

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 A Display Comparison Method on Multi-Platform Test System for GUI Applications

鶴崎 真理子†
Mariko Tsurusaki

阿倍 博信†
Hironobu Abe

1. はじめに

ソフトウェアの回帰試験や環境のリプレース時の GUI 試験において、人手の試験操作による工数が大きいという課題がある。本稿では、リモートデスクトップの一種である VNC(Virtual Network Computing)ベースの GUI 操作記録再生技術と、GUI 操作記録再生時にキャプチャした画面の画像照合技術による GUI 試験省力化方式を提案する。キャプチャ画面の画像照合の評価を実施したところ、輝度値の相関により類似度を算出する手法で、98%程度正しく照合できることを確認した。

2. GUI 試験における課題

ソフトウェアの回帰試験や環境のリプレース時の GUI 試験において、試験者が何度も同じ操作を行うため、人手の試験操作による工数がかかるという課題がある。さらに、操作の記録再生を行う際の、目視による GUI 画面確認にも工数がかかる。

前者の課題を解決するものとして、操作の記録・自動再生を行うツールである、UWSC[1]や Selenium[2]等が知られている。しかし、これらは OS が Windows に限定される、ブラウザ試験に制限される等の制限が存在する。したがって、Linux の組み込み機器や、Android 端末等について適用するためには、別の方式を検討する必要がある。また、後者については、画面の表示を自動で確認するための画像照合機能が搭載されているツールとして、HP の Unified Functional Testing[3]が知られている。しかし、この機能はビットマップチェックにより判定するものであり、1ビットでも異なると不一致と判定されてしまう。

3. GUI 試験省力化方式

前章で述べた GUI 試験における課題の解決策として、2つの方式を提案する。1つ目は、VNC ベースの操作記録再生[4]であり、2つ目は画面の自動照合により試験結果判定を効率化する、GUI 画面照合である。

3.1 VNC ベースの操作記録再生

本方式は、リモートデスクトップの一種である VNC に GUI 操作の記録機能及び再生機能を追加することにより、GUI 試験の省力化を実現するものである。具体的には、図 1 で表しているシステム構成図の操作端末 (VNC クライアント) から試験端末 (VNC サーバ) の遠隔操作を行う際に送信される操作コマンドを操作端末側で記録しておき、再生時には、操作端末で記録した操作コマンドを読み込み、試験端末に操作コマンドを再送信することによって自動再生を行う。VNC は Windows、Linux、Android 等の OS に対応しているため、本方式もこれらのプラットフォームに対応している。

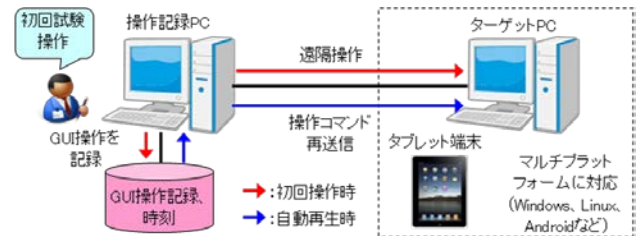


図 1. システム構成図

3.2 GUI 画面照合

図 2 で表しているように、操作記録時及び再生時にキャプチャした画像を照合することによって、試験結果の確認作業を効率化する。具体的には、照合する画面の類似している度合を表す類似度を算出し、操作再生時の画面が正常に表示されたかどうかを判定する。

関連研究として、類似画像検索がある。類似画像検索は、Google Similar Images[5]や Yahoo! ラボの Visual Seeker[6]などのアプリケーションに使用されている。これらは、画像を指定すると、類似した画像をインターネット上から探し、表示するサービスである。類似画像検索では、一般的に以下の手法が用いられている。

(1) 相関ベースの手法

正規化二乗誤差等、照合する二枚の画像各画素の輝度値の相関値を評価に用いる。

(2) 特徴ベースの手法

SIFT[7]や SURF[8]等の局所特徴量を二枚の画像からそれぞれ抽出し、その対応関係を評価に用いる。一般に、自然背景等の複雑な画像の照合を行う際に有効な手法として知られている。

GUI 画面照合では、(1)、(2)のアルゴリズムを使用して操作記録画面と再生画面の組の類似度を算出し、類似度より、試験結果 (OK/NG) を判定する。



図 2. GUI 画面照合イメージ

4. GUI 画面照合方式の評価

本評価では、評価データとして、Web システム 50 画面の操作記録・再生時のキャプチャ画面の組を使用した。評価データはログイン画面等の文字入力フォームを含む画面、リンク等の文字が多く表示されている画面、表が表示されている画面等で構成されている。

GUI 画面照合の評価を以下の手順で実施する。

(1) 評価データを学習データとテストデータに分ける。

† 三菱電機株式会社, Mitsubishi Electric Corporation

- (2)学習データについて、各画面の組の類似度を算出する。
 (3)学習データについて、(2)で算出した類似度と目視で確認した結果 (OK/NG) を入力データとして判定モデルを生成する。
 (4)テストデータについて、(3)で生成した判定モデルによる判定精度を算出する。

4.1 画面類似度算出

類似画像検索で一般的に用いられる、(1)相関ベースの手法と(2)特徴ベースの手法を用いて類似度を算出した。

(1) 相関ベースの手法

SSD (Sum of Squared Difference)を類似度とする。値が小さいほど類似していることを表す。

(2) 特徴ベースの手法

局所特徴量として一般的に使用されている、SIFT、SURF、ORB[9]を用いて、対応点の数により類似度を算出する。具体的には、下記①～④の順序を実施する。

- ①照合対象の画像において特徴点と特徴量を算出する。
- ②①で算出した特徴点及び特徴量より、対応点を求める。
- ③②の各対応点の座標の距離が閾値以上の場合、対応点のリストから除外する。
- ④類似度 = (対応点の個数×2) / (照合対象の2枚の画像内の全特徴点の個数) とする。値が大きいほど類似していることを表す。

4.2 判定モデルによる判定精度の算出

各画面について算出した類似度と目視の確認結果より、閾値を算出する。今回の評価では、学習データから一般的に使用されている機械学習の手法である SVM(Support Vector Machine)[10]により OK/NG の2クラス判定モデルを生成し、テストデータによる判定精度を求めた。判定精度とは、正常に判定されたデータ数を全体の数で割った割合(%)である。また、OKのデータをNGと誤判定する場合、NGのデータをOKと誤判定する場合の発生割合(%)を求めた。

今回の評価では、評価データが少ないため、クロスバリデーション(交差検証)による判定精度の評価を行った。表1に各手法による判定精度及び誤判定率を示す。

表1. 各手法による判定精度

| 類似度算出方式 | 判定精度 (%) | OK を NG と誤判定 (%) | NG を OK と誤判定 (%) |
|---------|----------|------------------|------------------|
| SSD | 98.04 | 0 | 1.96 |
| SIFT | 88.24 | 0 | 11.76 |
| SURF | 98.04 | 0 | 1.96 |
| ORB | 96.08 | 1.96 | 1.96 |

4.3 考察

評価結果に対して考察を行った結果、今回使用した評価データでは、SSDとSURFによる方式を使用した場合に最も精度よく判定できていることが分かった。

しかし、判定は失敗する場合があります、そのほとんどがNGのデータをOKと誤判定するものであった。実際に画面試験に適用する場合、機械的にOKと判定できるデータを除外し、それ以外のデータについて目視で照合を行うことによって人手による工数を削減することが望ましい。このときNGのデータをOKと誤判定してしまうと、誤っ

た照合結果が出力されてしまう。以下の誤判定のケースがあった。

(1) 差異のある領域が小さい場合

文字入力フォームやダイアログ等の、小さい領域での差異により、NGである画面をOKと誤判定していた。

(2) 特徴が少ない画面の場合

アイコンや文字がない等、特徴が少ない画面では、検出される特徴点の数が少ない。このとき、アイコンや文字がないが色が異なる等の、本来はNGである画面をOKと誤判定していた。

ケース(2)の誤判定は特徴ベースの手法を使用したときに発生した。今回の評価では、SURFではケース(2)の誤判定はなかったが、特徴ベースの手法なので、このような誤判定が生じる可能性がある。したがって、特徴点の少ない画面については、特徴点抽出の際のパラメータのチューニングが必要であることが分かった。

また、全ての方式において、ケース(1)の誤判定があった。この原因として、評価した方式では、画面全体の輝度値の差分を均等に使用して類似度を算出するため、画面全体に対してサイズが小さい領域の差異が類似度として反映されにくいことが考えられる。

5. おわりに

本稿では、GUIアプリケーションの試験効率化を目的とした、操作記録再生及びGUI画面照合を提案した。また、GUI画面照合について評価を実施し、SSD及びSURFによる手法で98%程度正しく照合できることを確認した。

考察を実施した結果、特徴点を検出されない画面の場合、局所特徴量で適切に類似度を算出できないという課題や、差異のある領域が小さい場合にNGのデータをOKと誤判定してしまう課題があった。

今後、これらの課題について、局所特徴量算出時のパラメータのチューニングや、OK/NG判定の際の閾値の調整を実施する。

参考文献

- [1]<http://www.uwsc.info/>
- [2]<http://www.seleniumhq.org/>
- [3]<http://www8.hp.com/jp/ja/software-solutions/unified-functional-automated-testing/>
- [4]鶴崎真理子, 阿倍博信, “マルチプラットフォーム対応GUI試験省力化方式,” 電子情報通信学会総合大会, 2015.
- [5]<http://googleblog.blogspot.jp/2009/10/similar-images-graduates-from-google.html>
- [6]<http://visseeker2.yahoo-labs.jp/vs/>
- [7]D. G. Lowe, “Object recognition from local scale-invariant features”, Proc. of IEEE International Conference on Computer Vision (ICCV), pp.1150-1157, 1999.
- [8]Herbert Bay, Tinne Tuytelaars, Luc VanGool “SURF: Speeded Up Robust Features”, computer vision-ECCV Lecture Notes in Computer Science, 2006.
- [9]E. Rublee, V. Rabaud, K. Konolige, and G. Bradski. ORB: an efficient alternative to SIFT or SURF. In Proc. of IEEE International Conference on Computer Vision, 2011.
- [10]津田宏治, “サポートベクターマシンとは何か”, 電子情報通信学会誌, Vol. 83, No. 6, pp. 460-466, 2000.