

科学研究費助成事業 研究成果報告書

平成 27 年 6 月 19 日現在

機関番号：12601

研究種目：若手研究(B)

研究期間：2012～2014

課題番号：24760629

研究課題名(和文) 安定的で効率的な推定制御のための適応型非線型回帰分析手法の開発

研究課題名(英文) Development of adaptive nonlinear regression methods for stable and efficient process control

研究代表者

金子 弘昌 (Kaneko, Hiromasa)

東京大学・工学(系)研究科(研究院)・助教

研究者番号：00625171

交付決定額(研究期間全体)：(直接経費) 2,600,000円

研究成果の概要(和文)：測定困難なプロセス変数をリアルタイムに推定するためソフトセンサーが広く使用されている。ソフトセンサーモデルの精度の劣化を低減するため、回帰モデルは逐次再構築されている。しかしモデルを逐次更新する際は、急激なプロセス変化に追従できず、プロセス変数間の非線型性にも対応できない。そこで本研究では、これらを解決する予測精度の高い適応型非線型回帰分析手法を開発した。複数の非線形回帰モデルを用いることで目的とする変数の値を精度良く予測することが可能となる。シミュレーションデータおよび様々な実プラントのデータを用いた解析により、従来手法と比較して提案手法が優位であることを示した。

研究成果の概要(英文)：A soft sensor predicts the values of some process variable y that is difficult to measure. To maintain the predictive ability of a soft sensor model, adaptation mechanisms are applied to soft sensors. However, even these adaptive soft sensors cannot predict the y -values of various process states in chemical plants, and it is difficult to ensure the predictive ability of such models on a long-term basis. Therefore, we propose a method that combines online support vector regression (OSVR) with an ensemble learning system to adapt to nonlinear and time-varying changes in process characteristics and various process states in a plant. Several OSVR models, each of which has an adaptation mechanism and is updated with new data, predict y -values. A final predicted y -value is calculated based on those predicted y -values and Bayes' rule. We analyzed simulation datasets and real industrial datasets, and demonstrate the superiority of the proposed method.

研究分野：プロセスシステム工学、プロセス制御、計量化学、化学情報学

キーワード：プロセス管理 ソフトセンサー モデルの劣化 適応型モデル サポートベクター回帰 アンサンブル学習 ベイズの定理 予測誤差

1. 研究開始当初の背景

産業プラントにおいては、測定困難なプロセス変数を推定する手法として、ソフトセンサーが広く用いられている[1]。ソフトセンサーとは、オンラインで測定可能な変数と測定困難な変数の間で数値モデルを構築し、目的とした変数の値を推定する手法である(図1)。ソフトセンサーを用いることで、オンラインで精度良く目的変数 y の値を推定可能である。このソフトセンサー技術は産業界において特に注目されており、プロセスシステム工学第143委員会におけるワークショップNo.29ソフトセンサーが発足し、産業界におけるソフトセンサーの現状と課題の整理や課題解決へ向けた技術開発が行われた。

上記のワークショップで企業側に実施したアンケートでも明確となったことであるが、解決が急務な問題点の一つが『モデルの劣化』である(図2)。ソフトセンサーは供給組成変化、外気温変化、機器や配管への汚れ付着等のプラントの運転状態の変化によって予測精度が劣化してしまう。この『モデルの劣化』を低減するため、現場では新しい測定データを用いて逐次モデルを更新している[2]。モデルの更新は、温度や圧力などの説明変数 X と y の関係が線形で表現できる場合、X と y の傾きを変更することに対応する。

しかし、モデルを逐次更新する場合、以下の二つの問題が存在する。

- ✓ 1. 急激なプロセス変化に追従できない
- ✓ 2. プロセス変数間の非線形性に対応できない

あるプロセス状態において測定されたデータを取り入れながら長期間更新されたモデルは、同じ状態における予測精度は高いが、プロセス変数の変動が起きた際に対応できない(問題点1)[3]。また、現状はモデルを更新する場合に計算負荷の小さい線形回帰分析手法が使用されているが、プロセス変数間に非線形関係が存在する場合は良好なモデルの構築が不可能である(問題点2)[4]。問題点1の急激な変化を、X と y の間に突然生じた非線形性と捉えれば、X と y の非線形関係を適切に表現しながらモデルを更新できれば、二つの問題点を解決可能といえる。

線形モデルと非線形モデルの特徴を図3に示す。

非線形回帰分析手法を用いたモデル構築には多くの時間がかかり、統計的学習理論が確立されており予測性の高いモデルを構築可能な手法である support vector regression (SVR) [5]も例外ではない。Parrella は、新しいデータを用いて SVR モデルを更新する際、データを選択することで効率的にモデル更新を行うアルゴリズムを開発した [6]。単純にモデルを再構築する場合と同等の精度で、計算負荷の小さいモデル更新が可能となる [7]。

しかし、SVR 法には最適化すべきパラメータが3つ以上存在するが、上記の SVR モデルを効率的に更新する場合は、パラメータの組を固定しなければならない。プラントの状態は急激に変化することもあり、あるパラメータの組ですべてのプラント状態に対応することは困難といえる。

2. 研究の目的

以上の背景を踏まえ、モデルの劣化問題の解決と急激なプラント変化やプロセス変数間の非線形性に対応可能な、予測精度の高い適応型非線形回帰分析手法を開発することを目的とした。

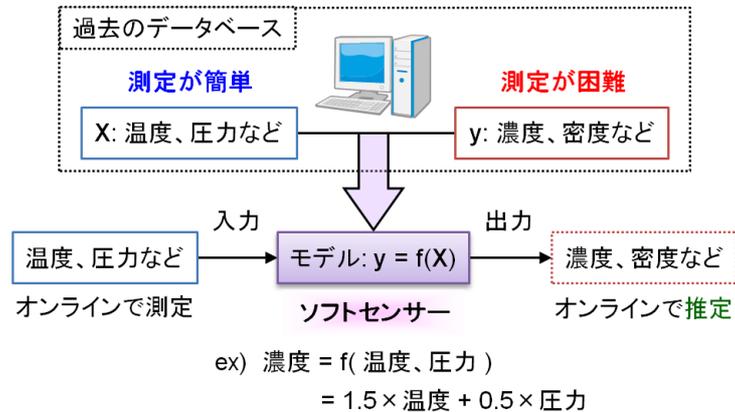


図1. ソフトセンサー

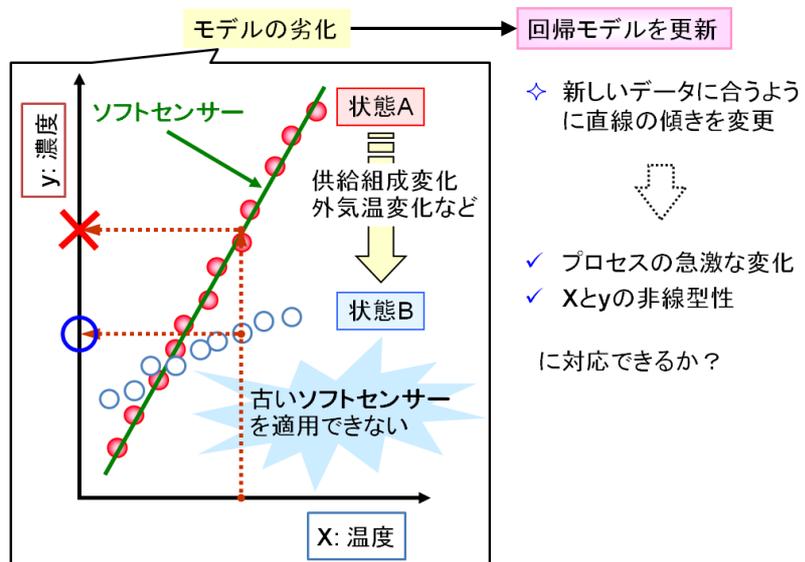
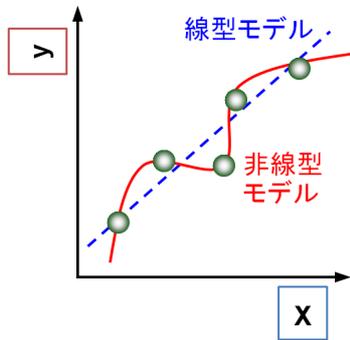


図2. モデルの劣化とその対策、問題点



- 線型モデル**
- モデルが単純
 - モデルの解釈がしやすい
 - ×精度を上げにくい
 - ×非線型性に対応不可能
- 非線型モデル**
- 精度を上げやすい
 - 非線型性に対応可能
 - ×予測性が低い場合がある
 - ×計算時間がかかる

図3. 線形、非線形モデル

3. 研究の方法

(1) Support Vector Regression (SVR)

SVR は support vector machine を回帰分析へと応用した手法である。SVR において、モデル構築用データにおける X のあるデータを $\mathbf{x}^{(i)}$ とした時に回帰式 f は、

$$f(\mathbf{x}^{(i)}) = \phi(\mathbf{x}^{(i)})\mathbf{w} + b \quad (1)$$

と表わすことができる。ここで、 ϕ はある非線形関数、 \mathbf{w} はウェイトベクトル、 b は定数項である。SVR においては、下記の式を最小化するように学習が行われる。

$$\frac{1}{2}\|\mathbf{w}\|^2 + C \sum_{i=1}^n E_{\varepsilon}(f(\mathbf{x}^{(i)}) - y_i) \quad (2)$$

ただし、 n はモデル構築用データ数であり、

$$E_{\varepsilon}(f(\mathbf{x}^{(i)}) - y_i) = \max(0, |f(\mathbf{x}^{(i)}) - y_i| - \varepsilon) \quad (3)$$

である。式(2)を最小化することによって、

モデル構築用データへの当てはまりと汎化能力とのバランスの取れた非線形回帰モデルが得られる。式(2)の C は、2つの項の間で重み付けを調整するための係数であり、式(3)の ε は許容する誤差を表わし、それぞれ学習データの特徴に応じて最適化を行う必要がある。

X の新しいデータ \mathbf{x} に対する y の予測値は以下の式で与えられる。

$$f(\mathbf{x}) = \sum_{i=1}^n (\alpha_i - \alpha_i^*) K(\mathbf{x}^{(i)}, \mathbf{x}) + b \quad (4)$$

ただし K は

$$K(\mathbf{x}^{(i)}, \mathbf{x}) = \phi(\mathbf{x}^{(i)})\phi(\mathbf{x})^T \quad (5)$$

でありカーネル関数と呼ばれる。本研究では以下のガウシアンカーネルを使用する。

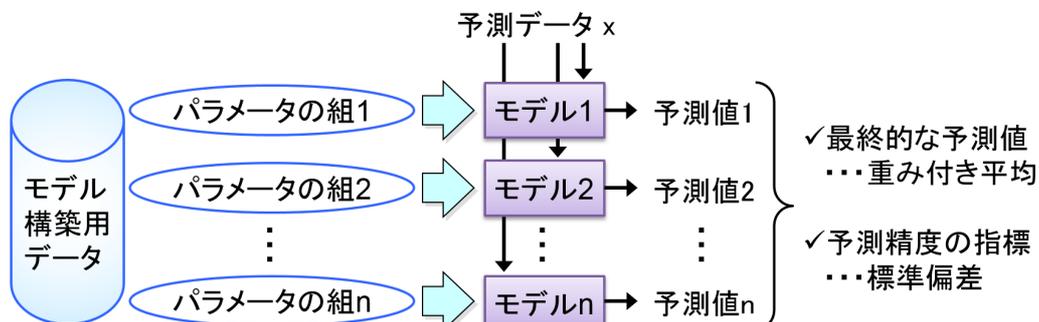
$$K(\mathbf{x}^{(i)}, \mathbf{x}^{(j)}) = \exp\left(-\gamma\|\mathbf{x}^{(i)} - \mathbf{x}^{(j)}\|^2\right) \quad (6)$$

はモデル構築用データから最適化すべきパラメータである。式(4)の α_i, α_i^* は式(2)の最小化により得られる。

以上より SVR モデルの構築のためには C, ε を事前に設定しなければならない。

(2) 提案手法 Ensemble Online Support Vector Regression (EOSVR) [8]

新しいデータを使用して SVR モデルを更新することは可能であるが、前述したように SVR パラメータ C, ε を一通りに固定しなければならない。実際のプラントではプロセス特性が急激に変化することもあり、一つのパラメータの組ではそのような変化に対応不可能であると考えられる。一方、SVR モデルのパラメータの組を最適化するためには、 C, ε についてすべての組み合わせでモデルを構築し、その中から予測的説明分散等の指標を基準にして適切なパラメータの組を選択するか、遺伝的アルゴリズム等の最適化手法を組み合わせるパラメータを最適化しなければならない。オンライン予測をするソフトセンサーにとって、新しいデータ



新しい測定データを用いて各モデルを更新

図4. 提案手法の概念図

が得られた際に毎回最適化を行うことは計算時間の観点から不可能である。

そこで本研究では、アンサンブル学習法 [9] を組み合わせることでこの問題の解決を試みた。アンサンブル学習法とは、事前に複数のモデルを準備することで、予測する際に予測データを各モデルに入力して得られる複数の予測値を総合的に評価して最終的な予測値とする手法である。提案手法の概念図を図 4 に示す。まず、データベース内のデータを時系列に並べ、 k 個ずつモデル構築用データとしてクロスバリデーションによりパラメータの組を最適化する。これにより n 個のパラメータの組が得られたとする。それぞれ同じモデル構築用データを用いて SVR モデルを構築する。

新しいデータが得られた際は、各モデルのパラメータを固定し効率的に SVR モデルを更新する。予測する際は、最新のモデル 1 からモデル n のすべてを用いて予測値を出力し、それらの重み付き平均により最終的な予測値とする。モデル i の重み z_i は以下のように与えられる。

$$z_i = \frac{1}{\text{RMSE}_i^2} \quad (7)$$

RMSE_i は直近のデータで計算されたモデル i の平均二乗誤差や直近のデータの中心によって計算されたモデル i の平均二乗誤差 [10] を表す。新しい y のデータが得られたらすべての SVR モデルを更新する。

さらに、アンサンブル学習により得られる複数の予測値のばらつきを評価する。図 4 における n 個の予測値の標準偏差を指標にすることで予測誤差を推定する。これにより、プロセスの状態ごとのモデルの信頼性を評価することができ、プラントの異常値診断や制御性能の向上に貢献できる。

4. 研究成果

提案手法 EOSVR を MATLAB 上で実装した。SVR モデルを構築するための最適化プログラムとして LIBSVM [11] を使用した。

提案手法の有効性を確認するため、 X と y の間に非線形性を持つシミュレーションデータと実際のプラントで測定されたデータを用いて解析を行った。以下の多様な手法と比較することで提案手法の有用性を検証した。

- ・JITPLS [12]: 予測データとユークリッド距離の近いデータを選択して線形の partial least squares (PLS) モデルを構築

- ・LWPLS [13]: 類似度行列に基づいてデータに重みを付けて PLS モデルを構築

- ・MWPLS [2]: 時間的に近いデータを用いて PLS モデルを構築

- ・OSVR [6]: 時間的に近いデータを用いて SVR モデルを構築

- ・EOSVR: 提案手法

シミュレーションデータは論文 [14] の X と

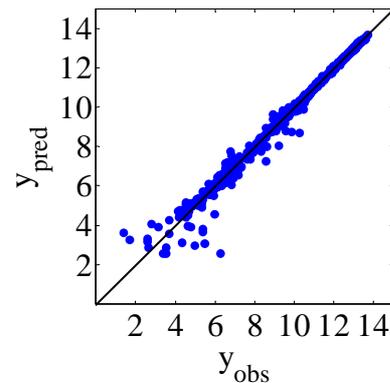
表 1. シミュレーションデータにおける RMSE_p の値 [8]

	RMSE_p
JITPLS	18.0
LWPLS	19.9
MWPLS	13.5
OSVR	11.5
EOSVR	9.5

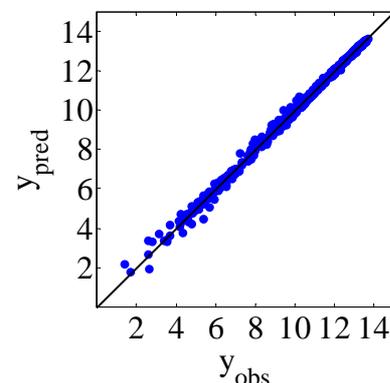
y との非線形関数を使用して生成し、平均 0・標準偏差 0.01 の正規乱数を追加した。モデル構築用データおよびモデル検証用データはそれぞれ 1000 データである。

表 1 にモデル検証用データにおける $\text{RMSE}(\text{RMSE}_p)$ の値を示す。 RMSE_p は予測誤差の指標でありこの値が小さいほどモデルの予測性能は高い。提案手法である EOSVR の値が最も小さく、提案手法を用いることで従来手法と比較して予測的なモデルを構築できることを確認した。

図 6 が EOSVR と RMSE_p 値が二番目に小さい OSVR における実測値と予測値のプロットである。OSVR の結果では y の値が小さい時に予測誤差が大きいのが、EOSVR では全体的に誤差が小さく予測できたことが分かる。提案手法により様々な X と y の間の関係を表現でき、



(a) OSVR



(b) EOSVR

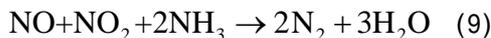
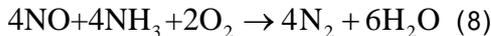
図 5. シミュレーションデータにおける実測値と予測値の関係 [8]

表 2. 排煙脱硝プロセスデータにおける RMSE_p の値 [8]

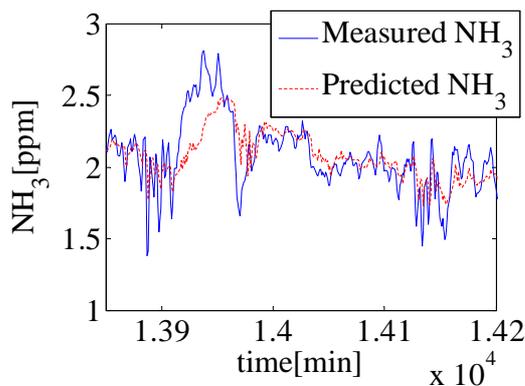
	RMSE _p
JITPLS	0.136
LWPLS	0.176
MWPLS	0.121
OSVR	0.119
EOSVR	0.087

高い予測精度を達成した。

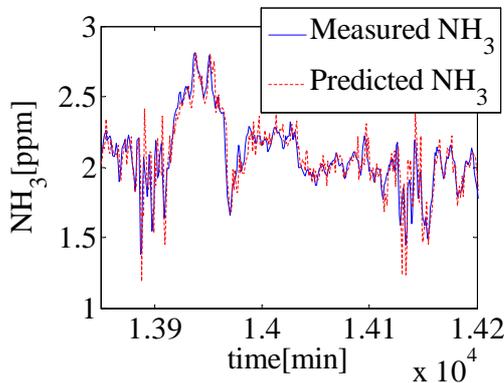
続いて実際のプラントにおいて測定された実データを用いた解析結果の例として、排煙脱硝装置における結果を示す。対象としたプロセスは三井化学大阪工場の排煙脱硝プロセスである。脱硝反応器において、アンモニアを注入して排ガスを触媒層に通すことにより、No_xを窒素と水蒸気に分解させる。反応式は以下の通りである。



脱硝反応器の出口での残存アンモニア濃度、No_x濃度、出口煙道での No_x濃度をオンラインで制御するため、ソフトセンサーが使用される。今回は目的変数を脱硝反応器出口でのア



(a) OSVR



(b) EOSVR

図 6. 排煙脱硝プロセスデータにおける実測値と予測値の時間プロット[8]。青い実線が実測値、赤い点線が予測値を表わす

ンモニア濃度とした。説明変数は混合器、反応器における温度、圧力、流量などの 23 変数である。2012 年 11 月のデータを 20000 データ(20000 分)準備し、最初の 10000 データをモデル構築用データ、次の 10000 データをモデル検証用データとした。

各手法を用いた際のモデル検証用データにおける RMSE_p の値を表 2 に示す。シミュレーションデータの場合と同様にして EOSVR によって RMSE_p 値が最も小さくなることを確認した。提案手法により精度良く予測可能なソフトセンサーを構築できた。

従来手法の一つである OSVR と提案手法である EOSVR における y の実測値と予測値の時間プロットの一例を図 6 に示す。OSVR においては時刻によっては実測値と予測値との差が大きかったが、EOSVR により全体的に誤差が小さく予測できたことが分かる。提案手法を用いることで大幅に精度が向上することを確認した。

図 7 に EOSVR における予測値の標準偏差と予測誤差の絶対値との関係を示す。予測値の標準偏差が大きくなるにつれて予測誤差の絶対値およびそのばらつきが大きくなっている。提案した予測値の標準偏差を指標にすることで事前に予測誤差を推定でき、ソフトセンサーの性能を監視しながら運用できることを確認した。

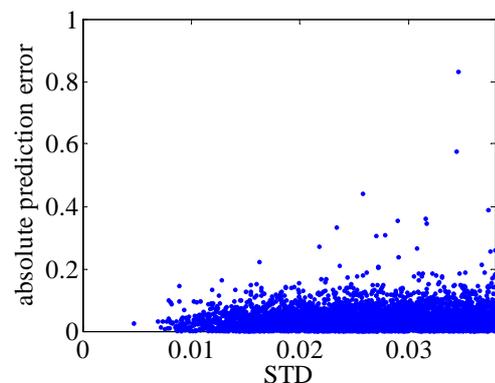


図 7. 排煙脱硝プロセスデータにおける EOSVR モデルの予測値の標準偏差と予測誤差の絶対値との関係[8]

<引用文献>

- [1] Kadlec, P., Gabrys, B., Strandt, S., Comput. Chem. Eng., 2009, 33, 795-814.
- [2] Kaneko, H., Arakawa, M., Funatsu, K., AIChE J., 2009, 55, 87-98.
- [3] Kaneko, H., Arakawa, M., Funatsu, K., AIChE J., 2011, 57, 1506-1513.
- [4] Kaneko, H., Arakawa, M., Funatsu, K., Comput. Chem. Eng., 2011, 35, 1135-1142.
- [5] Bishop, C.M., Pattern recognition and machine learning, Springer, New York, 2006
- [6] Parrella F., Master Thesis,

University of Genoa, 2007.

- [7] Ernst F., Schweikard A., Int. J. CARS., 2009, 4, 439-447.
- [8] Kaneko, H., Funatsu, K., Chemom. Intell. Lab. Syst., 2014, 137, 57-66.
- [9] Kaneko, H., Funatsu, K., J. Chem. Inf. Model., 2014, 54, 2469-2482.
- [10] Kaneko, H., Funatsu, K., J. Chem. Inf. Model., 2013, 53, 2341-2348.
- [11] Chang, C.C., Lin, C.J., LIBSVM: a Library for Support Vector Machines, (2001) Software available at <http://www.csie.ntu.edu.tw/~cjlin/libsvm>
- [12] Cheng, C., Chiu, M.S., Chem. Eng. Sci., 2004, 59, 2801-2810.
- [13] Schaal, S., Atkeson, C.G., Vijayakumar, S., Appl. Intell., 2002, 17, 49-60.
- [14] Li, G., Aute, V., Azarm, S., Struct. Multidiscip. O., 2010, 40, 137-155.

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文](計16件)

Hiromasa Kaneko, Kimito Funatsu, Applicability Domain Based on Ensemble Learning in Classification and Regression Analyses, Journal of Chemical Information and Modeling, 査読有, Vol. 54, pp. 2469-2482, 2014. DOI: 10.1021/ci500364e

Hiromasa Kaneko, Kimito Funatsu, Adaptive Soft Sensor Based on Online Support Vector Regression and Bayesian Ensemble Learning for Various States in Chemical Plants, Chemometrics and Intelligent Laboratory Systems, 査読有, Vol. 137, pp. 57-66, 2014. DOI: 10.1016/j.chemolab.2014.06.008

Hiromasa Kaneko, Kimito Funatsu, Application of Online Support Vector Regression for Soft Sensors, 査読有, AIChE Journal, Vol. 60, pp. 600-612, 2014.

DOI: 10.1002/aic.14299.

[学会発表](計21件)

金子 弘昌, [研究奨励賞] 化学プラントにおける制御性能向上のための推定制御手法に関する研究, 化学工学会 第80年会, 芝浦工業大学(東京都), 2015年3月21日
Hiromasa Kaneko, Kimito Funatsu,

Adaptive regression model for nonlinear and time-varying systems, The 5th French-Japanese Workshop on Computational Methods in Chemistry, 1 July 2014, Strasbourg (France)
Hiromasa Kaneko, Kimito Funatsu, Adaptive Soft Sensor Model Using Online Support Vector Regression and the Time Variable, 2013 AIChE Annual Meeting, 4 November 2013, San Francisco (U.S.A.)

[図書](計1件)

船津 公人, 金子 弘昌, コロナ社, ソフトセンサー入門 -基礎から実用的研究例まで-, 2014, 224 ページ

6. 研究組織

(1)研究代表者

金子 弘昌 (KANeko, Hiromasa)
東京大学・大学院工学系研究科・助教
研究者番号: 00625171

(2)研究分担者

()

研究者番号:

(3)連携研究者

()

研究者番号: