

第9章 ユーザビリティ評価手法の提案

9.1 研究の背景と目的

企業でのユーザビリティ評価は次機種のプロトタイプを他社競合製品や、代替案のプロトタイプとの比較において行われるケースが多く、このような現場での評価は少数のプロトタイプに加え、少数の被験者という制約上、統計的な手法を適用することが困難である。そのため、発話プロトコル等による定性的な分析が主に用いられている[1][2]。

これは「定量的調査は定性的調査の約4倍の費用が必要」とニールセンが指摘[3]するように、十分な数の被験者による評価実験は開発時のコストおよび時間が膨大になるということが主な理由である。しかしながら、定性的な手法は実験者の恣意的な判断が入りやすいため、筆者らの企業での経験からも、限られたプロトタイプと被験者数の制約の中で、たとえ精度はそれほど高くなくても定量的な手法の導入へのニーズは高いと言える。

そこで、本研究では企業の現場での活用が期待できる比較的低コストでかつ短時間で実施可能な定量的手法を導入し、かつ従来の定性的手法も活かした新たなユーザビリティ評価手法の提案を行う。本手法の最大の特徴は、これまでの定性的手法に加えラフ集合[4]および区間AHP法[5][6]という定量的手法を導入した点にある。また、これに加え、企業でのユーザビリティ評価はデザインへの改善提案のフィードバックを前提としているため、設計論的視点からは単なる問題点の発見に留まらず、評価結果とデザイン要因との関連性を明らかにすることが望ましい。

本手法ではこの視点も配慮している点も特徴として挙げられよう。また本研究では、デジタルミュージックプレーヤ（以下デジタルプレーヤと略記）を事例に、提案手法の有効性について検証し考察する。

9.2 新ユーザビリティ評価手法の概要

前述の目的に沿って構築した新ユーザビリティ評価手法（以下新評価手法と略記）のプロセスは次のとおりである。

- 1) 区間AHP法を用いた被対象製品（またはプロトタイプ）の「視覚的な使いやすさ」の5段階主観評価
 - 2) 対象製品を用いて被験者に複数のタスクを課すユーザビリティテストによるプロトコル解析、区間AHP法を用いた機種ごとの「実際の使いやすさ」およびタスク毎の「実際の使いやすさ」の5段階主観評価
 - 3) 上記1)、2)で得られたデータに基づき、ラフ集合理論を用いた各製品の「視覚的な使いやすさ」および「実際の使いやすさ」に対する貢献要因と阻害要因の分析
- なお、ここで用いている「視覚的な使いやすさ」とは被験者が対象製品に手を触れることなく見た目の印象で判断する使いやすさと定義し、「実際の使いやすさ」とは被験者が実際にユー

ザビリティテストを経験した後に判断する使いやすさと定義する。これらを評価対象とした理由は筆者らの企業での経験から、「視覚的な使いやすさ」と「実際の使いやすさ」は必ずしも一致せず、両者を一致させることが優れたインタフェースデザインの目標のひとつであると考えたためである。

次に上記の各プロセス毎に評価法の概要を説明する。

(1) 区間 AHP 法を用いた「視覚的な使いやすさ感」の主観評価

被験者に製品の「視覚的な使いやすさ」を5段階で評価させた後、その評点の幾何平均を基に、区間 AHP 法で計算する。田中英夫が提唱した区間 AHP 法[7]の特徴は、従来の AHP 法では、「 $A > B$, $B > C$ のとき、 $A > C$ 」となる強推移性または半推移性しか求められなかったのに対し、「 $A > B$, $B > C$ であってかつ $A > C$ でない」となる非推移性も求めることができる点にある。従来の AHP 法はこの非推移性は被験者の論理的な矛盾とみなして、手法的に強推移性また半推移性になるよう修正を求めている。だが、近年の感性の研究では行動経済学も含めて、非推移性の判断は人間の特徴であることが指摘されている。したがって、区間 AHP 法は従来の AHP 法を特殊な場合として包含するため、より精度の高い手法と考えられるため採用した。また、この手法では非推移性の曖昧性は区間で示されるため、視覚的な考察ができるという利点もある。

(2) ユーザビリティテストによるプロトコル解析および

区間 AHP 法を用いた「実際の使いやすさ」評価

プロトコル解析を用いる点は従来のユーザビリティテストと同様であるが、本手法では対象製品毎、タスク毎に求めた5段階評価点の幾何平均を基に区間 AHP 法を用いて「実際の使いやすさ」を算出する点、および発話プロトコルデータから「優位項目」と「問題項目」を抽出し、これを次項のラフ集合理論を用いた分析の元データとして利用する点が特徴である。

(3) ラフ集合理論を用いた「視覚的な使いやすさ」および「実際の使いやすさ」に対する貢献要因と阻害要因の分析

まず、各製品の「視覚的な使いやすさ」のデザイン的要因についてラフ集合を用いて分析する。なお、ラフ集合は、一般に企業で行われているユーザビリティ評価のように少ないサンプルを用いた実験の場合でも目的変数から説明変数を求めるいわゆる逆問題を扱うことができるため採用した[8] [9]。

ラフ集合計算の元である決定表を作成する際に、属性（説明変数）と結論（目的変数）を求める必要がある。ユーザビリティテストの被験者の発話データをもとに、実験者が被験者の認知している箇所を抽出し、認知部位として決定表にまとめる。この認知部位は一般に形態要素（属性値）で構成されている。また単独の属性値や属性値の組み合わせが決定ルールであり決定クラスへの貢献度合いを示す指標となる。

次に決定表を分析して、幾つかのグループの各決定ルールに共通して含まれる属性値（以降、共通属性値と略記）となる属性値を抽出する。新評価手法ではこの共通属性値を抽出する手法として、決定ルール分析法[12]を用いた。

以上のプロセスを経て「視覚的な使いやすさ」および「実際の使いやすさ」への貢献要因と阻害要因を特定することができ、この分析結果をデザインプロセスへフィードバックすることによりインタフェースデザインの改善を図ることができる。

9.3 新ユーザビリティ評価手法を適用した事例実験

本研究では事例実験として、2006年10月時点で販売されていたデジタルプレーヤ12機種をKJ法[10]により分類し、代表的なインタフェースタイプを持つデジタルプレーヤ5機種(図9.1の左から、A、P、H、S、V)を用いた。次に前述2項の新手法の手順に従って実験を行った。

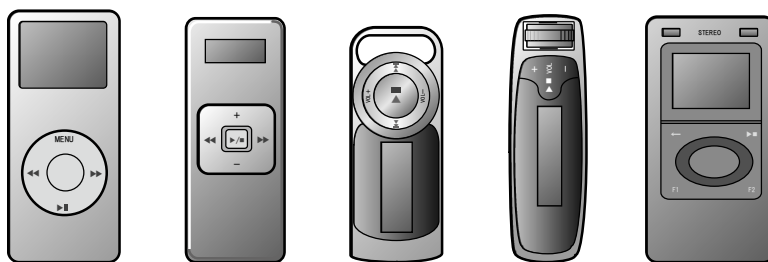


図 9.1 サンプル(製品)のイラスト画

(1) 実験の構成と内容

まず、被験者10名(男性5名と女性5名の大学生)によるサンプルを手にとらず、視覚的な使いやすさだけに着目しての5段階評価の一対比較を行った。なお本実験は企業のユーザビリティテストでの一般的なサンプル数と被験者数に準じた。

表 9.1 タスクの内容

<p>■タスク1 (初心者)</p> <p>本体の電源を入れ、電池があることを確認してから「歌手A」の「曲名B」を、直接聞いてください。その後、音量を適切な大きさに調整し、演奏を3分20秒付近まで飛ばし、サビが流れているのを確認してから音楽の再生を止め、画面を最初のメイン画面に戻してください。</p>
<p>■タスク2 (中級者)</p> <p>あなたの好きな一曲を繰り返し聴けるようにしてください。繰り返し、曲が流れるように設定が終わると、音楽の再生を止め、画面を最初のメイン画面に戻してください。</p>
<p>■タスク3 (上級者)</p> <p>ランダム(シャッフル)に曲を聴けるようにしてください。ただし、一曲目は聴きたい曲が出るまで、曲を飛ばしてから聴いてください。ある程度、曲を聴いてから音楽を聴くのを止めてください。最後に電源を切り、プレーヤを最初の状態に戻してください。</p>

その後、同一被験者を使ったユーザビリティテストを行なった。具体的には、実際の製品をサンプルとして使用し、表 9.1 に示す 3 つのタスクを課した。また、実験の様子をビデオカメラに収録し、被験者の発話データと各タスクの達成時間、各サンプルの問題項目、および優位項目の指摘を記録した。次に、各タスクの終了後に、表 9.2 に示す各タスク内の細目（タスク 1：15 項目、タスク 2：2 項目、タスク 3：14 項目）についての 5 段階評価を実施し、最後に実際の使用結果を踏まえての 5 段階評価による総合的な一対評価を行った。

なお、ユーザビリティテスト前に実施した視覚的な使いやすさの評価に関しては統計的な有効性を確保するために、本実験の後、さらに男性 10 名と女性 10 名の大学生を被験者とした 5 段階評価の一対比較を行った。

表 9.2 タスク毎の評価項目

<p>■タスク1</p> <ol style="list-style-type: none"> 1) 電源ONボタンがわかりやすい 2) 電源を入れやすい 3) 決定ボタンがわかりやすい 4) 決定ボタンが押しやすい 5) 音量調整ボタンがわかりやすい 6) 音量調整画面が見やすい 7) 音量調整がしやすい 8) 曲の早送り・巻き戻しボタンがわかりやすい 9) 早送りや巻き戻し中に、再生時間が分かり易い 10) 曲の早送り・巻き戻しがしやすい 11) 曲の再生ボタンがわかりやすい 12) 再生ボタンが押しやすい 13) 曲の停止ボタンがわかりやすい 14) 停止ボタンが押しやすい 15) 聴きたい曲を選曲しやすい <p>■タスク2</p> <ol style="list-style-type: none"> 16) 曲の繰り返し操作設定が行いやすい 17) 繰り返し演奏設定になったことが分かり易い <p>■タスク3</p> <ol style="list-style-type: none"> 18) ランダム(シャッフル)操作設定が行いやすい 19) ランダム(シャッフル)設定になったことが分かり易い 20) 前曲・次曲への操作がわかりやすい 21) 前曲・次曲への操作ボタンが押しやすい 22) 電源のOFFボタンがわかりやすい 23) 電源が切りやすい 24) 曲を選曲するときに他にどのような曲があるか分かり易い 25) 画面が見えやすい 26) アイコン(文字)が見えやすい 27) アイコン(文字)の並びが見やすい 28) アイコン(文字)の意味がわかりやすい 29) 前の画面に戻るキー操作がわかりやすい 30) 次の画面に進むキー操作がわかりやすい 31) 電池残量が分かりやすい 32) 操作ボタンが操作しやすい
--

(2) 一対比較評価の結果と考察

ユーザビリティテスト前の視覚的な使いやすさの一対比較の評価と、ユーザビリティテスト後の総合的評価の一対比較の評価を区間 AHP 法で分析シグラフ化した結果を図 9.2 に示す。

この図を見るとユーザビリティテストの前後では評価が大きく異なっていることが判る。まず、テスト前の視覚的な使いやすさは、サンプル H または S を除くと全順序の傾向が示されて

いる。また各サンプルの区間の幅はほぼ同じである。さらに、1位と2位は区間が重ならないため順序が確定しているが、3位以下は重なり合っているため明確な順序は確定していない。なお、視覚的な使いやすさに寄与しているデザイン上の特徴の抽出については次章で説明する。

一方、テスト後の総合評価では最下位のサンプルVが上位に移行し、またその区間は一番広い。これはサンプルVに対する被験者の評価にバラツキがあることを示している。また、テスト前の評価が中位のサンプルHは下位に後退して、評価のバラツキがほとんどない。これはサンプルHには評価の一致するような大きな欠点があることを示唆している。次に、1位と2位の順位がテスト前後で逆転し、かつテスト後の評価が一部重なり合っている。また、実験前と比べ、実験後は2つサンプルの区間も広がっている。したがって、実験後のサンプルVの高い評価への移行の要因と、サンプルHのバラツキがほとんどない評価の要因、サンプルAとサンプルPの高い評価の要因などについての分析と考察が必要となる。次章以降でそれらの要因の分析と考察を行う。

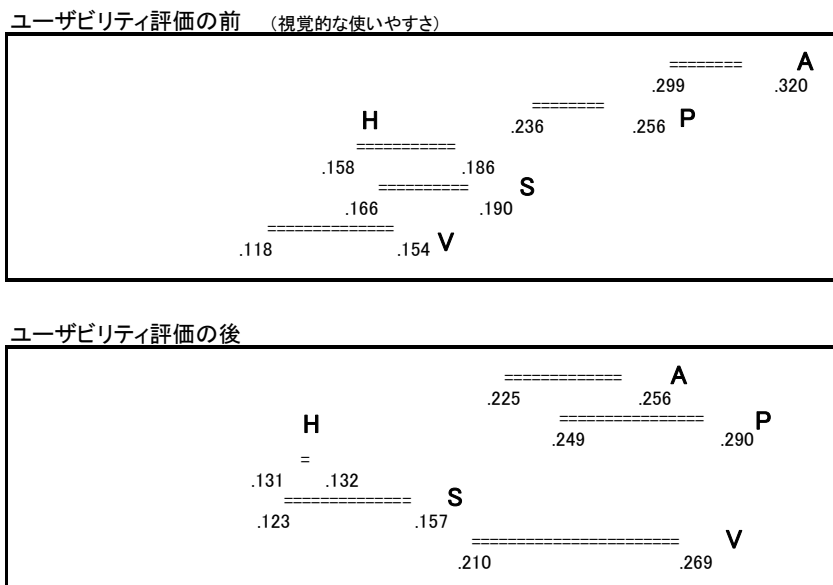


図 9.2 区間 AHP 法による分析結果

(3) タスク毎の評価結果分析と考察

表 9.2 に示す 3 つの各タスクの細目を 5 段階評価してもらい、その被験者 10 名の各平均値を求め、図 9.3 に示すように各サンプルについて昇順で並び替えた。図 9.3 の各枠内の右端の数値「1～3」はタスク番号を示す。また、平均値(a)の値「2.5」、「3.0」、および「4.0」に境界線を施した。尚、紙面の都合上、サンプル A 以外の平均値(a)は割愛した。

図の下端には、サンプルごとに、31 項目の評価の各平均値(a)の全体平均値(b)を記してある。例えば、図 9.3 の左端の下に記した平均値(b)3.76 は最上段の平均値(a)「1.7」から最下段の平均値(a)「4.6」までの平均値を示す。なお、表 1 に示すように、各タスク 1、2 および 3 はタスク内容がそれぞれ初心者、中級者、および上級者を対象としたものである。具体的には、入

り口のやさしい操作から難しい機能の操作までの各段階に対応している。

図 9.3 の実験結果から、サンプル A とサンプル V は、タスク 3 の上級者を対象にした細目の評価項目が高い評価となっている。一方、サンプル P はタスク 3 が低い評価であり、タスク 1 が高い評価となっている。一方、サンプル H とサンプル S の各タスクの評価はバラツキが大きく、はっきりとした傾向が認められない。つまり、サンプル A とサンプル V、サンプル P と比較して、インタフェースのデザインコンセプトが明確ではないことが見て取れる。それだけでなく、この 2 つのサンプルの評価の全体平均値(b)も前者の 3 つのサンプルと比較してかなり低い。特に、サンプル H は図 9.3 の左端の上から 3 番目の境界線が示すように、平均値「4.0」以上の項目がない。図 9.3 の下端の平均値も「2.73」と他と比べ顕著に低い値である。

以上の分析結果から、サンプル P は初心者には配慮し、操作の基本をきちんと守って設計されているため、図 9.2 のテスト後の高い評価につながったと考えられる。他方、サンプル A とサンプル V については上級者に配慮し、難しいタスクが比較的分かり易かったことがテスト後の高い評価に繋がったと考察できる。すなわち、両者とも対象ユーザーを的確に捉えた明確なデザインコンセプトを持っていることが示された。

	A	V	P	S	H	
	平均値 (a)					
2.5 ↓	22) 1.7 3	1) 1	29) 3	22) 3	32) 3	
	5) 2.5 1	16) 2	1) 1	16) 2	4) 1	
	32) 2.9 3	22) 3	22) 3	23) 3	3) 1	
3.0 ↓	16) 3.1 2	19) 3	25) 3	29) 3	21) 3	
	15) 3.2 1	24) 3	23) 3	24) 3	15) 1	
	7) 3.3 1	2) 1	15) 1	30) 3	24) 3	
	19) 3.3 3	11) 1	24) 3	1) 1	25) 3	
	1) 3.5 1	8) 1	27) 3	3) 1	12) 1	
	10) 3.5 1	5) 1	30) 3	15) 1	17) 2	
	13) 3.6 1	10) 1	10) 1	25) 3	26) 3	
	23) 3.6 3	13) 1	26) 3	32) 3	7) 1	
	6) 3.7 1	17) 2	32) 3	4) 1	29) 3	
	29) 3.7 3	14) 1	16) 2	26) 3	27) 3	
	18) 3.8 3	18) 3	28) 3	10) 1	30) 3	
	24) 3.8 3	23) 3	21) 3	27) 3	16) 2	
	4) 3.9 1	26) 3	6) 1	28) 3	10) 1	
	14) 3.9 1	27) 3	4) 1	18) 3	20) 3	
4.0 ↓	3) 4 1	32) 3	2) 1	2) 1	8) 1	
	17) 4 2	9) 1	3) 1	20) 3	14) 1	
	28) 4 3	15) 1	11) 1	31) 3	1) 1	
	30) 4 3	20) 3	17) 2	8) 1	5) 1	
	2) 4.1 1	12) 1	18) 3	19) 3	6) 1	
	9) 4.1 1	28) 3	19) 3	17) 2	19) 3	
	11) 4.1 1	25) 3	20) 3	9) 1	22) 3	
	26) 4.2 3	30) 3	31) 3	13) 1	2) 1	
	12) 4.3 1	31) 3	7) 1	14) 1	18) 3	
	20) 4.3 3	21) 3	12) 1	12) 1	28) 3	
	27) 4.3 3	3) 1	9) 1	21) 3	31) 3	
	21) 4.4 3	6) 1	13) 1	7) 1	13) 1	
	25) 4.4 3	7) 1	8) 1	11) 1	11) 1	
	8) 4.5 1	29) 3	14) 1	6) 1	23) 3	
	31) 4.6 3	4) 1	5) 1	5) 1	9) 1	
	平均値 (b)	3.76	3.73	3.59	3.2	2.73

図 9.3 タスク分析の評価項目の順位

(4) 発話データの結果分析と考察

被験者 10 人の実験の様子を録画したデータから、被験者が口頭で指摘した各サンプルについての「優位項目」と「問題項目」を抽出した。表 3 にその内容を整理した一覧表を示す。表 9.3 中の○印は複数のサンプルに共通する項目で、●印はそのサンプルの単独項目である。したがって、「優位項目」での●印は差別化項目とみなすことができる。他方、「問題項目」での●印は低い評価に強く影響する項目になる。なお、発話データのため、該当項目の個数はサンプル間に大きな差のあるときのみ意味を持つと考える。

まず、サンプル A とサンプル V が、「優位項目」の数から優れた項目と差別化項目が多いことが示されている一方で「問題項目」の数は少ない。これが図 9.2 の実験後の高い評価を裏付けている。他方、サンプル H は「優位項目」はゼロで、「問題項目」の数が多く、図 9.2 の低い評価を裏付ける結果となった。特に、録画データによれば、被験者のほぼ全員が、サンプル H にだけ採用されている「スライド式のボタン」を、押しボタンと間違えたとの指摘があった。この項目が、図 9.2 の一致した低い評価の要因と推定することができる。

そして、サンプル S は優れた項目と差別化項目が多いにも関わらず、「問題項目」でも、最も該当数が多い。このことが図 9.2 の実験後の低い評価とバラツキの大きさを裏付ける結果である。最後に、サンプル P の場合は、両者に独自項目もなく、「問題項目」の方が多数は多いものの、その問題点も、他社と共通しているため、欠点が顕在化しにくく、図 9.2 の高い評価になっていると考える。

表 9.3 優れた項目と問題項目

優れた項目		A	P	H	S	V
画面	画面サイズが大きい	○				○
	画面の情報量が多い	○				○
	カラー液晶	○			○	○
表示	GUI					●
	アイコンが理解できる	○	○		○	○
ボタン	ボタン数が少なく使いやすい	○	○			
	決定ボタンが大きい					●
電源	電源が分かりやすい	●				
	電源がONと同時に音楽再生				●	
機能	2種類以上でシャッフルができる	○			○	
	再生中にショートカットができる				●	
その他	慣れれば使いやすい	●				
	曲が選曲しやすい	○	○			○
	取説なしでも使用できる					●
	携帯電話に似ている					●

問題項目		A	P	H	S	V
画面	画面が小さい		○	○	○	
	画面の情報量が少ない		○	○	○	
	白黒画面		○	○	○	
表示	表示文字が小さい			○	○	
	英語表記が嫌い				●	
ボタン	電源ONボタンが分からない		○			○
	電源OFFボタンが分からない	○	○			○
	決定ボタンが分からない				●	
	クルクルの操作がやりにくい	●				
	スライド式のボタン			●		
	ボタン操作に統一感がない			○	○	
	ボタン操作がしにくい			●		
ボタン操作した実感がない		○	○			
メニュー	メニュー画面が分からない				●	
	メニュー画面への行き方が分からない				●	
機能	選曲の検索が分からない				●	
	早送り再生の時に微妙な調整ができない				●	
	曲の選曲が一曲づつしかできない		○			
その他	現在地が分からない			○	○	
	画面構造が分からない			○	○	
	電源が落ちるスピードが早い				●	

9.4 事例実験に基づく「使いやすさ」の要因分析と考察

(1) 視覚的な使いやすさの考察

まず、図 9.2 に示されたテスト前の各サンプルの視覚的な使いやすさのデザイン的な要因についてラフ集合を用いて分析する。ユーザビリティテストの被験者の発話データをもとに、実験者が被験者の認知している箇所を抽出して、表 9.5 の左側に示す認知部位にまとめた。この認知部位は形態要素（表 9.5 のアルファベットの属性値）で構成されている。そして、結論は 5 つのサンプルを決定クラス（サンプル A=1、サンプル P=2、サンプル S=3、サンプル H=4、サンプル V=5）とした。以上をもとに決定表にしたのが表 9.6 の上段の表である。

表 9.5 2種類の認知部位

ユーザビリティ評価の前			ユーザビリティ評価の後				
全体形状	本体形状	縦長四角 横長四角 横型構型	A1 A2 A3	文字	メニュー時の 文字表示	GUI+文字 文字のみ	A1 A2
	サイズ	標準 小	B1 B2		GUIの大きさ	大 小 なし	B1 B2 B3
	厚さ	厚い 標準 薄い	C1 C2 C3		文字表記	日本語表示 英語	C1 C2
画面	画面サイズ	標準 小	D1 D2	操作中の 文字サイズ	大 標準	D1 D2	
	画面の向き	縦型 横型	E1 E2	情報量	ディスプレイ 内の情報量	多い 標準 少ない	E1 E2 E3
	画面カラー	カラー 白黒	F1 F2		選曲時の情報量	1段 2段 4段 6段	F1 F2 F3 F4
メインボタン	メインボタン位置	真中 右側	G1 G2	操作音	操作毎の 操作電子音	あり なし	G1 G2
	メインボタン形状	丸系 四角 特殊	H1 H2 H3		ボタンの操作音	あり なし	H1 H2
	メインボタンの操作	押しボタン 押しボタン+回転 スライド ステック	I1 I2 I3 I4	機能	自動再生	あり なし	I1 I2
ボタン	ボタン数	多い 少ない	J1 J2		ショートカット	あり なし	J1 J2
	決定ボタン サイズ	標準 小	K1 K2	検索機能の数	多い 少ない	K1 K2	
	ボタンの 文字サイズ	標準 小	L1 L2	その他	スクロールバー の有無	あり なし	L1 L2
			自動OFFまで の時間		長い 短い	M1 M2	
			画面にHold表示		あり(全ボタン) あり(一部ボタン) なし	N1 N2 N3	

(視覚的な使いやすさ)

表 9.6 の上段の決定表からラフ集合の計算により求めたものが、図 9.4 の左側に示すサンプル毎の決定ルールである。例えば、サンプル A を見ると、上から独自の属性値の「I2: 押しボタン+回転」と、「C3:全体形状の厚さが薄い」かつ「B1:全体形状のサイズが標準」の属性値の組み合わせ等が、サンプル A の個別の特徴であることが示されている。単独の属性値や属性値の組み合わせが決定ルールである。なお、決定クラスに寄与している貢献度合いを示す指標である CI 値については、各サンプルの各決定ルールは「1」となるので、図 9.4 の左側では省略してある。

一方、同様の考え方で考察すると、順位の低いサンプル H とサンプル S、サンプル V では「J1: ボタン数の多い」が、順位が高い 2 社にない属性値である。特に、サンプル V は、「J1」の値が最も高いので、この特徴から第一印象が悪くなってしまい、順位が最も低くなったのではないかと考察される。次に、「J1」のように、この 3 社のすべてに共通する属性値ではないが、「C1: 全体形状の厚さが厚い」と「E2: 画面の向きが横型」、「G2: メインボタン位置が右側」が評価を下げる要因となったと考察される。特に、サンプル H とサンプル S の「G2」は値も高く評価を下げた要因と考えられる。また、「C1」ではサンプル V が特に高いので、前述の「J1」と「C1」がこのサンプルの評価を最下位にする属性値であると考えられる。

さらに、サンプル S に注目すると、「A3: 本体形状が横型樽型」と「C2: 全体形状の厚さが標準」、「H3: メインボタン形状が特殊」、「I4: メインボタンの操作がステック」、「L2: 文字サイズが小さい」がこのサンプルだけの独自の特徴となっている。サンプル S 全体の評価の低さを考え合わせると、この特徴が視覚的使いやすさを阻害しているものと考えられる。

ところで、図 9.4 の決定ルール数を見ると、サンプル A が極端に少ないことが見て取れる。特徴を示す数が少ないということは、サンプル A の特徴が明確であると言える。つまり、他のサンプルとのはっきりした差別化がなされている。この点でも、サンプル A が最も視覚的に使いやすく感じられる結果を説明することができよう。

Y= 1	CI値	Y= 2	CI値	標準化 コラムスコア	
A1	C1	1	C2	0.5	
A1	J2	0.67	E2	0.5	
A1	K2	0.67	J1	B2	0.5
A1	M1	0.67	J1	D2	0.5
A1	I2	0.67	J1	F1	0.5
G1	C1	0.67	J1	G1	0.5
G1	M1	0.67	J1	H2	0.5
G1	J2	0.67	J1	I1	0.5
N2		0.67	J1	L2	0.5
E1		0.67	K1	B2	0.5
H1		0.67	K1	F1	0.5
L1		0.67	M2	B2	0.5
			M2	D2	0.5
B1		0.33	M2	F1	0.5
D1		0.33	M2	G1	0.5
F3		0.33	M2	H2	0.5
C1	K1	0.33	M2	I1	0.5
(中省略)					
(中省略)					
			N3	B2	0.5
			N3	E3	0.5
			N3	G2	0.5
			N3	H2	0.5
			N3	K2	0.5
			N3	L2	0.5
N3	A1	0.33			
N3	G1	0.33			
N3	K1	0.33			
(総数: 47)					
(総数: 46)					
			A1	0	
			A2	0.32	
			B1	0	
			B2	0.96	
			B3	0	
			C1	0	
			C2	0.32	
			D1	0	
			D2	0.59	
			E1	0	
			E2	0.32	
			E3	0.48	
			F1	0.48	
			F2	0.32	
			F3	0	
			F4	0	
			G1	0.32	
			G2	1.28	
			H1	0	
			H2	0.96	
			I1	0.48	
			I2	0.96	
			J1	1.28	
			J2	0.27	
			K1	1.12	
			K2	0.48	
			L1	0	
			L2	0.96	
			M1	0.27	
			M2	1.28	
			N1	0.32	
			N2	0	
			N3	0.96	

図 9.5 表示画面デザインの分析結果

(2) 表示画面の使いやすさの考察

次に、テスト後の分析として、情報機器で重要となる分かりやすい表示画面の分析と考察を行った。そのためにまず、表示画面の使いやすさに貢献する属性値をビデオの録画データから作成した。その認知部位が表 9.5 の右側である。

そして、図 9.2 の下側から、結論部は操作性評価の高い群（1 群：サンプル P、V、A）と低い群（2 群：サンプル S、H）に分けた決定クラス（Y）とした（表 9.6 の下側）。この表 9.6 の下側の決定表をもとにラフ集合の計算を行った結果が図 5 である。

図 9.5 から、高い評価の 1 群（Y=1）の CI 値が最大「1」では「A1: GUI (Graphical User Interface) + 文字」かつ「C1: 日本語の文字表記」の組み合わせの決定ルールが求められた。また、CI 値が「0.67」の決定ルールの中でも、その「A1」の属性値が共通属性値（図 9.5 のアルファベットの四角枠）となっている。同様に、「G1: 操作電子音があり」の属性値も共通属性値となっている。そして、強い特徴となる独自の属性値では、「N2: 画面に Hold 表示が一部あり」と「E1: ディスプレイ内の情報量が多い」、「H1: ボタンの操作音があり」、「L1: スクロールバーがあり」が求められている。この結果から、「日本語文字表記を用いた GUI で、各種の電子音やスクロールバーがあり、ディスプレイ内の情報量が多い表示画面デザイン」が使いやすいことが示された。

一方、図 9.5 の低い評価の 2 群（Y=2）を考察すると、CI 値はすべて同じ値であるので、つまり、高い CI 値に焦点をあてることができないため、決定ルール分析法を用いて考察を行う。その分析結果が図 9.5 の右側である。

図 9.5 左側の四角枠に囲まれた共通属性値を示す標準化コラムスコアの「0.96」以上の値を求めると、「B2、G2、H2、I2、J1、K1、L2、M2、N3」となる。この結果から、「GUI の大きさが小さく、また、操作音や機能、その他のアイテムの中の属性値がなしのもの」が分かりにくい表示画面デザインと考察できる。また、独自の属性値としては、「A2、C2、E2、F2、N1」が求められた。この結果から、特に、メニューが英語の文字のみで、選曲時の情報量が 2 段である表示画面デザインが分かり難いと考察できる。

9.5 まとめと今後の課題

本提案手法では区間 AHP 法で求めたサンプルの使いやすさの順位をもとにして、ユーザビリティテスト結果から、サンプルの優位項目と問題項目を抽出した。つまり、従来のユーザビリティ評価手法とは異なり、順位を評価基準にすることで、問題項目と優位項目の両方を抽出することが可能となった。なお、解析で得られた結果も記録した発話データでも述べられていた内容と符合するところが多く納得できるものであった。

また、従来の統計的な手法では少ないサンプルの因果関係（逆問題）を解くことはできなかったが、ラフ集合を用いることで、少ないサンプルでの使いやすさの要因分析を行なった結果から、どのような操作に関する内容が使いやすさに寄与しているかがある程度は定量的に明らかになり、使いやすいインタフェースデザイン開発上の知識を得ることができた。

今後の展開としては、デジタルプレーヤだけでなく、デジタルカメラや携帯電話機を始めと

する多くの家電製品に対して同様の調査・実験を行い、実験手法の改良を行っていききたい。また、ラフ集合理論は、統計学とは異なり検定の考え方はないので、データ解析の信頼性の検討については今後の検討課題である。

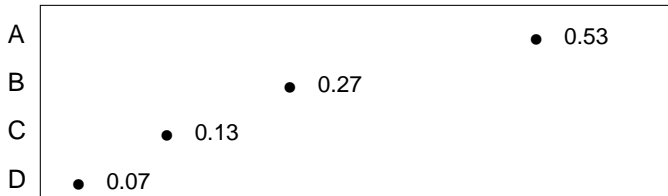
注と参考文献

- [1] 黒須正明、伊東昌子、時津倫子：ユーザー工学、共立出版、pp.1-29、1999
- [2] 三菱電機（株）デザイン研究所編：こんなデザインが使いやすさを生む、pp.14-64、2001
- [3] ヤコブ・ニールセン著、三好訳：ユーザビリティエンジニアリング原論、東京電機大学出版局、pp.134-135、1999
- [4] Z.Pawlak : Rough sets, International Journal of Information Computer Science, Vol.11,No.5, pp.341-356, 1982
- [5] K.Sugihara, H.Tanaka : Interval evaluations in the analytic hierarchy process by possibility analysis, Computational Intelligence, Vol.17,No.3, pp.567-579, 2001
- [6] 田中英夫、円谷友英、杉原一臣：意思決定における区間評価手法（解説）、知能と情報（日本知能情報ファジィ学会誌）Vol.17, No.4、pp.406-412、2005
- [7] 人間の主観的であいまいな感覚判断を用いて、各評価項目に対する重視度（ウェイト）を求める方法で有名なのが Saaty の提案する AHP 法である。この手法では、評価者が与える一対比較データに基づき評価基準のウェイトが求められる。その代表的な手法が固有ベクトル法である。一対比較行列から最大固有値を求め、その固有ベクトルを正規化したものを各評価項目のウェイトとして推定する方法である。この従来の AHP 法では、一対比較行列の不整合性がウェイトに反映されない。ただし、固有ベクトル法では不整合性を表す指標があり、この指標が大きければ、それを評価者の矛盾とみなし、評価者に再度一対比較を行ってもらふ必要がある。

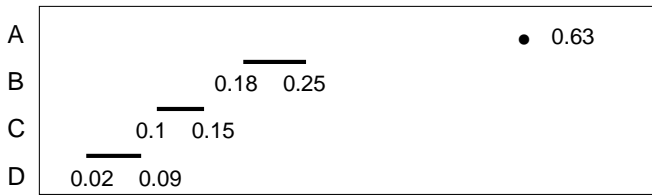
本論文で用いた田中英夫の提唱する区間 AHP 法は、一対比較行列の不整合性がウェイトに反映される手法である。従来の AHP 法は、各項目のウェイトがクリスプな値である一対比較値を仮定している。つまり、項目間の関係を強推移性また半推移性とみなしている。しかし、区間 AHP 法では、ウェイトを区間値とすることで、非推移性まで取り扱うことができる。

上記の強推移性と半推移性、非推移性の違いを例題で次に示す。まず、強推移性の場合、A から D までの各倍率の関係に整合性がある。したがって、従来 AHP 法と区間 AHP 法のウェイト値は同じになる。半推移性の場合、A から D までの大小関係はそのまま、各倍率の関係に整合性がない。したがって、両者のウェイト値は異なり、各倍率の不整合関係を区間 AHP 法では区間で表現される。例題の A と D の直接的な関係が逆（矛盾）になっている非推移性の場合、区間 AHP 法では交差する区間になり、例えば、全体としては $D < A$ であるが、D の右端と A の左端では、 $A < D$ となる。このように、矛盾関係が区間で視覚的に表現される。

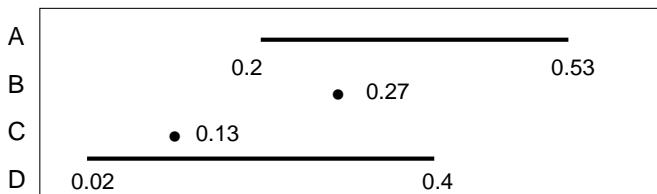
強推移律					区間		
	A	B	C	D	AHP	下端W	W上端
A		2	4	8	A	0.53	0.53
B			2	4	B	0.27	0.27
C				2	C	0.13	0.13
D					D	0.07	0.07



半推移律					区間		
	A	B	C	D	AHP	下端W	W上端
A		3	5	7	A	0.55	0.63
B			2	3	B	0.19	0.18
C				9	C	0.19	0.1
D					D	0.06	0.02



非推移律					区間		
	A	B	C	D	AHP	下端W	W上端
A		2	4	-2	A	0.33	0.2
B			2	4	B	0.31	0.27
C				2	C	0.16	0.13
D					D	0.2	0.4



他方、算出方法の考え方は、まず、推定区間のウェイトを、 $W_i = [w_i, \bar{w}_i]$ で表す。ただし、 w_i は区間ウェイトの下限で、 \bar{w}_i は区間ウェイトの上限である。いま評価項目が n 個あるとし、 $n \times n$ の一対比較行列を $A = [a_{ij}]$ と表す。 a_{ij} を近似する推定区間を、区間ウェイトを用いて次のように仮定する。

$$\forall I, j (i \neq j) \quad W_{ij} = [\underline{w}_i / \underline{w}_j, \bar{w}_i / \bar{w}_j]$$

そして、以下の3つの仮定を行い、ニューラルネットワーク等と同じ模索型解法である線形計画問題で計算する。

- (1) 与えられた一対比較値 a_{ij} は一対比較の推定区間 W_{ij} 内に存在する。
- (2) 推定される区間ウェイトベクトルは正規化条件を満足する。
- (3) 推定される区間ウェイトの幅はあいまいさを表現しており、出来る限り少ないほうがよいと考えられる。したがって、区間ウェイトの幅の合計を最小にする最適化問題を考える。

なお、仮定(1)と(2)が制約条件で、仮定(3)が目的関数になる線形計画問題である。

[8] 森典彦、田中英夫、井上勝雄編：ラフ集合と感性、海文堂出版、pp.3-50、2004