

# ユーザーレビューを用いた要求品質の構造化と重要度算出に関する研究

後藤 正幸 研究室

0532014 井口 琢斗

## 1. 研究の背景と目的

消費者の要求品質を把握し、製品の品質機能へとリンクさせるための方法論として、品質機能展開(以下、QFD)の利用価値は高い[1]。QFD では、要求品質展開表を作成する際、一般的にアンケート調査を用いて消費者の要求品質を抽出している。しかし、アンケート調査では、質問項目にない情報を収集できないことや、多くの回答数を収集するには膨大な労力と時間的なコストがかかるという問題がある。先行研究[2]では、ユーザーレビューの自由記述文章に対して形態素解析を行い、抽出した単語で構造図を作成している。しかし、単語だけでは文脈を判断することができず、消費者の要求を正確に抽出できない。また、重要度の算出方法においても改善すべき点がある。そこで、本研究では、従来の品質機能展開表作成における問題解決のために、係り受け分析を用いた構造化法、並びにベイズ推定と確率分布間の距離の概念に基づく重要度算出の方法を提案する。さらに、携帯電話を事例とした分析を行い、提案する方法の検証を行う。

## 2. 品質機能展開と先行研究

QFD は、1970 年代に赤尾教授らによって提唱された品質管理手法であり、現在では製品開発の分野において世界中の企業で活用されている。アンケート調査から抽出された要求品質と品質特性の対応関係と要求品質の重要度を用いて、品質特性の重要度を算出することにより、顧客の要求を満たす品質特性を把握しようとする方法である。しかし、アンケート調査に基づく QFD には分析コストなどの面で問題点も存在する。

これらの解決策として、Web サイト上のユーザーレビューは、自由に意見を書きこめる「自由記述文章」と数段階で製品を評価付けできる「評価ポイント」で構成されている。すでに、石塚ら[2]は、ユーザーレビューに対して形態素解析を行い、ブランド・エクイティの要素の1つである知覚品質を抽出する方法を示した。QFD を応用して、横軸に知覚品質、縦軸にブランド・アイデンティティをとり、マトリックス表を作成している。知覚品質の手がかりとなる単語の抽出は、ユーザーレビューの自由記述文章の出現頻度が 15 回以上かつ単語の重要度が基準値を超えたものとしている。また、単語の重要度算出には、以下の方法を示している。

$S$  段階の評価ポイントを  $i = 1, 2, \dots, S$  で表し、各評価ポイントで出現した単語の頻度の和を  $N_i (i = 1, 2, \dots, S)$  と定義する。

抽出した  $j$  番目の単語 ( $j = 1, 2, \dots, K$ ) が、評価ポイント  $i$  において出現した頻度を  $X_{ij}$  とすると、

$N_i = \sum_{j=1}^K X_{ij}$  になる。また、単語の出現頻度を  $\sum_{i=1}^S X_{ij}$  とすると、重要度  $Z_j$  は、

$$X_{ij}^* = \frac{X_{ij}}{N_i} \quad (1) \quad Z_j = \sum_{i=1}^S X_{ij} \times \frac{\max_i \{X_{ij}^*\}}{\min_i \{X_{ij}^*\}} \quad (2)$$

となる。評価ポイント数により出現頻度が異なるため、 $N_i$  で割ることにより出現頻度を相対化し、評価ポイント各カテゴリ間での最小値から最大値までの変化の程度を倍率の形で求めている。そして、最小値と最大値の倍率と単語の出現頻度の積により重要度を算出する。しかし、以下の点が課題として挙げられる。

形態素解析から単語を抽出し、これらを独立的に分析しているため、肯定・否定の文脈が考慮されていない。

類似性のある単語が多数存在する可能性があり、これが結果に影響を与えている。

評価ポイント間での単語頻度の最小値と最大値の倍率しか考慮しておらず、他の評価ポイントのばらつきが考慮されていない。

## 3. 要求品質の重要度算出方法提案法の手順

### 3-1. 先行研究の問題解決策

2 節で示した、の問題点の解決策として、抽出した単語に対し係り受け分析を行い、KJ 法を用いて、類似性の高い単語の結合・整理を行う。なお、本研究では、係り受け分析で得られた単語のセットを“キーワードセット”と呼ぶ。次に、の問題点の解決策として、重要度を考えた場合、次の 2 点をみたくことが合理的と考えられる。

- ) キーワードセットの出現数自体が多いものほど重要度が高い。
- ) キーワードセットの評価ポイント上の確率分布とキーワードセット全体の評価ポイントの確率分布との差が大きいほど重要度が高い。

そこで、本研究では、確率分布間の差とキーワードセットの出現頻度を考慮した重要度算出式を提案する。

### 3-2. 要求品質の構造図と重要度算出法の手順

本研究では、Web サイト上に存在するユーザーレビューの自由記述文章と評価ポイントを用いた要求品質展開表の作成と重要度算出を行う。以下に、要求品質の構造図作成と重要度の算出の手順を示す。

- [手順 1] Web 上のユーザーレビューから自由記述文章と評価ポイントを抽出
- [手順 2] 収集した自由記述文章に対して形態素解析を行い、名詞、動詞、形容詞を抽出し、不要語を除去
- [手順 3] 手順 2 で抽出した単語に対して係り受け分析を実行し、キーワードセットを抽出
- [手順 4] KL 法を用いてキーワードセットを類似しているもの同士にグルーピング
- [手順 5] 評価ポイントごとのキーワードセットの出現頻度をカウント
- [手順 6] キーワードセットの出現頻度を基に重要度を算出

### 3.3. 要求品質重要度の算出方法

全キーワードセットの出現頻度における各評価ポイント  $i$  ( $i=1,2,\dots,S$ ) の出現率推定量  $p_i$ 、抽出したキーワードセット  $C_j$  ( $j=1,2,\dots,K$ ) の各評価ポイント上の出現率推定量  $q_i^j$  をベイズ推定で構成すると、

$$p_i = \frac{N_i + 1}{N + S}, \quad q_i^j = \frac{X_{ij} + 1}{\sum_{i=1}^S X_{ij} + S} \quad (3)$$

となる  $p_1, p_2, \dots, p_s$  は、抽出した全てのキーワードセットの出現割合推定量であるので、これと  $q_1^j, q_2^j, \dots, q_s^j$  の差が大きいほど、キーワードセット  $C_j$  は評価ポイント  $i$  の出現割合に影響を与え、重要度 ( $z_j$ ) が高いと言える。そこで、本研究では、キーワードセットの出現頻度  $\sum_{i=1}^S X_{ij}$  も考慮し、次式の重要度算出法を提案する。

$$Z_j = d_j \times \sum_{i=1}^S X_{ij} \quad (4)$$

ただし、 $d_j$  は  $p_1, p_2, p_3, \dots, p_s$  と  $q_1^j, q_2^j, q_3^j, \dots, q_s^j$  の分布間距離であり、次に 3 つの計算方法を示す。

<方法 1> ユークリッド距離

$$d_j^E = \sqrt{\sum_{i=1}^S (p_i - q_i^j)^2} \quad (5)$$

<方法 2> Kullback-Leibler 情報量

$$d_j^K = \sum_{i=1}^S p_i \log \frac{p_i}{q_i^j} \quad (6)$$

<方法 3> コサイン尺度

$$d_j^C = \frac{\sum_{i=1}^S p_i q_i^j}{\sqrt{\sum_{i=1}^S p_i^2 \sum_{i=1}^S q_i^j{}^2}} \quad (7)$$

## 4. 結果

本研究では、携帯電話を事例に提案法を適用し、評価を行った。対象のデータは、価格.com における携帯電話のユーザーレビュー 2000 件分である。係り受け分析と KJ 法を用いて抽出された 116 組のキーワードセットから作成された要求品質展開表が表 1 である。また、ブランドへの関心度を出現頻度の割合から算出し、比較したものが図 1 である。各項目軸に付与されている  $z$  の数字は、キーワードセットの出現頻度の和から大項目の出現頻度を求め、KL 情報量(6)式を用いて算出した重要度である。この結果から、ユーザーは汚れ・傷を重視しているが、両ブランド共にこの項目の頻度は小さいことがわかる。CASIO 社のユーザーは、機能の項目に関心が高いことがわかり、SHARP 社は画面表示の関心が高いことがわかった。

## 5. 考察

ユーザーが重視している製品特性と評価が高い製品特性に差異があることがわかった。図 1 から SHARP 社のボタンについての関心が低いことがわかる。しかし、関心が低くても重要度が低い場合は、企業はその製品特性に重点をおかなくてもよい。そのかわり、重要度の高い製品特性についての関心をあげることに重点をおくべきである。SHARP 社の場合、重要度の高い汚れ・傷についての関心が低いので、評価を高める必要がある。詳細に述べると、「指紋 - つきにくい」などについて消費者から評価を得られるような企業戦略をたてる必要があると考えられる。

## 6. 結論

本研究では、Web サイト上におけるユーザーレビューの自由記述文章を用いた要求品質展開表の作成手順と重要度算出法を示した。ユーザーレビューを用いたことで従来よりも大量のデータを扱うことが可能となり、調査コストの削減につながると考えられる。また、本研究で提案した手法は、分布の推定と分布間の類似度から重要度を求めるものであり、合理性を有していると考えられる。

## 参考文献

- [1] 赤尾洋二:「品質展開入門」, 日科技連出版社, (1990)
- [2] 石塚隆男, Jamsranjav Baasankhuu: “知覚品質機能展開によるブランドマネジメント”, 経営情報学会平成 19 年秋季全国研究発表大会予稿集, pp. 214-217, (2007)

表 1. 要求品質展開表

大項目	中項目1	中項目2	小項目	満足度の数					
				1	2	3	4	5	
デザイン	GOOD	全体	デザイン	良い	7	10	10	32	61
			デザイン	かっこいい	0	0	0	4	7
			デザイン	シンプル	0	1	1	2	9
		デザイン	可愛い	0	0	1	2	0	
		部分	携帯	良い	0	0	2	0	5
			携帯	悪い	1	1	4	11	16
	携帯		小さい	0	0	0	2	0	
	BAD	全体	デザイン	悪い	0	0	2	17	19
			携帯	良い	0	0	2	3	8
			携帯	悪い	2	0	1	5	1
		部分	携帯	悪い	0	1	0	5	4
			携帯	悪い	0	0	2	2	2
携帯			大きい	0	0	1	7	8	
ボタン	GOOD	全体	ボタン	押しやすい	1	1	2	10	13
			ボタン	大きい	0	0	1	9	2
			決定ボタン	押しやすい	0	0	0	0	0
		部分	決定ボタン	大きい	0	0	1	0	0
			サイドボタン	押しやすい	0	0	0	0	0
			サイドボタン	押しにくい	2	5	5	13	10
	BAD	全体	ボタン	小さい	0	1	7	10	12
			ボタン	悪い	0	1	2	2	2
			ボタン	開閉	0	3	2	0	2
		部分	決定ボタン	押しにくい	1	1	0	6	1
			決定ボタン	小さい	0	1	0	1	0
			サイドボタン	押しにくい	0	0	0	0	1

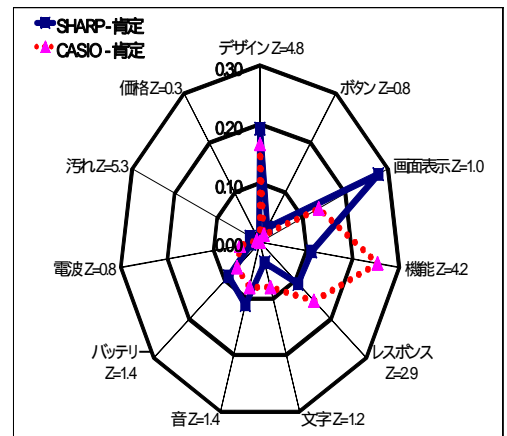


図 1. 重要度と企業の関心度