

GUI ソフトウェアの回帰テストにおける画面差異特徴とドキュメント情報を用いた不具合原因分類

鶴崎真理子^{†1} 阿倍博信^{†1}

概要: GUI ソフトウェアの回帰テストにおいて、改修前後の画面の比較による結果確認作業の効率化を目指している。画面を比較して完全一致でない場合、人が目視で表示が異常であるかどうか、どのような原因で異常になっているかを確認し、障害票等のドキュメントを作成して報告する必要がある。本研究では、画面比較時の差異特徴と仕様書やテストシナリオ等のドキュメントから抽出した情報により、表示異常のパターン毎に分類する方式を検討した。

キーワード: 回帰テスト, 画像比較, 画面検証, 原因分類

A Method of the Classification of Errors Using Different Pixels and Information Extracted by Documents on Regression Tests

MARIKO TSURUSAKI^{†1} HIRONOBU ABE^{†1}

1. はじめに

GUI ソフトウェアの回帰テストや環境のリブレース時のテストの現場では、繰り返し行う操作や画面の表示確認による労力が大きいという課題がある。例えば、GUI ソフトウェアの小規模な改修時やプラットフォームの更新時に、変更がない箇所についても不具合やバグが発生しないか確認することで品質を担保する必要があるため、膨大な数の画面の目視確認が実施されている。

GUI ソフトウェアの開発現場では、初回テスト時の画面スクリーンショットをマスタとして2回目以降の画面のスクリーンショットと比較し、画面が正常に表示されているか自動で判定する技術が使用されている[1]。実際の運用では、マスタと完全一致している画面については除外し、残りの画面について目視で確認する。しかし、それらの画面についてどこに異常が出ているか、なぜ表示に不具合がでているか等目視で確認・調査し、障害票や報告書等のドキュメントにまとめる作業が必要であり、この作業に労力がかかっている。著書らは不具合画像の差異の傾向から原因毎に分類する方式を検討、評価し、約80%の精度で正しく分類できることを示した[2]。

本研究では、GUI ソフトウェアの回帰テストにおける画面確認作業の効率化を目的として、画面内の差異の傾向と仕様書等のドキュメント情報から、表示不具合の原因毎に分類する方式を検討している。評価として、画面差異の傾向から、機械学習の手法の一種でクラスタリングによく用いられるランダムフォレスト[3]により分類モデルを生成し、不具合の原因毎に分類できるか確認した。

2. 回帰テストの画面確認における課題

GUI ソフトウェアの回帰テストにおける表示画面の確認作業を効率化することを目的として、操作記録時の画面スクリーンショットをマスタとして操作再生時の画面スクリーンショットとの画素値の比較を行い、完全に一致している画面は目視で確認する必要がないため確認対象から除外し、それ以外の画面の組について目視で表示を確認する。

目視で表示を確認する際、画面内のどこに差異が出ているのか、どのような原因で不具合が出ているのかを確認し、報告書等にまとめる作業を実施しており、これらの作業に労力がかかることが課題となっている。

3. GUI 画面の表示不具合原因分類

3.1 概要・目的

課題への対応策として、完全一致でない画面の組について、画面内の差異の傾向やドキュメント情報から表示不具合の原因毎に分類する方式を検討する。

具体的には、日付、テキストボックス、ボタン等の領域上に出現する差異画素の輝度値の差分から抽出される差異特徴量と、仕様書等のドキュメント情報からモデルを生成する。日付やテキストの場合、文字が異なる差異が多いので、小さい領域に差異が集中している、ボタンの差異の場合、フォーカスによる色の変化による差異が多いので、輝度差が小さい等の特徴があることが考えられる。

モデル活用時には、画面の組から抽出した差異特徴量とドキュメント情報から、どの構成要素に差異が出ているかをモデルにより推定する。不具合原因の推定により、報告

^{†1} 三菱電機株式会社
Mitsubishi Electric Corporation.

書の自動生成等の負荷の軽減が期待できる。

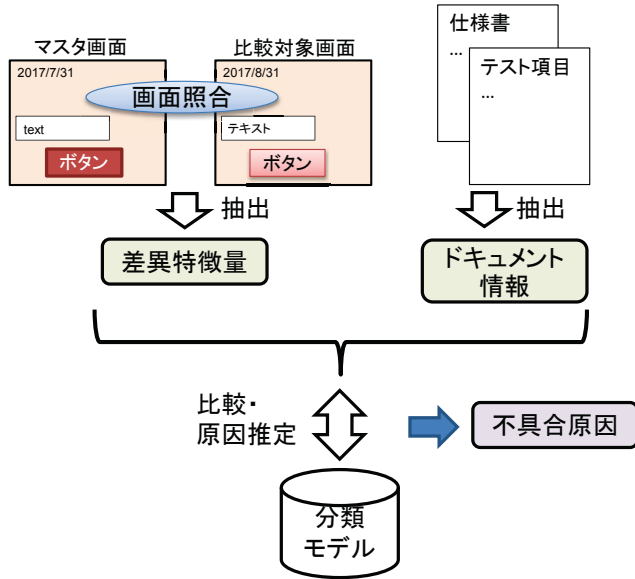


図 1 表示不具合原因分類のイメージ
 Figure 1 Image of Classification of Errors.

3.2 処理の流れ

以下の(1)~(5)の手順で表示不具合の原因分類を行う。

(1) 領域毎に分割

画面を大まかな領域に分割する。例:ステータスバー, タイトルバー, ボタンが集中している領域毎に分割する。

(2) 差異特徴量の抽出

NG と判定された画面の組について, 差異の特徴から表示不具合パターン分類を行う。具体的には, 差異の特徴として差異領域の座標や大きさ, 輝度差の最大/平均/分散, 差異領域内の差異画素の個数等が考えられる。

(3) ドキュメント情報の抽出

画面仕様やテストシナリオ等のドキュメントから, 画面の構成要素, 画面表示直前の操作内容等の情報を抽出する。これらの情報を各領域に属性として付与する。

(4) 学習

照合データを学習データとテストデータに分割する。学習データについて, 表示不具合のパターンをラベル, 特徴量及び領域の属性をデータとして, 教師あり学習により, 不具合原因分類モデルを生成する。

(5) 分類による原因推定

不具合原因分類モデルを基に, テストデータについて分類を行う。差異画像をテストデータとして入力すると, 差異の出ている領域と不具合原因が出力される。

4. 評価

前章の(1)~(5)の手順, 表 1 の評価条件で評価したところ, 91%の精度で正しく分類できることがわかった。各特徴量, 補助情報の寄与率を計算したところ, 領域の位置, サイズ, 差異画素数, 補助情報の各パラメータについてほぼ同様の

値であり, 突出して高い寄与率のパラメータはなかった。

誤判定していたケースでは, テキストボックスとコンボボックスが含まれる領域について, 表 1 の③⇔⑤を誤判定していた。原因として, 評価に使用したテキストボックスとコンボボックスの形状と差異画素の特徴が類似していることが考えられる。今回の誤判定のようなケースは実際に発生することが予想されるので, 定義した異常クラスを分類できるような特徴量の検討が必要である。

また, 今回は領域を手動で分割したが, 実際に適用する際には画面内の部品を認識して領域とする等, 自動で領域を抽出する方式と併せて使用することが考えられるため, 画面内の部品の認識方式についても検討する必要がある。

表 1 評価条件

Table 1 Condition of Evaluation.

項目	内容
画面の構成要素	スクロールバー, 表, テキストボックス, ボタン, コンボボックス, 日時
差異画像数	100
画面種類	1
分割領域数	3
補助情報	表更新, スクロールバー操作, テキスト入力
異常クラス	① スクロール位置が異なる ② 表内の表示内容が異なる ③ テキストボックス内のテキストが異なる ④ ボタンフォーカスがきかない ⑤ コンボボックスの中身が異なる
差異特徴量	領域座標, 幅, 高さ, 類似度 (差異画素の個数/全画素の個数), 輝度差最大値, 輝度差平均値
評価方法	ランダムフォレスト (5-fold cross validation)

5. まとめ

本稿では, GUI ソフトウェアの回帰テストにおいて, 操作記録・再生時に取得した画面スクリーンショット画像を照合し, 画像内の差異画素の傾向及びドキュメント情報から不具合の原因に分類する方式を検討, 評価した。具体的には, 提案方式において画面自動照合で不一致と判定された画面の組の差異の特徴量から, ランダムフォレストにより表示不具合のクラスに分類できるか評価を実施し, 約 90%の精度で正しいクラスに分類できることが分かった。

今後, 機能として実現するために, 差異領域の自動検出方式やドキュメント情報の抽出の仕組み等を検討する。

参考文献

- [1] 鶴崎真理子, 阿倍博信, “マルチプラットフォーム対応 GUI 試験システムにおける画面照合方式,” 第 14 回情報科学技術フォーラム(FIT2015), 2015
- [2] 鶴崎真理子, 阿倍博信, “差異画素抽出に基づく画面照合機能の特徴とする GUI 試験システム,” マルチメディア, 分散, 協調とモバイル(DICOMO2016)シンポジウム, 2016
- [3] Breiman, Leo (2001). “Random Forests”. *Machine Learning* 45 (1): 5-3