

レビュータスクにおける情報量と視線の関係分析

斉藤 功樹^{†‡}, 日高 昇平[‡]

Koki Saito, Shohei Hidaka

[†]日本ユニシス株式会社, [‡]北陸先端科学技術大学院大学

Nihon Unisys, Ltd., Japan Advanced Institute of Science and Technology

koki.saito@uniys.co.jp

概要

ソフトウェア開発の要件定義書レビューでは、定義書が顧客の求めるシステムを必要十分に表現しているかを判別することが重要である。先行研究では、その判別能力をはかる課題を開発し、視線との関係を分析したところ、判別成功時には特定の顧客要求文で *fixation* が増加傾向にあった。そこで、本研究では顧客要求文の持つ情報量を独自に定義し、視線との関係を分析した。その結果、情報量が多い文をより注視する場合に、レビューの成功可能性が高いことが示唆された。

キーワード：要件定義、レビュー、必要十分性、情報量、*fixation*

1. はじめに

ソフトウェア開発において、上流工程の成果物である設計書や仕様書の品質が重要であることはよく知られている[1]。レビューが成果物の欠陥を検出/修正する手法であり、様々な手法が開発され、品質に与える影響の比較が行われている[2]。レビュー手法にはそれぞれ一長一短があり、全ての成果物に対して有効なレビュー手法は存在しない。

さらに、Uwano ら[3]はレビュー品質に与える影響は手法により違いよりも個人差の方が大きいと指摘し、視線情報を用いてソフトウェアのコードレビューの品質を評価するシステムを開発した。コードレビュー時の品質を評価する取り組みは多く検討され始めているものの、自然言語で記述された上流工程の成果物のレビュー品質に関する研究は少ない。

斉藤ら[4][5]はレビューを、要件定義書を満たすシステムと顧客が求めるシステムの突合と考え、システムを幾何学図形で象徴化したレビュー課題を開発し(図1)、課題の正誤と視線の関係について分析した。その結果、正答時には特定の顧客要求文において *fixation* (固視, 1箇所を注視している視線の集まり)が増加しており、情報量が多い文に着目することで正答している可能性が示唆された。しかし、顧客要求文に含ま

れる情報量は定量化されておらず、正答時にどの情報に着目しているのかは明らかになっていない。

そこで、本研究では Saito らのデータセットを用いて、顧客要求文の情報量を定義し、視線情報との関係を分析することで正答時のレビューア特性を調査した。Saito ら[5]はシステムを図形事例の集合で象徴化することで、レビューを図形事例の突合と捉えた。その際、顧客要求文から作成される図形事例の集合を **R**、顧客が求める図形事例の集合を **P** と定義した。集合 **P** は顧客が求めるシステムを象徴化しているため、システムごとに要素数は一定である。一方で、集合 **R** は、顧客要求文を読み進めるにつれてその要素数が減少するという特徴を持つ。したがって、顧客要求文がもつ情報量は、要素数の変化に起因するものであると考えられる。そこで、本研究では顧客要求文が進むにつれて変化する要素数などの特徴量に着目し、情報量の候補として採用した。それらの候補と視線情報の関係を分析することで、正答時のレビューア特性を明らかにした。

2. 方法

2.1. データセット

Saito らの研究[5]にて構築されたデータセットを用いた。データセットには、7名の参加者(5名の男性と2名の女性、平均年齢は41.9 (SD = 10.9))が、計13問のレビュー課題に取り組んだ結果とその際の視線情報が含まれる。視線情報は、gazepoint社のGP3 HD Eye Tracker (サンプリングレート: 150Hz, 正確度 (Accuracy): 0.5-1.0°, 非接触型)を用いて計測された。計91問(7名×13問)のうち、1問は視線情報が正しく取得できていなかったため、計90問分を分析に用いる。

■本題01

顧客要求

1. 大は正三角形または正四角形または正五角形である
2. 小は正三角形または正五角形である
3. 大が正四角形の場合、大は黒塗りである
4. 大が正四角形の場合、大の点と小の点は1点のみ接する
5. 小が正三角形の場合、大の辺と小の辺は1辺以上接する
6. 大が黒塗りの場合、小は大と同じ形ではない
7. 大が白塗りの場合、小は大と同じ形である
8. 大が正五角形の場合、小の色は大の色と違う
9. 大が黒塗りの場合、大の辺と小の辺は2辺のみ接する
10. 小が正五角形の場合、大の辺と小の辺は1辺以上接する
11. 大が正五角形の場合、大の点と小の点は1点のみ接する
12. 大と小の塗りが同じ場合、大の辺と小の辺は1辺のみ接する

正例

負例

タイマー : 04 : 59 中断

ルール/問題に戻る

■回答欄
顧客要求を満たす「図形事例の集合」と「正例」が過不足なく一致するか？

- 一致する
- 一致しない

確信度 :

- 全く自信がない
- あまり自信がない
- どちらともいえない
- まあまあ自信がある
- かなり自信がある

次の問題へ進む

図 1 開発されたレビュー課題（左側が顧客要求文，中央が幾何学図形で象徴化されたシステム，右側が回答欄）[5]

2.2. 情報量の定義と分析手法

レビュー課題にて，顧客要求文の番号が進むにつれて変化する集合 R に着目し，その要素数を基に3つの特徴量を情報量の候補として定義した．3つの特徴量は，集合 R の要素数，要素の減少数および減少率とした．ここで，要素数の初期値は組合せ可能な図形事例の最大値である 267 とし，最初の顧客要求文においては初期値からの要素数の変化を減少数/減少率とした．2番目以降の顧客要求文においては，一つ前の文で作成される集合を基として要素数を計算し，その変化を減少数/減少率とした．

3つの特徴量を基に，顧客要求文との相関と fixation との関連を分析し，3つの特徴量が顧客要求文の情報量として適してるかを調査した．Fixation は，gaze point 社の Gaze point analysis に実装されている独自アルゴリズム (Jacob らの研究[6]を基に作成) を用いて算出された．最初に，顧客要求文ごとの特徴量の推移を確認する目的で，顧客要求文番号と特徴量の相関関係を算出した．第二に，正答時のレビューア特性を調査する目的で，正答/誤答それぞれで特徴量における単位時間当たりの fixation 数のヒストグラムを算出した．単位時間当たりの fixation 数は，各要求文を注視している際の fixation の総数を注視時間で割ったものである．一般に

fixation の総数は，注視時間に比例して増加するため，文章の長さや読みやすさなどの影響を受ける．それらの影響を排除するため，注視時間で割った単位時間当たりの fixation 数を用いた．ヒストグラムを基に，正答/誤答で特徴量による fixation 数の違いを分析し，3つの特徴量が情報量として適してるかを評価した．

3. 分析結果と考察

3.1. 特徴量と顧客要求文との相関関係

13問において，3つの特徴量は，顧客要求文の番号との間に負の有意な相関がみられ（要素数と顧客要求文番号のスピアマンの相関係数 $r = -0.950, p < 0.001$ ，減少数と顧客要求文番号のスピアマンの相関係数 $r = -0.853, p < 0.001$ ，減少率と顧客要求文番号のスピアマンの相関係数 $r = -0.443, p < 0.001$ ），要素数が最も相関が強く，次いで減少数，減少率であった．顧客要求文の番号ごとの減少数の箱ひげ図を図 2 に示す．

Saito らの研究[5]では，正答の場合には顧客要求文の番号3において最も fixation 数が多い傾向にあったため，3つの特徴量の値の大小が情報量を示していないことが示唆された．Saito らの研究を基にすると，fixation

が大きいほど、それに費やす認知負荷が大きいことから正答/誤答で有意に差がみられた文番号3付近において最も情報量が高いと想定される、しかし、3つの特徴量はすべて有意な負の相関がみられたため、単に3つの特徴量の大きさが情報量を示すのではなく、特定の値にて情報量が多いと考えられる。そこで、3.2では3つの特徴量と fixation の関係を分析し、正答/誤答にて特徴量による fixation の違いを調査した

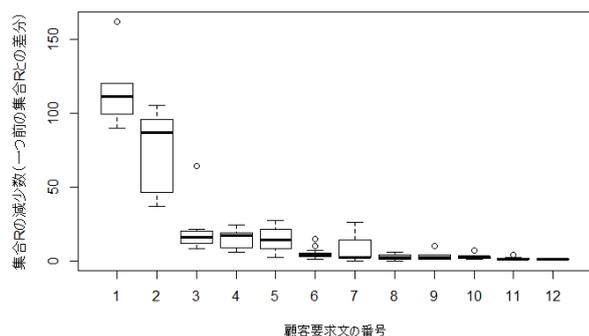


図 2 顧客要求文の番号ごとの減少数の箱ひげ図

3.2. 特徴量と fixation の関係

特徴量ごとの、単位時間当たりの fixation 数のヒストグラムを基に、正答/誤答での違いを分析したところ、減少数でピーク位置に違いがみられた (図 3)。要素数と減少率においては、正答/誤答での単位時間当たりの fixation 数に大きな違いはみられなかった。減少数では、正答の場合には誤答と比較して減少数が 15-25, 95-100, 110-115 の幅にてピークがみられた。これらの減少数をもつ顧客要求文にて情報量が大きいと考えられる。そこで、減少数と顧客要求文の番号との対応関係を調査した。

その結果、fixation 数のピークがみられた減少数に対応する顧客要求文の番号の多くは前半に集中しており、前半の文にて情報量が多いことが示唆された。さらに、減少数 0-5 では誤答時の fixation 数が多く、これに対応する顧客要求文は後半の番号であり、後半の文は情報量が少ないことが示唆された。これらの結果は、Saito らの研究[5]で示唆された報告 (前半の文で情報量が大きく、後半の文で情報量が少ない) と一致する。

したがって、情報量としては減少数が最も適していると考えられる。正答時には情報量の多い文 (減少数 15-25, 95-100, 110-115 に対応する文) により多くの認知リソースを費やしており、一方で情報量の少ない文

(減少数 0-5 に対応する文) には少ないリソースを費やしていることが示唆された。

以上を整理すると、顧客要求の必要十分性を判別する上での重要な情報量は、顧客要求文における減少数であり、正答時のレビューアは情報量の多い文 (減少数 15-25, 95-100, 110-115 に対応する文) に着目しレビューすることで正答している可能性が示唆された。

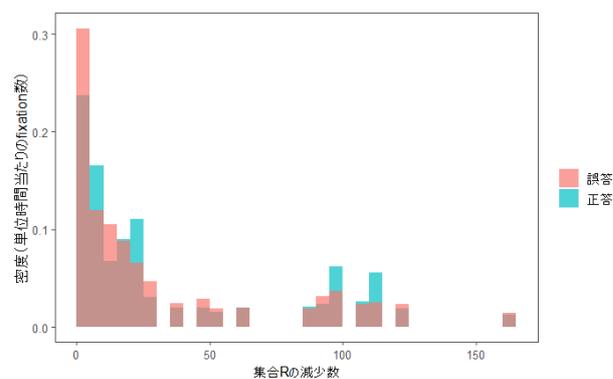


図 3 減少数ごとの単位時間当たりの fixation 数のヒストグラム

4. 今後の予定

本研究では、参加者の問題を解く際の方策や文以外の視線情報 (図形事例の視線情報など) を考慮にいれていないため、それらによって重要な情報が変化する可能性がある。問題を解く際の方策によって、重要な情報が異なる可能性もあり、文以外の図形事例がより重要な情報を持っている可能性も否定できない。

しかし、それらの影響を考慮した分析にはより多くの参加者の視線情報が必要である。今後は、Saito らの研究[5]にて構築されたレビュー課題を用いて、新たに参加者を増やして更なる分析を実施予定である。

文献

- [1] B. Boehm and V. R. Basili, (2001) "Top 10 list [software development]," Computer, vol. 34, no. 1, pp. 135-137.
- [2] T. Thelin, P. Runeson, and C. Wohlin, (2003) "An experimental comparison of usage-based and checklist-based reading," IEEE Transactions on Software Engineering, vol. 29, no. 8, pp. 687-704.
- [3] H. Uwano, M. Nakamura, A. Monden, and K. Matsumoto, (2007) "Exploiting Eye Movements for Evaluating Reviewer's Performance in Software Review," IEICE Transactions on Fundamentals of Electronics

Communications and Computer Sciences, vol. E90-A, no. 10, pp. 317–328.

- [4] 斉藤功樹・日高昇平, (2019) “定義の必要十分性判別能力を試す課題の構築,” 日本認知科学会第36回大会論文集, pp. 188–192.
- [5] K. Saito and S. Hidaka, (2020) “Gaze behavior in a review-a-definition task,” The 42nd Annual Meeting of the Cognitive Science Society (CogSci2020).
- [6] R. J. K. Jacob, (1995) “Eye Tracking in Advanced Interface Design,” Virtual Environments and Advanced Interface Design, W. Barfield and T. A. Furness, Eds. New York, USA: Oxford University Press, Inc., pp. 258–288.