検査員の印象表現に基づいた欠陥候補像の良否判定*

片山隼多** 吉村裕一郎** 青木公也*** 舟橋琢磨*** 奥水大和***
加藤久幸† 池野 純† 山村良太† 小口幸成† 福澤満保†

Defect Classification based on Smart Analysis of the inspector's Impression Expressions of Defects

Hayata KATAYAMA, Yuichiro YOSHIMURA, Kimiya AOKI, Takuma FUNAHASHI, Hiroyasu KOSHIMIZU,

Hisayuki KATOU, Makoto IKENO, Ryota YAMAMURA, Yukinari OGUCHI and Mitsuyasu FUKUSAWA

In this paper, we describe about classification of a defect image for automatic visual inspection. Conventionally, a machine learning approach has been effectively utilized as a method for determining the quality of a defect candidate image in recent years. However, because the logic constructed in the classifier is black box generally, it is difficult to perform the maintenance and operation by inspectors. Therefore, we propose a method of machine learning based on inspector's impression expressions about a defect image, such as shape, density, texture. The logic constructed by our method is not only easy to understand by inspectors, but also expected to abstract a tacit knowledge of inspection. Experimental results indicate that our approach is applicable to the defects provided from the production line.

Key words: visual inspection, machine learning, defect classification, impression expressions, tacit knowledge

1. 諸 言

一般に目視・外観検査の現場では、欠陥候補像のサイズ(長さ、面積等)や密度(一定範囲での個数等)、濃度等を計測したり、限度見本と比較したりすることによって、候補像の良否判定や種類判別を行っている。また欠陥候補像を見た際に受ける印象表現(「細長い」「黒っぽい」「かすれた感じ」といった形や濃度、質感に関する表現)からも判別が行われる。どちらも形式知を用いた判別だが、前者は定量的で明確な形式知であることに對して、後者は定性的で曖昧性を含む形式知である。つまり、印象表現は十分に形式知となっておらず、暗黙知が含まれている可能性がある。このような暗黙知となっている検査技術を継承する際には、暗黙知から形式知への変換が重要な課題となる¹⁾²⁾。この課題が解決されれば、検査の現場における熟練技能等のコツ・カンに頼っている技能の継承が容易になると考えられる。

一方、古くから画像による欠陥の良否判定や種類判別を機械学習によって解決する試みは多い。機械学習において用いる特徴量の設計は重要であり、サンプル画像の調査や現場の形式知に基づき、画像処理の観点から行われる。特徴量の設計はシステムの性能に大きく影響する。これは画像を用いた機械学習全般の問題である。近年では、この問題を回避する手段として Convolutional Neural Networks³⁾等、特徴量の設計をヒトが行う必要がない手法が提案されている。

いずれにせよ、機械学習は外観検査の自動化において有効に機能している。ただし、従来手法の多くは、ヒトが機械学習による判別理由を理解できないという問題点がある。この問題が

Fig.1 Example of defect candidate images

解決されれば、検査員と機械の相互学習が期待され、システムの性能向上に繋がると考えられる⁴⁾⁵⁾。また、システムの運用・保守の点において、発生した分類誤りに対策を講じる際、その原因を究明することも容易になる。つまり、検査の場において機械学習を用いた判別を行う際には、入力したデータから判別結果に繋がる理由を明確にすることが重要であると言える。

2. 研究 目 的

本研究では、図1に示すような金属部品の表面に発生する、目視検査による判定が難しい欠陥候補像を対象とする。具体的には、検査員と同様に欠陥像を見た際に受ける印象の組み合わせから、欠陥判定を行っていることに着目し、欠陥候補像の良否判定(図1において(a)であるか(b)であるか)を自動化することを試みる。製造ラインにおいて検査対象の表面に異常部が検出された際に、図1(a)はその異常部を拭き取る等といった処理

* 原稿受付 平成 28 年 5 月 9 日

* 掲載決定 平成 28 年 8 月 17 日

** 学生会員 中京大学大学院 (愛知県名古屋市中区八事本町 101-2)

*** 正 会 員 中京大学工学部 (愛知県名古屋市中区八事本町 101-2)

† 賛助会員 YKK(株)工機技術本部 (富山県黒部市吉田 200)

Table 1 Example of impression expressions

Sample No.	Inspection image	Answer
1		straight, uniform width
2		blur pattern
⋮	⋮	⋮
N		clear outline

Fig.2 Generation of the mask image

により除去可能であるもの、図 1 (b) は除去不可能なもののことを指す。本研究で扱う製品は金属部品であるため、簡単に除去が可能なものは良品とされる。

前述の通り、現場における検査員の知見(形式知)には、欠陥部のサイズ(〇〇mm以上)など定量的に表現されているものと、濃度や形状(濃い薄い, 細い太い)といった印象表現によって定性的に表されているものが存在する。そこで本研究では、知見のうちの印象表現と画像の特徴量を機械学習によって紐づけ、さらに印象表現と良否判定を紐づける、2 段階のシステムを提案する。最終的な良否判定を行うための情報を検査の現場と同様の言葉とすることにより、検査員が識別器による判別理由を理解しやすくなる。また、生成された識別器に、調査では明らかにならなかった暗黙知(印象の組み合わせ)が含まれることが期待できるシステムの学習方法を提案する。

3. 提案手法

まず、現場における検査員の知見(形式知)の調査を行った。結果の一例を表 1 に示す。検査員は普段、実物を目視することによって検査を行っているが、今回は単一画像を用いた良否判定を行うため、同様に単一画像を見た際に受ける印象が必要となる。そこで、検査員に欠陥候補画像を提示し、どのような印象を受けるかを自由記述のアンケートにより調査した。この際、印象表現の属人性を抑制するため、複数の検査員の合意により意見を決定して頂いた。その後、調査によって得られた印象表現の中から実験に用いる印象群の抽出を行った。

次に、単一検査画像から画像特徴を多数設計し、欠陥候補像について計測を行った。抽出した印象表現を表すには、特徴量として色の濃淡、欠陥部の形、欠陥数等が必要となると考えら

= Discriminator

Fig.3 Learning model of relation between image characteristics and impression expressions

Fig.4 Learning model of relation between impression expressions and quality determination

れる。例えば、平均・分散・エッジ強度等の輝度特徴、面積・長さ・周囲長等のサイズ特徴、円形度・複雑度等の形状特徴である。その後、Floodfill アルゴリズム⁹⁾を用いて、検査画像より欠陥候補像領域のみを抽出したマスク画像を生成し、マスク画像に基づき特徴量の計測を行う。この際、図 2 に示すように生成したマスク画像に膨張処理を行い、元となったマスク画像との排他的論理和により差分領域を抽出することで、マスク領域の外淵部について輝度特徴の計測を行うための画像を生成する。同様に収縮処理を行い差分をとることで、マスク領域の内淵部の輝度特徴の計測を行うための画像を生成する。

次に、曖昧性を含む形式知の要素(印象表現)について、検査員が各サンプル画像から受ける印象を画像特徴との対応において学習する。このモデルを図 3 に示す。印象群の数と同数の識別器を生成し、各識別器が印象群内の 1 つの印象について判定を行う。その後、生成した各識別器に画像特徴を入力することで、その画像特徴を持つ画像はどの印象に当てはまるか判定を行う。この際、各識別器における出力は 0, 1 による 2 値信号ではなく、0.0~1.0 までの実数値による強度とした。これにより、検査員による表現の程度(少し〇〇と思う、非常に××とを感じる等)を再現できると考えられる。

さらに別途、形式知(ここでは曖昧性を含む形式知であり、前述での印象表現)による良否判定を学習する部分を生成する。このモデルを図 4 に示す。この良否判定の結果は印象表現の組み合わせにより判定が行われる。検査員は画像特徴の組み合わせを提示されても解釈しづらい。一方、印象表現の組み合わせは、自らが良否判定を行う際に用いているため、解釈が可能であると考えられる。そして図 5 に示す通り、前者を前段、後者

Fig.5 Machine learning model

Table 2 Proposed Image features

1	Ave. brightness inside a mask	13	Brightness SD at vicinity outer an outline
2	Med. brightness inside a mask	14	Min. brightness at vicinity outer an outline
3	Brightness SD inside a mask	15	Max. brightness at vicinity outer an outline
4	Min. brightness inside a mask	16	Size
5	Max. brightness inside a mask	17	Perimeter
6	Ave. brightness at vicinity inner an outline	18	Complexity
7	Med. brightness at vicinity inner an outline	19	Degree of circularity
8	Brightness SD at vicinity inner an outline	20	Short side length of OBB of a mask
9	Min. brightness at vicinity inner an outline	21	Long side length of OBB of a mask
10	Max. brightness at vicinity inner an outline	22	Number of blobs
11	Ave. brightness at vicinity outer an outline	23	Var. of x-coordinate of masks
12	Med. brightness at vicinity outer an outline	24	Var. of y-coordinate of masks

Table 3 Impression expressions

1	Blackier than a flaw	10	Whitish
2	Blackier than a blot	11	Uniform width
3	Blur pattern	12	Linearly
4	Floating	13	Peeling off
5	Pale	14	Splash
6	Clear outline	15	Hit
7	Lateral length	16	Clearly
8	Multiple defects	17	Usual place
9	Visible shadow	18	Uneven

を後段とした良否判定システムとする。以上より良否の判定が前段の各識別器の印象表現の強度出力により判定されるため、それを見ることによって、識別器による判定基準を検査員が理解できるシステムとなる。

一方、前段の画像に対する印象の学習結果と、その際の候補像の良否判定(正解の判定)を紐付けて後段を学習することを考える。これによって正答率が向上した場合、判定ごとの印象入力を分析する。その判定に用いられた印象入力の強度、つまり暗黙知であった印象の組み合わせを形式知として抽出することが可能であると考えられる。

4. 実験と考察

まず、検査員の識別能力と提案手法の性能との比較を行うために、前述のアンケート調査と同様にして検査員に欠陥候補画像を提示し、その際の正答率を調査した。この実験における判別率は70%程度であった。単一画像ではなく実物を観察した場合の判別率は100%である。正答率の低下の原因は、実際の検査では様々

Table 4 Discrimination rate

Experiment method	Accuracy
Visual judgment by inspectors	70.1%
Experiment 1	79.7%
Experiment 2	88.8%

な方向からワークを観察すること等、実験では普段に比べて判定に必要な情報が欠けていることが考えられる。

次に、提案手法を用いて良否判定を行った。欠陥候補像は検査員による判別実験と同様の画像から143個の欠陥候補像を抽出し、用いる画像特徴には表2に示す24種類、印象群は表3に示す18種類を設計した。つまり、実験に用いる基本的なデータセットは1個の欠陥候補像について、

- ① 欠陥候補像から計測される画像特徴(24種類)
- ② 欠陥候補像に対する検査員の画像印象表現(18種類)
- ③ 欠陥候補像に対する検査員の良否判定結果

の3要素によって構成される。

まず図5中の図3に当たる前段の識別器に Extremely Randomized Trees(ERT)⁷⁾を用い、欠陥候補像から計測される画像特徴(①)を入力データ側、欠陥候補像に対する検査員の画像印象表現(②)を出力データ側として学習する。これによって、画像特徴と印象表現が紐付けられ、

- ②' 欠陥候補像に対する識別器(前段)による画像印象表現(18種類)

が欠陥候補像に対する情報として観測可能となる。

ここで、実験1として後段の識別器にSVMを用い、欠陥候補像に対する検査員の画像印象表現(②)を入力データ側、欠陥候補像に対する検査員の良否判定結果(③)を出力データ側として学習する。この際、SVMにはガウシアンカーネルを用いて非線形識別を行う⁸⁾。そして、前段より得られた判定結果、つまり、欠陥候補像に対する識別器による画像印象表現(②')を入力し、良否判定を行う。判別率は、検査員に調査を行った場合を上回り、約79%であった。

次に実験2として、後段の識別器に実験1と同様にSVMを用い、欠陥候補像に対する識別器による画像印象表現(②')を入力データ側、欠陥候補像に対する検査員の良否判定結果(③)を出力データ側として学習する。つまり、実験1とは後段の学習データが異なる。この場合の判別率は実験1を上回り、約88%であった。各実験の結果を表4に示す。なお、以上の学習・判定実験ではLeave-One-Out法を用いた。

実験1において生成された識別器は、検査員の形式知のみを学習させたものであることに対して、実験2で生成された識別器は検査員の形式知のみではなく、暗黙知が含まれている可能性がある。実験2における印象判定結果の一例を図6に示す。この図は、例として挙げた欠陥候補像(キズ)に対して行った印象判定によって出力された結果、つまり印象の強度のグラフである。アンケート調査の際には検査員は「欠陥の太さが均一である」、「欠陥部が直線的である」といった、2つの印象のみを述べていた。それに対して、提案手法による前段の識別器の出力は、アンケート調査によって得られた2つの印象強度が高くなっていることに加え、「こすれたあとがない」といった、印象の強度も高くなっていることが分かる。また、印象判定から新たに得られた印象は、ヒトが欠陥部を目視した際にも妥当であると判断できる印象表現であ

Fig.6 Example of impression expressions

Fig.7 Experimental image example

る。このことから、アンケート調査では明示されなかった印象表現(暗黙知)を抽出できていると考えられる。

実験2では、アンケート調査では得られていない印象の組合せ、つまり暗黙知を含めて学習することにより、真に良否判定に用いるべき情報を含む識別器となり、判別率が実験1と比べて9%向上したと考えられる。

また、この際に判別に用いている特徴量は、検査の現場と同様に印象表現の組み合わせを用いているため、識別器による判定基準を検査員が理解することができる。

図7に実験画像の一例を示す。ここでは欠陥候補像における良品の例としてヨゴレ、不良品の例としてキズを挙げている。図8は図7に例として挙げた2つの欠陥候補像より画像特徴を計測し、それらを比較したグラフである。また、図9は図7に示した2つの欠陥候補像に対して、提案手法により得た印象判定結果(強度)をグラフとして示している。良否判定を行うにあたっては、特徴空間において類似度が高いサンプルではクラス間の境界面の設定が困難になる場合がある。そのため、これらの2つのサンプルの類似度は、低くなることが望ましい。ここで計測を行った画像特徴による2つのサンプルのcos類似度が0.724であることに對して、印象判定結果におけるcos類似度は0.355であった。つまり前者に比べ後者の分離度が高くなっていると言える。この分離度の高さは図8、図9を比較しても明らかである。このことから、良否判定を行う際に、印象判定結果を用いることは有用であると言える。

また、良否判定に失敗したサンプルに注目すると、実験1、2ともに判定に失敗したサンプルの多くが印象判定による結果(強度)が、設計した全ての印象において0.5未満となっていることが分かった。この際の印象判定の結果の例を図10に示す。良品と不

Fig.8 Comparison among image characteristics

Fig.9 Comparison among the impression expressions

Fig.10 Comparison among impression expressions (a failed example)

良品の印象判定の結果が低い値で揃い、似通ってしまっていることが見て分かる。したがって、特徴量空間において良品と不良品の距離が近くなってしまい、識別に失敗したと考えられる。この要因として、設計した印象表現が十分でなかった可能性が挙げられる。このことは、設計し切れていない潜在的な印象表現が存在しているとも考えられる。つまり、印象表現の組み合わせのみではなく、印象表現そのものが暗黙知となっているため、判別に失敗している可能性がある。そのため、今後は生成された識別器から暗黙知を抽出する方法、設計できていない潜在的な印象表現の抽出方法等の検討に取り組む予定である。

5. 結 言

検査員の印象表現に基づき、検査員が見てもロジックが理解しやすい良否判定システムを提案した。単一検査画像について、検査員と提案手法による判別実験を行った。判別率は、検査員：約 70%，実験 1：約 79%，実験 2：約 88%であった。前段の出力、つまり印象判定の強度に注目することで、暗黙知となっていた良否判定に必要な印象の組み合わせを明確にできる。したがって、暗黙知の形式知への変換が期待できる。

今後は生成された識別器から暗黙知を抽出する方法、設計できていない潜在的な印象表現の抽出方法等の検討に取り組む。

参 考 文 献

- 1) 森和夫, “暗黙知の継承をどう進めるか”, 特技懇, **268**, 43.
- 2) 丁園鎮, “知識開発モデルに関する一考察: SECI モデルの発展を試みて”, 青森公立経営経済学研究, **9**, 2(2004), 43.
- 3) Jonathan Maschi, Ueli Meier, Dan Ciresan, Jurgen Schmidhuber, Gabriel Fricout, “Steel Defect Classification with Max-Pooling Convolutional Neural Networks”, IJCNN2012, 1.
- 4) 渋谷久恵, 岡部隆史, 中川泰夫, “決定木を利用した欠陥分類ルール自动生成技術”, 画像電子学会誌, **36**, 5 (2007), 731.
- 5) Christoph H. Lampert, Hannes Nickisch, Stefan Harmeling, “Learning To Detect Unseen Object Classes by Between-Class Attribute Transfer”, CVPR, 2009, 951.
- 6) Shaw, J.R. QuickFill: An efficient flood fill algorithm. Online article. <http://www.codeproject.com/gdi/QuickFill.asp> [Accessed 08 May 2016]
- 7) Pierre Geurts, Damien Ernst, Louis Wehenkel, “Extremely randomized trees”, Machine Learning, **63**, 1(2006), 3.
- 8) B.E.Boser, I.M.Guyon, and V.N.Vapnik, “A Training Algorithm for Optimal Margin Classifiers”, Proc. Fifth Ann. Workshop Computational Learning Theory, ACM Press, New York, 144