

# 計算機と人間との協調作業による事例データベース構築の省力化

## Reduction of Workload of Human Labeling using Computer and Human Interaction

村田 博士\*<sup>1</sup>      山名 美智子\*<sup>2</sup>      小野田 崇\*<sup>1</sup>  
 Hiroshi Murata      Michiko Yamana      Takashi Onoda

\*<sup>1</sup>(財) 電力中央研究所      \*<sup>2</sup>(独) 産業技術総合研究所  
 Central Research Institute of Electric Power Industry      Advanced Industrial Science and Technology

A large sample database which consists of measurement data and humans' judgment information is essential element to realize the *Condition Based Maintenance* (CBM). Against such a background, we propose a efficient method for creating and maintaining database. A sample database are created and maintained by cooperative work between human and machine in this method.

This report presents a reduction method of workload of human labeling keeping high accuracy of automatic labeling. We applied our method to labeling handwritten images for digit recognition and labeling rust images for crossarm reuse judgment. In the result, this method is able to reduce the workload of human about labeling rust images for crossarm reuse judgment by 70 ~ 89 % without labeling error of machine.

### 1. はじめに

設備産業である電力会社では、電力自由化の中、設備・機器の再利用による有効活用、異常予兆発見に基づく設備・機器の交換などによるコストダウンの実施を行っている。特に、設備・機器の保守・管理を従来の使用年数に基づく時間基準保全 (Time Based Maintenance (TBM)) から、設備・機器の劣化状態等に基づく状態監視保全 (Condition Based Maintenance (CBM)) への移行を、近年、図り始めている。

CBM では、設備・機器があらかじめ定めた性能劣化基準範囲内での設備・機器の運用が可能となるため、TBM より、大幅なコスト削減が期待できる。しかし、CBM によって、コスト削減効果を高めるには、(1) 設備・機器の様々な状態に関するデータを蓄積していること、および、(2) そのデータに対する専門家のノウハウを蓄積していること、の 2 つの条件が重要となる。

現在、電力会社では、デジタルカメラ、ビデオ等の IT 関連機器を利用して、設備状態を表した画像データなどのデータの収集を開始している。しかし、専門家の知見により付与される、対象機器の定性的状態を示す情報 (ラベル) がないと、そのデータの有効利用は困難である。

最近の情報技術の発展により、デジタル画像データの撮影は、精度良く、安価に、簡単に行えるようになってきており、大量の設備・機器画像データを収集できる環境が整ってきた。しかし、設備・機器の画像データは簡単に収集できるものの、収集されたデータに対し、専門家が自分の知見に基づき、ラベルを付与する「ラベル付け」には、膨大な労力が必要となる。図 1 に、事例データベースに基づく、画像自動判定のイメージを示す。図 1 における自動判定の精度は、ラベル付けされた画像データからなる事例データベースの善し悪しに大きく左右される。しかし、何千、何万という画像データを専門家がラベル付けするのは不可能ではないにしろ、長時間の集中力を要する困難な作業となる。

実際、腕金錆画像に基づく再利用判定システムに関する我々の研究において、600~700 枚の腕金錆画像データに対し、専門家にラベル付けを行ってもらったところ、ラベル付けに一週

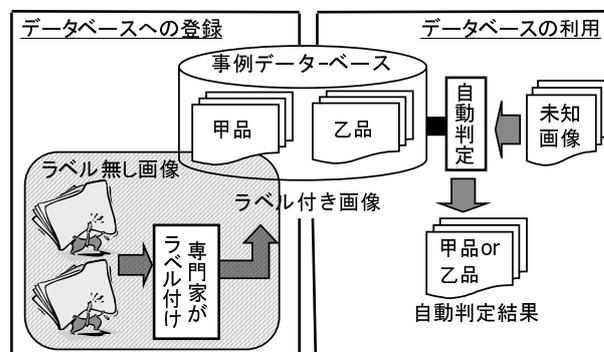


図 1: 事例データベースに基づく画像自動判定のイメージ

間を要した。このラベル付け作業は、長い時間の集中を要する大変な作業である。また、長時間にわたる作業であるため、作業途中でラベル付けのための判定基準が振れる傾向が見られた [1]。

このような背景から、本報告では、腕金錆画像などの事例データベース構築時に専門家による少ないラベル付け回数で、精度の高い判定を行える方法を提案する。

### 2. 提案するデータベース構築手法

本章で、事例データベース構築作業 (図 1 灰色縞部分) を計算機と人間との協調作業とすることにより、省力化をはかる方法を提案する。この方法は次のような手順からなる。

1. 図 2 に示すように、計算機は大量のラベル無しデータから少量のデータだけを提示し、専門家がラベル付けを行う。
2. そのラベル付けされたデータを用いて、計算機はサポートベクターマシン (SVM) [2] を用いて学習し、残りの大部分のデータへのラベル付けを行う。
3. その後、計算機ではラベル付けの難しいデータを自動選択し、専門家に提示し、専門家がラベル付けを行う。

上記手順を計算機ではラベル付けの難しいデータがなくなるまで繰り返すことで、専門家が関わるラベル付け作業は少量で、計算機が大量のラベル付けを行うことができる。

A: (財) 電力中央研究所 システム技術研究所, 〒 201-8511 東京都狛江市岩戸北 2-11-1, TEL:03-3480-2111, FAX:03-5497-0318, murata@criepi.denken.or.jp

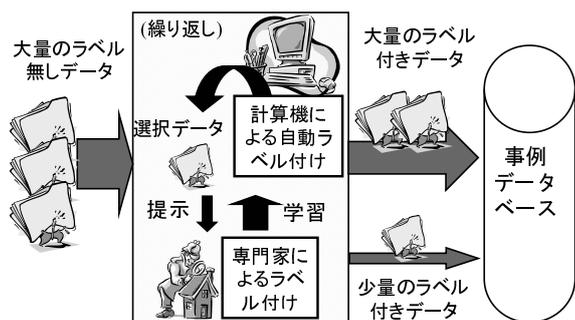


図 2: 計算機と人間との協調作業によるラベル付けシステム

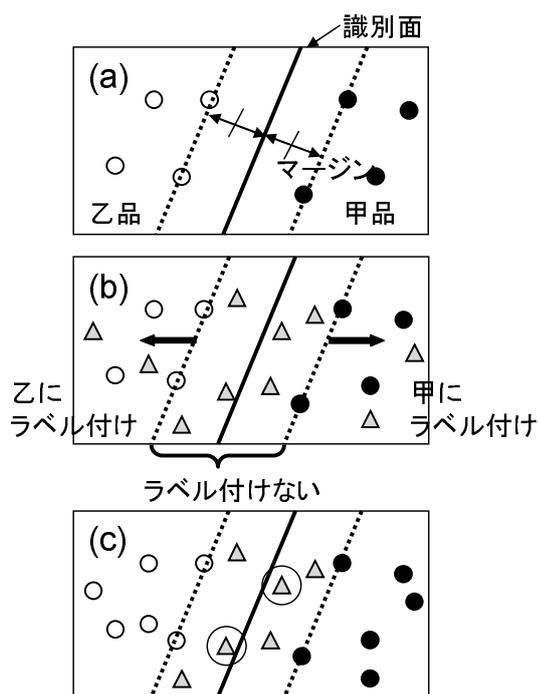


図 3: サポートベクターマシンの概念図とラベル付け

以上の計算機と人間との協調作業による事例データベース構築方法は、専門家による少量のラベル付けされたデータを計算機が学習し、そのラベル付け基準を事例データベース全体に反映させる。従って、データベース構築後に設備・機器の状態判定基準を変更する場合にも、新たに少量のデータに専門家がラベル付けするだけで事例データベース全体を変更できる利点もある。

ここで例として、電柱部品である腕金の表面錆画像判定事例データベース構築を考える。SVM では、各画像は、色を表す RGB 画像の画素値などの画像の特性値を要素とするベクトル空間上 (特徴ベクトル空間) の、ある一点として表される。このとき、特徴ベクトル空間上で甲品と乙品のクラスを分ける面が識別面となる (図 3(a): 図中の黒丸が甲品の画像、白丸を乙品の画像とする)。ラベル付けがなされたデータに基づいて識別面を決めることを学習と呼ぶ。SVM では、識別面からラベルの付いているデータまでの最短距離をマージンと呼び、識別面はマージンが最大となるように学習により定める。マージン内は、与えられたデータだけではどちらのクラスに属するか不明な領域 (マージン領域) である。ここでさらに、専門家がラ

ベル付けしていない、複数の新たな画像データを追加する (判定する) ものとする (図 3(b): 図中の三角印をまだラベル付けしていない画像とする)。二つのクラスの境界に位置するマージン領域のデータにはラベル付けに注意を要する。なぜなら、マージン領域のデータは僅かな違いでどちらのクラスにも属する可能性があり、ラベル付けが難しい画像データだからである。従って、甲品領域に入った新しい画像データには、甲品と、乙品領域に入った新しい画像データには、乙品とラベルを付け、マージン領域の新たな画像データはラベル付け不可能とする (図 3(c))。そして、マージン領域の特に識別面上に近い画像データを専門家に提示し、ラベル付けを促すのが効果的であると考えられる。

SVM を用いた腕金錆画像に基づく判定を行った我々の研究結果から、RBF カーネル  $K(x, x') = \exp(-\gamma \|x - x'\|^2)$  を用いた判定精度が最も高くなるのがわかっている [1, 3]。一般的に、SVM の識別能力を決めるパラメータ  $\gamma$  を大きくすると、よく似たデータにも異なるラベルを与えることが可能になるが、学習に必要なデータ数は多くなる [4]。

そこで、必要とされるラベル付けの精度を維持しながら、人の作業量をできるだけ少なくするように、パラメータを適切に制御する方法について検討を行った。

### 3. 自動パラメータ調節方法

本章では、計算機が自動でパラメータを調節する、効率的なラベル付け方法について述べる。

#### 3.1 パラメータ調節の考え方

識別器の性能を制御するパラメータ  $\gamma$  は小さいほど、データ特徴ベクトル空間上での差異を区別しないため、計算機は少ない学習で、多数のデータのラベルを付けられるが、ラベルの誤りも多い。従って、必要な識別能力を有するだけ小さい  $\gamma$  を求める必要がある。この目的のために、 $\gamma$  の初期値から逐次改良して、最適な値を求めるアルゴリズムを提案する。以下に詳細を述べる。

#### 3.2 パラメータ初期値の決定方針

計算機はまず、人間がラベルを付けた少数の初期のデータセットから  $\gamma$  の初期値を決める。パラメータの初期値決定に先立ち、 $\gamma$  が満たすべき条件について考える。

条件 1. (計算機によるラベル付けができる条件)

ラベル無しのデータを判定すると、計算機がラベルを付けられるデータ (マージン領域より外に出るデータ) が最低でも一つはある (図 4)。

条件 2. (学習済みデータは正しくラベル付けできる条件)

学習したデータは正しくラベル付けできる。すなわち、全ての学習済みデータはマージン領域の外に存在し、ラベルの誤りはない。

条件 1 は、計算機によるラベル付けが行われるようにするための条件である。 $\gamma$  が大きくなりすぎると、学習したデータに関しては全て、ラベル付け結果が学習データと一致する。その代わりに、ラベル付けしていないデータは全て、マージン領域に含まれるため、計算機はラベルを付けられなくなる。従って、条件 1 は  $\gamma$  の初期値の上限を制約する (図 5)。一方、条件 2 は人間がラベル付けした結果を正しく学習するために課す条件である。 $\gamma$  が小さくなり過ぎると、特徴ベクトル空間上での微細な差異が区別できなくなり、学習したデータに関しても正しいラベル付けができなくなる。従って、条件 2 は  $\gamma$  の初期値の下限を制約する (図 5)。

パラメータの初期値は、誤ラベル数ができるだけ小さくなるように、計算機がラベルを付ける数が最小 (但し、1 個以上) と

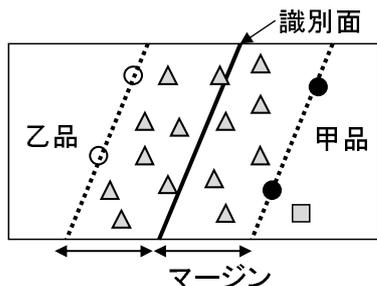


図 4: 条件 1 を満たす状態例 (四角印は計算機がラベルを付けられるデータ)

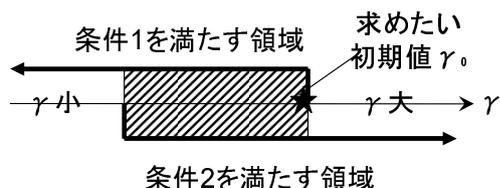


図 5: 条件 1 と 2 を満たすパラメータ  $\gamma$  の領域 (斜線部分). 星印が求めたい初期値.

なる, 条件 1,2 を満たす  $\gamma$  の最大値をとることにする (図 5). ここでパラメータの最大値を取る理由は, 条件 2 で得られるパラメータの最小値を低く見積もる可能性が高いためである. 最初の段階でユーザがラベル付けを行うデータは数個であるため,  $\gamma$  をかなり小さくしても, すべてのデータを正しく学習することができる. しかし, ラベル付けを行っていないデータを正しく学習できる保証はない. そこで, 条件 1,2 を満たす  $\gamma$  の最大値を取り, 条件 2 における誤りのリスクを回避する.

### 3.3 パラメータの更新方針

条件 2 は, 計算機によるラベル付け結果が学習した結果と同じであることを保証するものであり, アルゴリズムの全過程において満たされる必要がある. そこで, もし, 計算機が学習データを正しくラベル付けできなくなったら,  $\gamma$  を自動的に大きくするものとする.

### 3.4 提案する効率的なラベル付け手順

図 6 に提案する効率的なラベル付け手順のフローを示す. 詳細を以下で説明する.

人間によるラベル付け

安定したラベル付け作業とするために, 一番最初に計算機

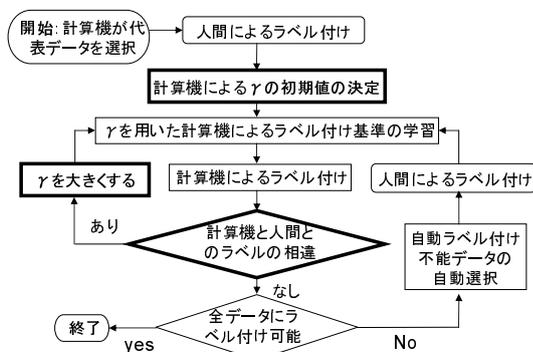


図 6: 計算機と人間との協調作業による事例データベース構築手法

が学習する画像は, 各クラスの典型となる画像が望ましい. ここで, 最初に人がラベルを付ける画像 (指定枚数) を, 計算機がクラスタリング手法である k-means 法 [5] を用いて代表画像を自動的に選択する. 各クラスから 1 枚以上の画像を選択するものとし, できれば同数ずつ選択するのが望ましい.

例えば, 3 クラスのラベル付けを行う場合には, 計算機はまずラベルを付けたい画像データ全体を k-means 法を用いて, 3 クラスに分類する. そして, 特徴ベクトル空間上における各クラスの中心点から, 平均二乗距離で近い順に同数ずつ画像を選択する. 但し, k-means 法はおおまかな分類であるため, 各クラスの中心点近傍にも, 誤分類されている画像が存在する. 選択した画像は, 人間に提示し, 各クラスから同数ずつになるよう選択してもらう.

初期パラメータの選定

次に, 計算機は, 人間が最初にラベルを付けたデータを用いて, パラメータの初期値を選択する. 計算機のラベル付けの誤り数を少なくするため, 上述の条件 1,2 (3.2 節参照) を満たす中から, 最大値を  $\gamma$  の初期値として選択する.

初期値の具体的な決定方法は, いくつか考えられるが, 例えば二分法などを用いると容易に決定できる. また, パラメータ  $\gamma$  は RBF カーネルに用いられるガウス分布の分散の逆数である. そのため, 分布の標準偏差が  $1 \sim 1e+4$  程度とすると,  $\gamma$  は  $1e-8 \sim 1$  程度となる.

学習と終了条件

次に計算機は, 選ばれた  $\gamma$  を用いた SVM で, 人がラベル付けした全画像を学習する. ここで使用する SVM は, 人間のラベル付けに誤りがないものとして, ハード マージンを用いる. 残りの画像を SVM で判定し, すべての画像がマージン領域の外に出たらラベルを付け, 確定して終了する.

新たな画像の提示

マージン領域に画像が残っている場合には, 計算機は識別面に最も近い画像を一枚選択し, 人に提示する. 3 クラス判定ラベル付け (A, B, C) の場合には, A と, B または C の 2 クラスに分けた後, B と C を分ける. 新たな画像の提示は, A と, B または C に分ける識別面近傍から 1 枚と, B と C を分ける識別面近傍から 1 枚と一度に 2 枚ずつ行うものとする.

パラメータ  $\gamma$  の変更と繰り返し

そして, 計算機が提示した画像に人がラベル付けを行い, その結果を計算機は SVM を用いて学習する. このとき, 学習した画像データが誤りなくラベル付けできなければ, パラメータ  $\gamma$  を大きくする. このパラメータを用いて, 学習と終了条件に戻り, 以降の行程を繰り返す.

## 4. 実験と評価

### 4.1 実験条件

提案したラベル付け手法の性能評価実験を以下の二つのデータを対象に行った.

(1) 米国郵便番号手書き数字認識用公開テストデータ [6] (白黒画像  $16 \times 16 = 256$  画素):

数字「0」1550 枚, 数字「1」1268 枚, 数字「2」929 枚 (合計 3747 枚).

(2) 腕金鍍画像データ (カラー画像  $640 \times 480$  画素を, 粗視化した画像 (画像サイズ  $20 \times 15 = 900$  画素) を用いる):

「再利用」183 枚, 「めっき後再利用」281 枚, 「廃棄」267 枚 (合計 731 枚).

上記データセットに対して, 2 クラス判定 (「0」と「1 または 2」, 「再利用」と「めっき後再利用または廃棄」と, 3 クラ

表 1: 提案方法によるラベル付け結果 (10 試行平均, 括弧内は全体に対する割合)

データ	初期値	最終値	ラベルの誤り数	人がラベル付けした数
数字 (2 クラス)	$\gamma=3e-3$	$\gamma=3e-3$	0 枚 (0%)	182±0.3 枚 (4.9%)
数字 (3 クラス)	$\gamma_1=3e-3, \gamma_2=1e-3$	$\gamma_1=3e-3, \gamma_2=1e-3$	0 枚 (0%)	214±0 枚 (5.7%)
錆 (2 クラス)	$\gamma=1e-2$	$\gamma=1e-2$	0 枚 (0%)	83±1.9 枚 (11.4%)
錆 (3 クラス)	$\gamma_1=3e-2, \gamma_2=3e-2$	$\gamma_1=3e-2, \gamma_2=3e-2$	0 枚 (0%)	216±0 枚 (29.5%)

ス判定 (「0」「1」「2」「再利用」「めっき後再利用」「廃棄」) を行う。

最初に人がラベルを付ける枚数は、各クラスから 2 枚ずつとし、2 クラス判定では 4 枚、3 クラス判定では 6 枚とする。また、3 クラス判定ラベル付け (A, B, C) の場合には、A と B または C の 2 クラスに分ける際のパラメータを  $\gamma_1$  とし、B と C を分ける際のパラメータを  $\gamma_2$  とする。この実験では、パラメータの取り扱いを簡単にし、また実験結果を明快に示すために (log スケールで等間隔になるように)、離散値  $\gamma=1e-9, 3e-9, 1e-8, 3e-8, \dots, 1, 3$  : 計 20 値をパラメータ値として採用することにする。よって、パラメータの初期値は上記 20 値から決め、パラメータの変更は、上記 20 値の中で初期値から次に大きい値に変更していくこととする。すべてのラベル付けが終わるまでに、人がラベル付けする枚数と計算機がラベル付けの際に誤った枚数を測定する。この試行を 10 回行い、その平均結果を測定する。

#### 4.2 実験結果

提案方法により、ラベル付けを行った結果について、表 1 に示す。パラメータの初期値は、提案するラベル付け手順で計算機により自動的に選ばれた値とした。また最終値は、その後のパラメータ自動調節で得られたラベル付け終了時のパラメータ値とした。ラベルの誤り数と人がラベル付けした数については、全画像数に対する割合を括弧内に示した。数字画像データ、錆画像データ共に、計算機によるラベル付けの誤り数は 0 で、人間のラベル付け作業量を低く押さえることができている。

今回はいずれの場合も、結果的にパラメータは初期値から変わらなかった。

#### 4.3 省力化効果

ここで、得られた結果から、計算機と人間との協調作業によるラベル付けを行う場合と、人間が全ての画像をラベルを付ける場合 (人海戦術法) について作業量の比較を行う (表 1)。数字の 3 クラス判定ラベル付け作業では、人が全ての画像にラベル付けを行う (人海戦術) では 3747 枚必要なのに対して、提案手法では、わずか 6% にあたる 214 枚で誤ラベルがなく作業が完了する。また、実データである錆画像についても人海戦術の 731 枚に比べて約 30% の作業量で、誤ラベルがなく作業が完了する。この原因は、白黒手書き数字画像では、数字の形の識別という分類問題としては比較的簡単な例である一方、錆画像判定は、微妙な色の違いなども含めて判定する必要があり、人間でもラベル付けするのが困難な問題であるためと考えられる。

本実験では、ラベルの誤りは生じず、作業量は削減されることが示されたが、一般の問題に対しては、ラベルの誤りが生じる可能性がある。

### 5. まとめと今後の課題

本報告では、人間と計算機の協調作業による効率的なデータベース構築手法を提案した。提案手法では、計算機は全対象データから代表的なデータを数個選択し、人間に提示する。人

間に提示された数個のデータにだけラベル付けを行えば、計算機は、ラベル付け精度を高くし、人間がラベル付けなければならないデータ数を削減できる。

設備診断に用いられる実データの例として、腕金再利用判定のための 731 セットの事例データベースの構築を行った「再利用」「めっき後再利用」「廃棄」のクラスへとラベル付けする 3 クラス判定ラベル付けの場合に、ラベルの誤りなく全画像にラベルを付けることができた。

また、ラベル付けの作業量削減効果について、手書き数字認識用の公開データへの適用結果と比較した結果、識別の比較的簡単な手書き数字画像データでは、人が約 6% の枚数をラベル付けるだけで、作業が完了したのに対し、識別の難しい腕金錆画像データでは約 30% のラベル付けを必要とした。

実フィールド上の腕金錆画像のように識別の難しいデータに対するラベル付けでは、誤ラベルが生じる可能性もある。今回検討した事例に関しては、適切な自動パラメータ調節は有効であった。しかし、パラメータの挙動に関する理論的な分析はまだ十分ではない。パラメータを用いて、誤ラベルを減少させることについて、理論的に検証することが今後の課題である。

#### 謝辞

本報告書の作成において使用した腕金錆画像データは、中部電力 (株) 本店販売本部配電部より提供して頂きました。ご協力に深く感謝致します。

#### 参考文献

- [1] 山名 美智子, 村田 博士, 小野田 崇, 大橋 徹, 加藤 誠二: 腕金錆画像に基づく腕金再利用判定精度の向上, 2004 年度人工知能学会全国大会論文集:2F2-02, (2004).
- [2] V. N. Vapnik. *The Nature of Statistical Learning Theory*, Springer, (1995).
- [3] M. Yamana, H. Murata, T. Onoda, T. Ohashi and S. Kato: Development of System for Crossarm Reuse Judgment on the Basis of Classification of Rust Images Using Support Vector Machine, Proceedings of IEEE ICTAI2005, pp.402-406, (2005).
- [4] G. Ratsch, B. Scholkopf, A. J. Smola, S. Mika, T. Onoda, and K.-R. Muller. *Robust ensemble learning*. A.J. Smola, P.L. Bartlett, B. Scholkopf, and D. Schuurmans, editors, *Advances in Large Margin Classifiers*, p.207-220, MIT Press. (2000).
- [5] 高木幹雄, 下田陽久監修, 「新編 画像解析ハンドブック」, 東京大学出版会, (2004).
- [6] Dataset of Handwritten Characters, created by the US Postal Service, URL:http://www.kernel-machines.org/data/usps.mat.gz.