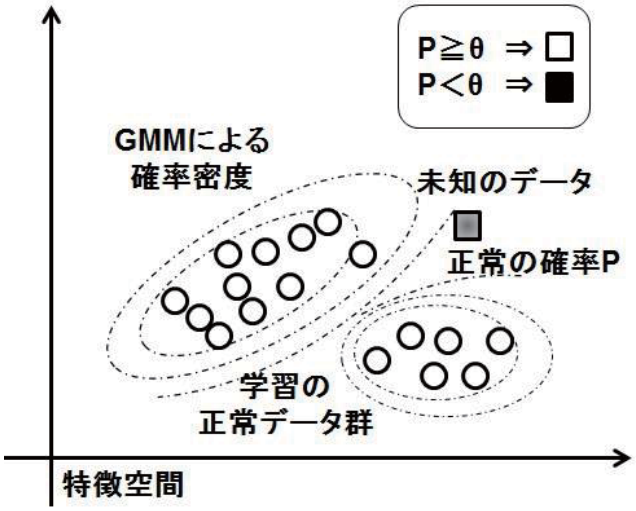

図 3:混合正規分布による異常検知

3. ATM 異常行動の映像データを用いた検証

3.1 検証データ

ATM 異常行動検知への適用可否を検証するために、正常 205 動作(学習用 175 動作＋検証用 30 動作)および異常 24 動作を用いる。本稿では、図 4 のように ATM の上部にカメラを設置し ATM 取引の動作を撮影する。また、カード挿入口付近の領域に絞って検証を行う。

●正常動作として、通常の引きだし・預け入れが含まれた複数パターンのデータを用いる。

●異常動作として、スキミング装置取り付けの複数パターンのデータを用いる。

3.2 検証の結果

表1は手法ごとの AUC(Area Under the Curve)・EER(Equal Error Rate)といった指標[Davis 2006]で評価した結果を表す。AUC は異常検知の全体の良さを表す数値であり、EER は正常を誤検知するエラー率が異常を未検知するエラー率が等しい時のエラー率を示す。

本稿では、部分空間法及び混合正規分布モデルのハイパーパラメータを複数試し、学習・評価を行うことにより最適なハイパーパラメータを選定する。評価結果により、部分空間法と混合正規分布モデルは同等の精度で異常の検知ができることがわかる。それぞれ約 13.5%の正常動作が誤検知されたときに、約 86.5%(=100%-13.5%)のスキミング装置取り付け動作を異常として検知することができる。

表 1:評価結果

指標	部分空間法	混合正規分布モデル
AUC (高いほうが良い)	95.0%	94.4%
EER (低いほうが良い)	13.5%	13.6%

3.3 考察

いずれの手法においても比較的高い異常検知精度となり、実運用への活用が期待できる。しかし、撮影時に想定していない正常の動作パターンは誤検知となる共通の傾向がみられる。例として、柄の財布や紙類を持ちながら ATM 取引を行うという動作などの異常値が高く誤検知されやすい。要因としては、学習対象のデータにこのような動作パターンの割合が非常に低いことが考えられる。全体と比べ割合の少ないパターンに対しては、学習により求める正常モデルに反映されず、評価時の異常値が高くなってしまうためである。

混合正規分布モデルは複数の正規分布で構成されるため、部分空間法よりも多様性のある正常動作を表現することができる。ただし、今回の正常データにはパターンが少ない特異な正常行動が含まれていたため、部分空間法および混合正規分布の手法の得意・不得意に大きな差が見られなかった。しかし、様々なバリエーションに対して大量にデータがある場合は混合正規分布モデルのほうが効果的だと考えられる。さらに、混合正規分布モデルにより、部分空間への変換処理を省略しつつ、比較的高い検知精度が実現できることを確認した。今後、実運用に向けては混合正規分布モデルの検討を進めていく予定である。

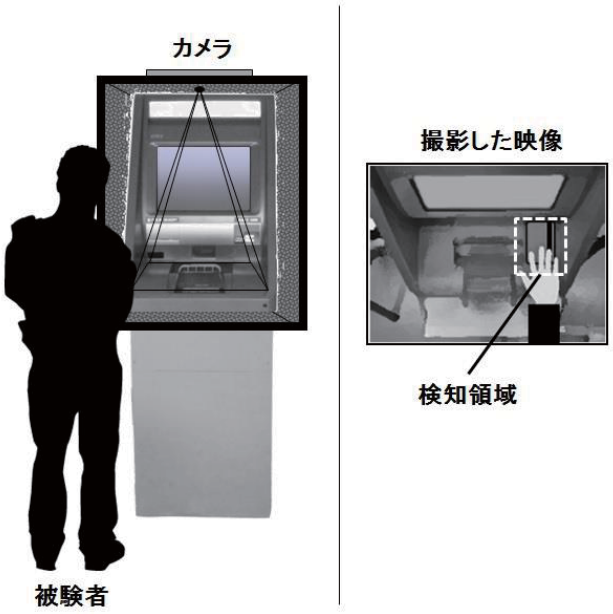

図 4:(左)撮影時のシステム構成;(右)映った操作部の映像

4. 実運用に向けて

4.1 ATM 取引情報の活用

ATM 操作において操作の位置や取引の内容によって起こり得る動作が全く異なるため、単体の正常モデルでの異常検知は困難である。そこで、我々は検知領域を複数設けることと共に、ATM 取引内容に合わせた正常モデルを複数学習することで高い検知精度を実現しようとしている。

4.2 追加学習の実現について

実運用では、設置現場の環境(照明や ATM の外観など)と学習時の環境が異なる状況が生じる。そのため、事前に学習した正常モデルは設置現場で起こる正常動作が表現できなくなり、誤検知につながると考えられる。そこで、我々は設置現場のデータを追加して学習する技術の導入を検討している。

●部分空間法(2.1 節)の場合:CCIPCA[Weng 2003]や IPCA[Skocaj 2003]のアルゴリズムを用いることで、正常動作を構成する主成分を更新することが可能である。

●混合正規分布(2.2 節)の場合:[Engel 2010]や[島田 2007]が提案した手法により、混合正規分布モデルのパラメータ(正規分布の数・重み・平均・分散)を更新し、追加のデータに適応させることができる。

5. おわりに

我々は ATM セキュリティの強化を目的に、ATM 操作部の映像からスキミング装置取り付け動作の検知(異常検知)を実現するための技術開発を行っている。本稿では、機械学習を ATM 操作部の映像に適用することにより、正常と異常の境界線があいまいである端末の操作などに対し、異常行動が検知できることを示した。今後、実運用に向け、ATM 取引情報の活用や追加学習の導入を検討していく予定である。

参考文献

[Weng 2003] Juyang Weng et al., “Candid covariance-free incremental principal component analysis”. PAMI 2003.

[Skocaj 2003] D. Skocaj and A. Leonardis, “Weighted and robust incremental method for subspace learning”. ICCV 2003.

[南里 2004] 南里卓也, 大津展之, "複数人動画像からの異常動作検出" , 電子情報通信学会技術研究報告. PRMU 2004.

[Kobayashi 2004] T.Kobayashi and N.Otsu, “Action and Simultaneous Multiple-Person Identification Using Cubic Higher-Order Local Auto-Correlation”. ICPR 2004.

［Davis 2006］ Jesse Davis et al., “The relationship between Precision-Recall and ROC curves”. ICML 2006.

[Dalal 2006] Navneet Dalal et al., “Human Detection Using Oriented Histogram of Flow and Appearance”. ECCV 2006.

[Murai 2007] Yasuhiro Murai et al., “Combined Object Detection and Segmentation by Using Space-Time Patches”. ACCV 2007.

[島田 2007] 島田敦士、有田大作、“適応的な分布数の増減法を利用した混合ガウス分布による高速な動的背景モデル構築”.信学論 2007

[Chaudhry 2009] Rizwan Chaudhry et al., “Histograms of Oriented Optical Flow and Binet-Cauchy Kernels on Nonlinear Dynamical Systems for the Recognition of Human Actions”. CVPR 2009.

[Omar 2013] Salima Omar, “Machine Learning Techniques for Anomaly Detection: An Overview”. IJCA 2013.

[Yu 2016] Hao Yu et al., “Video Anomaly Detection Based on Mixed Statistic Feature”. International Journal of Signal Processing Systems 2016.