

動画像を用いた顧客満足度推定における3D-CNNの有効性検証

Validity validation of 3D-CNN in customer satisfaction estimation using moving images

中野 智文 *¹ 加藤 昇平 *¹ *²
Nakano Tomofumi Kato Shohei

*¹ 名古屋工業大学工学研究科情報工学専攻

Dept. of Computer Science and Engineering, Graduate School of Engineering, Nagoya Institute of Technology

*² 名古屋工業大学情報科学フロンティア研究院

Frontier Research Institutes for Information Science, Nagoya Institute of Technology

A customer satisfaction survey is conducted in order for companies and public entities to find business problems. However, it is a burden for investigators to devise questionnaire items and to sum up them, and it is a bother for customers to answer the questionnaire. There is also a problem that the collection rate of the questionnaire is low. I would like to estimate customer satisfaction by sensing technology to solve these problems. In this research, we hypothesize that satisfaction will appear in human facial expression and body movement, and we validate the hypothesis by applying 3D-CNN to a moving image.

1. はじめに

顧客満足は企業にとって非常に重要な概念であり、企業は自社製品およびサービスに対する顧客満足度を定期的に調査することが求められている。顧客満足度をアンケートなどの調査によって数値化し、客観的に評価できる指標としたものを顧客満足度と呼ぶ。企業はこの顧客満足度を調査することで業務上の課題を見つけ、新製品開発やサービス向上施策を実施する。しかし、顧客満足度調査は容易ではない。法人側にはアンケート項目の考案から、配布、回答依頼、集計などの負担とそのためコストが生じ、顧客側には回答の手間が生じる。さらに、アンケートを実施しても回収率や即時性が低いといった問題点もある。

こうした問題を解決すべくセンシング技術によって顧客満足度を推定した先行研究が存在する。松川ら [松川 09] は、スポーツを観戦している複数の観客の顔の向きや表情から満足度を推定するシステムを提案した。しかし、こうした満足度のみを出力するシステムでは、どこをどう改善していけばいいかが分かりづらく、顧客満足度調査の本来の目的である業務の改善に役立てることが難しい。

そこで本研究では、知覚矯正仮説 [Howard 73] という顧客満足度に関する理論仮説の考えを盛り込み、満足度に加えて期待感もセンシングすることで、適切なマーケティングがなされているかどうかを評価するシステムを開発する。このシステムによって、顧客満足度を測る際に企業側にかかるコストや顧客側の手間を軽減するとともに、業務施策の考案を補助する。将来的に飲食店や小売店などで利用することを考えると、光学映像を用いることがシステムの導入をもっとも容易かつ安価にすると考えた。そこで本研究では、人の表情や体動に期待感や満足感が表れるという仮説を立て、動画像に3D-CNN [Ji 13] を適用することでその有効性と仮説を検証した。

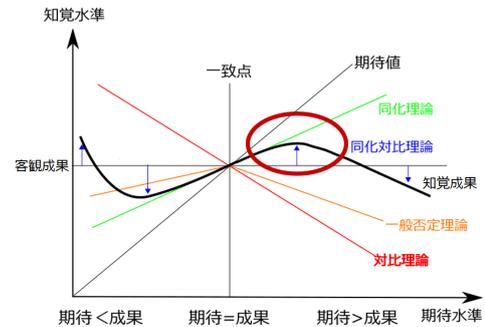


図 1: 知覚矯正仮説概念図 [嶋口 81] p.57 を修正

2. 知覚矯正仮説

顧客満足度は行動前に抱かれる期待感と行動後に知覚された成果のギャップにより決定されると言われている [上原 09]。顧客満足度と期待と成果の関係について、4つの理論仮説が考えられている。4つの理論仮説はそれぞれ「同化理論」「対比理論」「同化対比理論」「一般否定理論」と呼ばれ、まとめたものを知覚矯正仮説と言う。

同化理論とは、顧客が抱いた期待と知覚した成果にギャップが存在するとき、認知的不協和を解消すべく、成果が期待に寄って知覚されるというものである。

対比理論とは、上記のギャップに対して、成果が期待からより際立つように知覚されるというものである。期待が成果を上回ったとき（負の不一致）は不満、成果が期待を上回ったとき（正の不一致）は満足となる。

同化対比理論とは、同化理論と対比理論にはある程度の範囲が存在し、期待と成果のギャップが小さいとき顧客満足度は同化理論により決定され、ギャップが大きいときは対比理論により決定されるという考え方である。

一般否定理論とは、ギャップの正負に関わらず、すべての成果を知覚に否定的に作用するという考え方である。

以上をまとめると、知覚矯正仮説は図1のような概念図で表される。縦軸は知覚水準であり、期待感および知覚成果の大き

連絡先: 加藤昇平, 名古屋工業大学, 名古屋市昭和区御器所町,
052-735-5625, shohey@katolab.nitech.ac.jp

さを表す。右肩上がりの黒線は期待感の大きさを表しており、横軸と水平な黒線が一定である客観成果を表している。図中央の一致点で期待と客観成果は等しくなり、それよりも横軸正の位置で「期待>客観成果」、負の位置で「期待<客観成果」となっている。緑、赤、黄色の実線はそれぞれ同化作用、対比作用、一般否定作用を表しており、波線で描かれた知覚成果はそれらに一定の割合で引き寄せられている。また、青の矢印4本は、それぞれの期待と客観成果の数値関係における同化-対比作用を表している。期待が実際の成果に比べて著しく大きい場合、対比作用によって知覚成果が下がり、期待が実際の成果に比べて若干大きい場合、同化作用によって知覚成果が上がる様子が表されている。また、期待が実際の成果に比べて若干小さい場合、同化作用によって知覚成果が下がり、期待が実際の成果に比べて著しく小さい場合、対比作用によって知覚成果が上がる様子が表されている。

企業は図1に示す赤丸の部分を目指すべきである。実際よりも成果が低く知覚されてしまう右端や左から2つ目の矢印の部分は損と言える。また、著しく期待感が低い左端の矢印の位置では、たとえ知覚成果が大きくなるとしても製品やサービスを購入してもらえないという問題が生じるため利益に繋がらない。よって企業は実際の成果よりも若干高めの期待を抱かせることで、消費者の購買意欲を掻き立てるとともに同化作用による知覚成果の引き上げを図るべきである。そして期待感および知覚成果の双方を把握することは、マーケティングが適切になされているかどうかを評価することに繋がると言える。

3. 予備実験

3.1 実験手順

知覚矯正仮説の検証および学習用データ獲得のため、以下のような実験を行なった。Jacobyら [Jacoby 71] がビールを用いて行った実験を参考に、あるりんごジュースに対して段階的な価格(90円から190円の間10円ごとの11項目)を設定する。商品名を隠した同じ成果物を、異なる価格で提供して期待感を操作した。実験協力者に期待感および満足感を評価させ、その結果を比較することで知覚矯正仮説を検証する。なお期待感および満足感推定器の学習データとするために、実験の様子をビデオカメラで撮影した。実験の手順を以下に示す。

1. 温度設定を一定にした冷蔵庫でりんごジュースを容器ごと1日以上冷やしておく。
2. 実験室の空調を25.5度に設定し、30分以上時間を置く。
3. 実験協力者に満足するまで水を飲んでもらい、その後実験終了までの飲食と激しい運動を控えてもらう。
4. 実験協力者は水を飲み終えてから30分後に実験者とともに実験室に入室する。
5. 実験協力者は健康状態等に関するアンケートに回答する。
6. 実験者は期待感および満足感の記入方法を説明する。
7. 実験者はりんごジュース50mlを実験協力者の前に置く。
8. 実験者は280mlペットボトルを見せつつ価格を呈示する。
9. 実験協力者は色と香りと価格を参考に期待感を記入する。
10. 実験協力者はりんごジュースを飲む。
11. 実験協力者はりんごジュースを飲み終えた後に満足感を記入する。

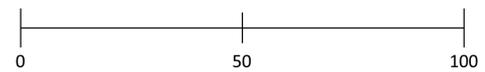


図2: 評価用グラフィックツール

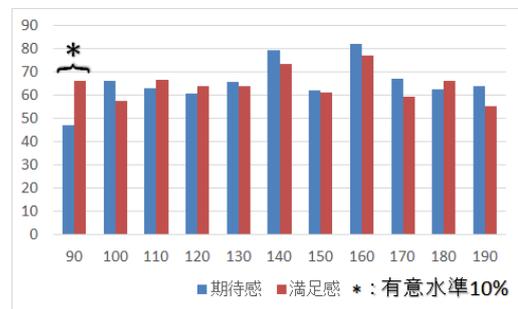


図3: 実験結果

期待感と満足感の評価には図2に示すグラフィックツールを用いた。このグラフィックツールは左ほど「不味そう」もしくは「不味い」場合を表し、右ほど「美味しそう」もしくは「美味しい」ことを表す。中央の50点はどちらでも無い場合である。実験協力者は該当箇所に縦線を記入する。この評価方法も実験方法同様Jacobyら [Jacoby 71] の実験を参考にした。

3.2 実験結果

各項目実験協力者5名ずつで平均を取った。結果を図3に示す。縦軸が期待感と満足感の大きさ、横軸はそれぞれの呈示価格を表している。呈示価格に伴って期待感が増加することを期待したが、そうはなっていないことが見て取れる。考えられる原因として、色や香りといった情報による期待感への影響が予想以上に大きかったことがあげられる。Changら [Chang 96] は、商品に関する価格以外の情報が多いほど価格による品質推定への影響力が小さくなることを示している。飲む直前に期待感が変わってしまうことを防ぐために事前に色や香りを確認させたことが仇となり、個人の主観や誤差といったものが大きく出てしまったことが考えられる。

しかし、最低呈示価格である90円の項目においては、他の項目と比べても期待感が著しく低いことが見て取れる。90円という呈示価格が、実験協力者の品質推定に影響を及ぼすほど低いものであったと考えられる。90円の項目に関して期待感が極めて低いために対比作用が働き、それ以外のすべての項目に関しては、満足感が期待感に沿う同化作用が働いているように見受けられる。実際に、呈示価格90円の項目に関してはt検定によって有意水準10%で有意差が見られ、それ以外の項目では有意差が見られなかったことがこれを裏付けている。しかし、図1における著しく期待感の高い領域が確認できなかったため、知覚矯正仮説全体を確認するには至らなかった。

4. 期待感と満足感の推定

本節では人の表情や体動に期待感および満足感が表れるか検証する。3.節の予備実験にて収集した動画データを用いて期待感と満足感を推定する実験を行った。

4.1 3D-CNN

畳込みニューラルネットワーク (Convolutional Neural Network, 以下CNN) は、隣接層間の特定のユニットのみが結合している

特別な層を持つ順伝播型ネットワークである。その構造は人間の視野の仕組みを模しており、画像認識において高い精度を示すことで知られている。通常の CNN が 2 次元画像を入力とするのに対し、Shuiwang ら [Ji 13] は縦横奥行き の 3 次元配列データを入力とする 3 次元畳込みニューラルネット（以下、3D-CNN）を提案した。この奥行きを時間と見なすことで、CNN を動画に応用することが可能となる。本研究では、3. 節の実験記録動画を入力データ、評価値を 5 段階評価に変換したものを目標出力として、3D-CNN によるクラス分類を試みた。学習方法としては Dropout [Srivastava 14] を用いた確率的勾配降下法によるミニバッチ学習を採用した。

4.2 事前準備

4.2.1 入力データ

3D-CNN の入力データとするためには、縦横の画像サイズだけでなく時間長も固定しなければならない。しかし人によって飲むスピードが異なるにも関わらず、無理やりリサイズすることは動画中の人物の動作スピードが変わってしまうこととなり、期待感や満足感と関係のある特徴量が変化してしまう危険性がある。そこで、期待感および満足感に対応する入力データとして、記録からそれぞれ一部を切り出すこととした。期待感に対応するデータはコップに口をつける瞬間までの直前 2 秒間、満足感に対応するデータはコップに口をつけた瞬間の直後 5 秒間である。フレームサイズ 320pixel × 240pixel、ビデオレート 15fps である期待感と満足感の動画から、実験協力者が写っている正方領域 160pixel × 160pixel をトリミングして入力データとした。図 4 に満足感の入力データの例を示す。

4.2.2 データ拡張

記録した動画は全部で 83 回分であり、3D-CNN の入力とするにはあまりに少ない。そこで、入力データに対して以下のような幾何学的データ拡張を行い、データ数を 4 倍に増やした。変形量はデータごとにランダムである。

- ±4 度以内の回転
- 左右上下に 2% 以内の平行移動
- 10% 以内の拡大縮小
- 50% の確率で左右鏡像反転

これによってカメラの傾き具合や実験協力者との位置など、期待感および満足感に関係がないと思われる特徴量を排除するとともに、データ数の充実を図った。また、重み更新のたびに訓練データに対してのみ同様のデータ拡張を行った。訓練データが更新ごとに違ったものになるため、実質的な訓練サンプル数が増えることとなり、過学習となるリスクを減らすことができる。

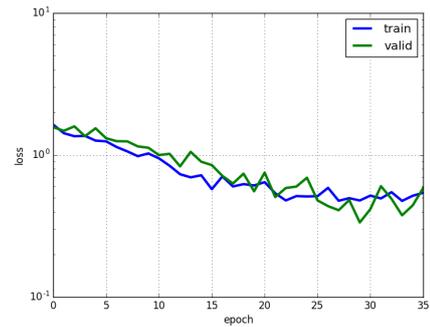
4.3 実験

4.3.1 学習進行の確認

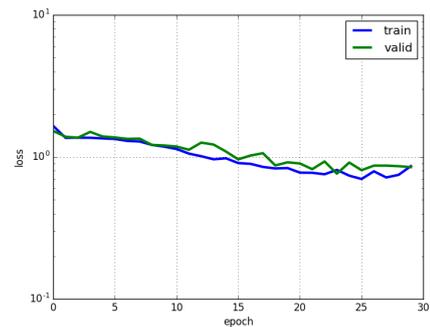
入力データのうち、ランダムに 9 割を訓練データ、残りの 1 割を検証データとして学習を行った。ここで訓練データとは、実際に重みの更新に用いられるデータである。また検証データとは、重みの更新には用いられず、重みの更新が終了する度に推定器に適用することで汎化性能を検証して学習を続けるかどうかを判断するためのデータである。本研究では検証誤差の改善が見込まれなくなった時点で学習を切り上げた。期待感、満足感ともに推定器の学習を行ったところ、図 5 に示す学習曲線が得られた。この図から、検証誤差が訓練誤差に伴って下がっていていることが確認できる。また、期待感や満足感と関係のある特徴をある程度捉えられていることが分かった。最終的な検証結果は表 1 のようになった。



図 4: 入力データの例



(a) 期待感



(b) 満足感

図 5: 誤差の学習曲線

4.3.2 leave-human-out

4.3.1 節の学習では、検証データによって学習の進行度を判断しているにも関わらず、最終的な誤差や正答率を検証データから算出している。これは未知のデータに対する汎化性能とはいえない。そこで訓練データ、検証データとは別にテストデータを用意する、leave-human-out によって推定器の汎化性能を検証した。leave-human-out の手順を以下に示す。

1. データ中のある人物のサンプルを取り出してテストデータとする。
2. 残ったデータの 9 割を訓練データ、1 割を検証データとして学習を行う。
3. 学習終了後、テストデータを適用してテスト誤差とテスト正答率を出力する。
4. 1.~3. を人数分繰り返す。
5. テスト誤差とテスト正答率の総和をそれぞれ計算し、人数で割ったものを汎化性能とする。

表 1: 推定結果

	誤差	正答率
期待感	0.59	78.8%
満足感	0.85	69.7%

表 2: leave-human-out の結果

	誤差	正答率
期待感	1.67	27.7%
満足感	1.51	31.3%

結果を表 2 に示す。4.3.1 節と比べて、芳しくない結果となっている。図 6 に leave-human-out 中の学習曲線の一部を示す。この図から、訓練誤差すら低下しておらず、学習がまったく進んでいないことが分かる。他のサンプルにおいても、おおむね同様であった。この原因として学習係数と訓練データの依存性が挙げられる。この依存性のために、4.3.1 節において手動で調整した学習係数のままでは、訓練データが学習の度に変わる leave-human-out でうまく学習が進まないことが考えられる。この問題は、Adam[Kingma 14] などの学習係数を自動で調整するような学習方法によって解決されると考えられるため、今後は学習方法をより適切なものに切り替えることで汎化性能の改善を図る。

5. 考察と今後の展望

本研究では、人の表情や体動に期待感や満足感が表れるという仮説を立てたが、4.3.2 節の結果から、仮説の検証はまだ不十分といえる。しかし 4.3.1 節において、訓練誤差の減少にしたがって検証誤差が下がっていったことから、消費行動の外見的特徴に期待感や満足感が表れる可能性と、3D-CNN の有効性が示唆された。

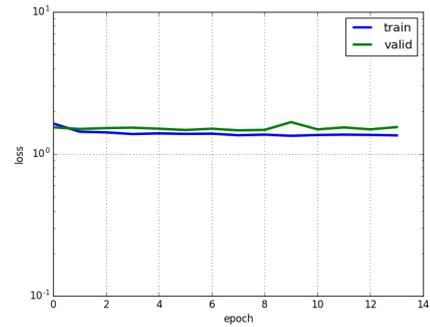
今後は 4.3.2 節で述べた学習係数を自動で調整する手法を取り入れることで、汎化性能の向上を目指す。また、今回は 3D-CNN の特性から入力データの時間成分を固定長で取り出した。これにより期待感と満足感に関係のある一部の特徴量が失われていることが考えられる。これを防ぐために、3D-CNN と再起型ニューラルネットワーク、特に Long Short Term Memory を組み合わせることで自由長データを扱えるニューラルネットワークを構築する。さらに、このシステムが実際に使われるときのことを考えて、人が写っている領域を自動で切り出せるようにすべきである。トリミングだけでなく、OpenCV と組み合わせることで自動で背景を除去することも有効だと考えられる。

6. おわりに

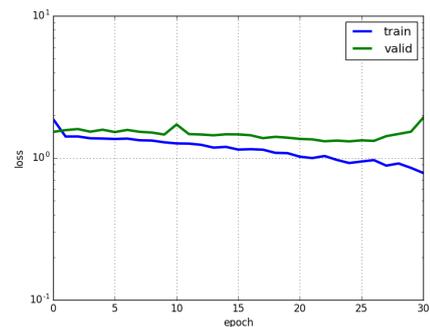
本稿では、人の表情や体動に期待感や満足感が表れるという仮説を立て、動画像に 3D-CNN を適用することでその有効性と仮説を検証した。また予備実験の結果から知覚矯正仮説を検証した。顧客行動を入力とした 3D-CNN による学習の有効性が示唆された。今後は上述した改善案を実装することで汎化性能の向上を目指す。

参考文献

[Chang 96] Chang, T.-Z. and Wildt, A. R.: Impact of product information on the use of price as a quality cue, *Psychology & Marketing*, Vol. 13, No. 1, pp. 55–75 (1996)



(a) 期待感



(b) 満足感

図 6: leave-human-out 中の学習曲線の一部

[Howard 73] Howard, J. A. and Hulbert, J.: Advertising and the public interest (1973)

[Jacoby 71] Jacoby, J., Olson, J. C., and Haddock, R. A.: Price, brand name, and product composition characteristics as determinants of perceived quality., *Journal of Applied Psychology*, Vol. 55, No. 6, p. 570 (1971)

[Ji 13] Ji, S., Xu, W., Yang, M., and Yu, K.: 3D convolutional neural networks for human action recognition, *IEEE transactions on pattern analysis and machine intelligence*, Vol. 35, No. 1, pp. 221–231 (2013)

[Kingma 14] Kingma, D. and Ba, J.: Adam: A method for stochastic optimization, *arXiv preprint arXiv:1412.6980* (2014)

[Srivastava 14] Srivastava, N., Hinton, G. E., Krizhevsky, A., Sutskever, I., and Salakhutdinov, R.: Dropout: a simple way to prevent neural networks from overfitting., *Journal of Machine Learning Research*, Vol. 15, No. 1, pp. 1929–1958 (2014)

[松川 09] 松川徹, 日高章理, 栗田多喜夫 他: 複数人物の顔方向・表情認識に基づく映像中の観客の満足度の自動定量評価, *情報処理学会論文誌*, Vol. 50, No. 12, pp. 3222–3232 (2009)

[上原 09] 上原聡: 顧客満足に対するサービス品質の影響に関する考察, *嘉悦大学研究論集*, Vol. 52, No. 1, pp. 1–15 (2009)

[嶋口 81] 嶋口充輝: 消費者満足構造の理論仮説とマーケティング戦略行動, *慶応経営論集*, Vol. 2, No. 3, pp. p43–69 (1981)