

料理写真の魅力度推定手法の改良 — 選好実験時の注視領域を反映した特徴抽出 —

佐藤 陽昇^{1,a)} 平山 高嗣^{1,b)} 高橋 和馬^{1,c)} 道満 恵介^{2,1,d)} 川西 康友^{1,e)} 井手 一郎^{1,f)}
出口 大輔^{1,g)} 村瀬 洋^{1,h)}

概要: 近年、料理レシピサイトや SNS の普及により Web 上への料理写真の投稿が増加している。Web 上に投稿される料理写真は美味しそうに撮影されていることが望ましいが、その撮影は必ずしも容易ではない。従来研究では、料理を美味しそうに撮影するための支援技術として、料理写真の魅力度、つまり撮影された料理が美味しそうに見える度合いを推定する手法が提案されている。この手法では、魅力度付きの料理画像群から画像特徴を抽出し、回帰の枠組みにより任意の料理画像に対して魅力度を推定する。本研究では、閲覧者が一般的に注目する領域を分析し、それに基づいて画像特徴を抽出する領域を選択することにより、料理写真の魅力度推定手法を改良した。まず、画像の選好を評価する被験者実験を行い、魅力度を判断する際の視線情報を計測した。次に、視線の停留状態が続いた領域、すなわち注視領域を分析し、それに基づいて画像特徴の抽出領域を選択した。そして、画像全体から特徴抽出する従来手法と比較した結果から、注視領域外から特徴抽出することが有効であることが確認された。

1. はじめに

近年、料理レシピサイトや SNS の普及により Web 上への料理写真の投稿が増加している。Web 上に投稿される料理写真は美味しそうに撮影されていることが望ましい。しかし、Web サイトや SNS に投稿される料理写真は、同一の料理でも美味しそうに見える度合いが様々である。例えば、図 1 は同一の料理を撮影した料理写真であるが、図 1(a) よりも図 1(b) の方が、写真全体に占める大きさや構図の点で料理が美味しそうに撮影されている。このように、被写体の大きさや撮影構図、色合いなどの違いによって料理写真に対する美味しさの印象が変わると考えられる。そのため、美味しそうに見える料理写真を撮影するためには、これらの要因を考慮して適切な撮影方法を選ぶ必要がある。

しかし、非専門家にとって撮影方法を迅速に決定することは必ずしも容易ではない。そのため、料理写真の撮影方法を推薦するシステムがあれば有用であると考えられる。



(a) 魅力に欠ける構図で撮影された料理写真 (b) 魅力的な構図で撮影された料理写真

図 1: 同一の料理を被写体とした料理写真の例。

そのようなシステムを実現するためには、まず撮影された料理が美味しそうに見える度合いを定量的に分析する必要がある。

我々は従来研究 [1] において、料理が美味しそうに見える度合いを「魅力度」と定義し、被験者が魅力度の対比較を行う選好実験に基づいて料理画像に魅力度を付与したデータセットを構築した。そして、料理画像の色彩調和、エッジの向きと強度、Convolutional Neural Network (CNN) から得られる特徴量である Deep Convolutional Activation Feature (DeCAF) [2] を特徴量として用いることで、料理写真の魅力度を推定する手法を提案した。この手法では、画像全体もしくは料理領域全体から均一に抽出される画像特徴で魅力度を推定している。本研究では料理写真の閲覧者が一般的に重視する領域に着目する。上記の選好実験に

¹ 名古屋大学 Nagoya University, Nagoya 464-8601, Japan

² 中京大学 Chukyo University, Nagoya 466-8666, Japan

a) satoa@murase.is.i.nagoya-u.ac.jp

b) hirayama@i.nagoya-u.ac.jp

c) takahashik@murase.m.is.nagoya-u.ac.jp

d) kdoman@sist.chukyo-u.ac.jp

e) kawanishi@i.nagoya-u.ac.jp

f) ide@i.nagoya-u.ac.jp

g) ddeguchi@nagoya-u.jp

h) murase@i.nagoya-u.ac.jp

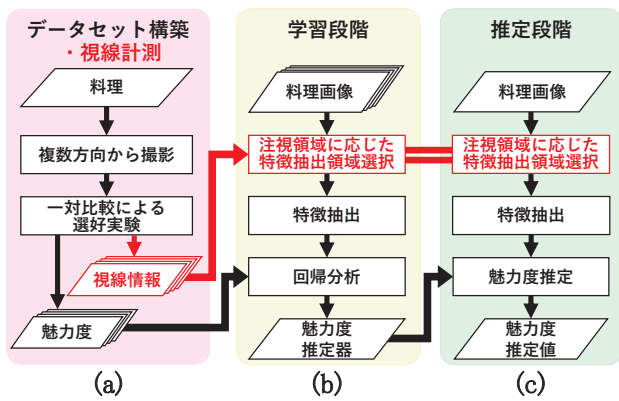


図 2: データセット構築と魅力度推定手法の処理手順. 赤字が従来研究 [1] に対する差分である.

においては、被験者が注目した領域が料理間で異なる可能性がある。つまり、一般的な注目領域を料理ごとに分析し、それに基づいて画像特徴を抽出することで、料理写真の魅力度の推定精度を向上させることができる可能性がある。

関連研究として、下条らは、左右に並べられた 2 枚の顔画像から好みに基づいて 1 枚を選択する選好実験において、被験者の視線配布が選択する方の画像に偏ることを示した [3]。この研究の工学的な応用を行った菅野らは、閲覧者の視線情報を用いて自然画像の選好を機械的に推定し、画像選好における視線情報の重要性を確認した [4]。また、尾崎らは、閲覧者の視線情報を利用して画像特徴を抽出する領域を制限することで画像の選好を推定し、注視点周辺の領域、すなわち注視領域に応じた特徴抽出により推定精度の向上を確認した [5]。

以上より、従来研究で定義した魅力度 [1] は一対比較による選好実験によって定められるものであることから、料理写真の魅力度推定精度向上のためには、一対比較時の視線情報を考慮することが有効であると予想される。そこで本研究では、被験者による画像選好時の注視領域に基づいて画像特徴を抽出することで従来の料理写真の魅力度推定手法を改良する。現状の技術では、料理写真の撮影時に視線を計測することが容易ではなく、提案手法は学習データセットの構築時にのみ視線情報を計測すれば良いため実用的である。

以降、2 章で視線情報に基づく画像特徴抽出による料理画像の魅力度推定手法について説明する。次に 3 章で被験者実験による魅力度の評価値と視線情報を付与した料理画像データセットの構築について述べ、4 章で提案手法の有効性を評価した実験について述べる。最後に 5 で本報告をまとめる。

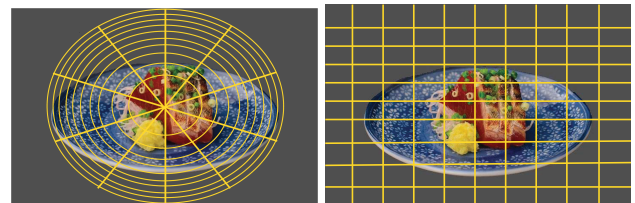
2. 視線情報に基づく画像特徴抽出による料理画像の魅力度推定手法

提案手法では、視線情報に基づいて画像特徴を抽出する



(a) 処理前 (b) 処理後

図 3: GrabCut による前景抽出の例.



(a) 放射状分割 (b) 格子状分割

図 4: 画像特徴抽出のための局所領域分割方法.

領域を選択する。図 2 に魅力度推定手法の処理手順を示す。まずデータセット構築では、被験者に料理画像の一対比較による選好課題を課し、その際の視線情報を計測する。視線情報は、料理画像上における注視点の座標の時系列情報として計測される。学習段階では、魅力度の評価値が与えられた料理画像群から、視線情報に基づいて選択した領域の画像特徴を抽出し、魅力度推定器を構築する。推定段階では、構築した魅力度推定器を用いて、魅力度が未知で視線情報が付与されていない料理画像に対して魅力度を推定する。以降、学習段階と推定段階の処理について詳述し、データセット構築については 3 章にて詳述する。

2.1 学習段階：魅力度推定器の構築

図 2(b) に示した処理手順により、魅力度推定器を構築する。まず、予め魅力度の評価値が与えられた料理画像について、皿を含めた前景領域を GrabCut [6] により抽出する。GrabCut とは、前景として抽出したい領域を大まかに囲む矩形を与えることにより、背景と前景を自動的に分離する領域分割手法である。抽出された前景画像の例を図 3 に示す。本研究では、皿を含めた料理の領域を料理領域と定義する。次に、視線情報を分析して、分割領域における視線の累積停留時間に応じて注視領域を検出する。これに基づいて、学習に利用する画像特徴抽出領域を設定する。

この一連の処理を全ての料理画像に適用し、各料理画像の全ての画像特徴抽出領域から抽出した画像特徴量から魅力度を回帰するパラメータを Random Regression Forest [7] を用いて学習することで、魅力度推定器を構築する。

2.1.1 色特徴量の抽出

一般的な写真の審美性を判定する際に、色彩調和が有効であるという知見 [8] がある。それに基づき、食材の色合

いや料理と皿の色の調和を考慮することが料理写真の魅力度推定に重要であると考えられる。そこで、色特徴として、 $L^*a^*b^*$ 色空間における色差の分布を利用する。 $L^*a^*b^*$ 色空間は、色空間上の2色の距離が人間の知覚する色の違いに近いとされており、色差を表すために最も多く使用されている表色系である。明度を L^* 、色相と彩度に相当する色度を a^* 、 b^* で表し、 a^* と b^* の値が大きくなるに従って色が鮮やかになる。

具体的には、まず料理領域全体から $L^*a^*b^*$ の3次元ヒストグラムを作成し、最頻出色 $G = (L^*, a^*, b^*)$ を計算する。ここで、 $L^*a^*b^*$ の各成分を0から7の8段階に量子化する。次に、図4(a)に示すように入力画像に対数極座標系を適用し、放射状に $10 \times 10 = 100$ 個の局所領域に分割する。このように分割することによって、皿の形状を考慮した食材の色配置を表現することができる。そして、各局所領域から最頻出色 $R_i = (\tilde{L}_i^*, \tilde{a}_i^*, \tilde{b}_i^*)$ およびその出現数 F_{R_i} を計算する。ここで、 i は局所領域の番号を表し、 $i \in \{1, \dots, 100\}$ 、 $\tilde{L}_i^*, \tilde{a}_i^*, \tilde{b}_i^* \in \{1, \dots, 8\}$ である。そして、 G と R_i の色差 D_i を最頻出色の出現数 F_{R_i} を考慮して次式で計算し、色特徴量として100次元ベクトル D を作成する。

$$D_i = F_{R_i} \sqrt{(L^* - \tilde{L}_i^*)^2 + (a^* - \tilde{a}_i^*)^2 + (b^* - \tilde{b}_i^*)^2} \quad (1)$$

$$D = (D_1, D_2, \dots, D_{100}) \quad (2)$$

2.1.2 形状特徴量の抽出

魅力的な料理写真の撮影には、盛り付けや料理の向きが重要であるという知見 [9] がある。それに基づき、料理の立体感や向きを考慮することが料理写真の魅力度推定に重要であると考えられる。そこで、形状特徴としてエッジの向きと強度を利用する。

具体的には、まず図4(b)に示すように入力画像を $10 \times 10 = 100$ 個の格子状の局所領域に分割し、料理領域内の各局所領域の最大エッジ強度 e_j とその勾配方向 n_j を乗じる。ここで、 j は局所領域の番号を表し、 $j \in \{1, \dots, 100\}$ である。格子状に分割することによって、最大エッジ強度とその勾配方向の位置関係を表現することができる。なお、勾配方向を計算する際に、皿を含めた料理領域から5画素内側の領域に対して処理を行う。これは、皿の縁に現れるエッジの影響を軽減するためである。そして、下記の100次元のベクトル M を作成し、形状特徴量とする。

$$M_j = e_j n_j \quad (3)$$

$$M = (M_1, M_2, \dots, M_{100}) \quad (4)$$

2.1.3 視線情報の分析による画像特徴抽出領域の設定

まず、視線情報に基づいて学習に利用する画像特徴量を抽出する領域を設定する。注視領域は、2.1.1項、2.1.2項で定義した各局所領域における視線の累積停留時間に基づいて検出する。ここで、視線の停留とは、ある一定箇所に

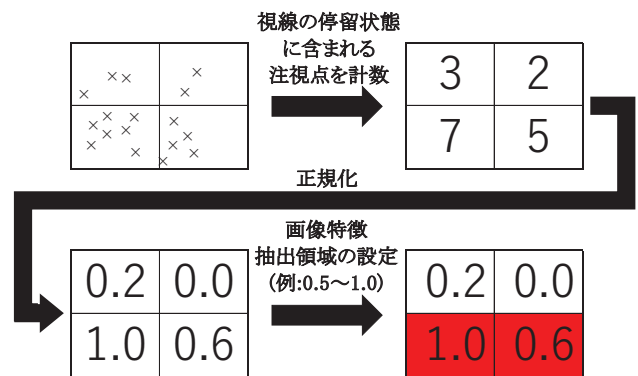


図 5: 画像特徴抽出領域の設定方法の例。

連続して注視点が留まっている状態を指す。本報告では、視線の移動角速度が 30 deg/s 未満である連続した2つの注視点を視線停留状態にあると定義する。視線の移動速度は2つの注視点座標のフレーム間差分により算出する。事前処理として、3章で述べる被験者実験を通して得られる注視点に関して、各時刻で視線の停留を判定し、停留と判定されなかった注視点を除去する。また、料理領域内に着目して分析するために、料理領域外にある注視点も除去する。

図5に具体的な処理の例を示す。まず2.1.1項、2.1.2項で述べたように入力画像を局所領域に分割する。次に、視線の累積停留時間を計算する。本報告では、3章で述べる被験者実験を通して得られる視線情報として、一定の間隔で注視点が計測されると仮定し、局所領域内で計測された視線停留状態に分類された注視点の数の合計を視線の累積停留時間とみなす。そして、得られた視線の累積停留時間の最小値が0、最大値が1となるように全局所領域にわたって正規化する。最後に、着目する視線の累積停留時間の範囲を設定することで画像特徴抽出領域を選択する。そして、回帰パラメータの学習に画像特徴抽出領域から抽出した画像特徴量を用いる。それ以外の領域から抽出した画像特徴量は使わない。

2.1.4 魅力度推定器の学習

提案手法では、2.1.3項で述べた方法で抽出した画像特徴量と魅力度の評価値との関係を学習することにより、魅力度推定器を構築する。すなわち、3.1節で紹介するデータセット内の各料理画像の魅力度を目的変数、2.1.3項で述べた方法で抽出した画像特徴量を説明変数として回帰モデルを学習する。その学習にはRandom Regression Forest [7]を用いる。

2.2 推定段階：魅力度の推定

図2(c)に示した処理手順により、料理画像の魅力度を推定する。入力画像に視線情報は付与されない。まず、入力された料理画像にGrabCut [6]を適用し、皿を含む料理領域を抽出する。そして、抽出された料理領域から2.1節で述べた方法で設定した画像特徴抽出領域のみから画像特徴

量を抽出する。その後、2.1.4項で構築した魅力度推定器を用いて魅力度を推定する。

3. データセットの構築

提案手法では、被験者に画像選好課題を課して計測した視線情報の分析結果に基づいて画像特徴抽出領域を設定する。そのため、料理画像の対比較による選好課題を被験者に課し、同時に被験者の視線を計測する。以降、実験用データセットの構築方法について詳述する。

3.1 魅力度付き料理画像データセット

本報告では、高橋らによる魅力度付き料理画像データセット [1] を使用した。このデータセットは、同一の料理について仰角と回転角を変更して36方向から撮影した料理画像群に対する選好実験の結果に対して、Thurstoneの対比較法 [10] を適用して算出した魅力度を付与したものである。

各料理画像は、3.2節で述べる視線計測実験において被験者に提示する画像として、また、4章で報告する評価実験において入力画像として使用する。また、各料理画像に付与されている魅力度は、4章で報告する評価実験において魅力度推定器の目標値として使用する。

以下、このデータセットの構築方法を紹介する。

3.1.1 対象料理

色合いや立体感の違いを考慮して5種類の料理が選ばれた。具体的には、鰹のたたき、カレーライス、鰻丼、ビーフシチュー、ハンバーグである。撮影の利便性と再現性の点から、時間経過に伴う状態の変化や盛り付けの変化が生じない食品サンプルが用いられた。

3.1.2 料理画像群

図6に各料理の画像群の抜粋を示す。撮影角度として設定された仰角は、撮影装置の回転皿と同じ平面を仰角0度とし、その面を基準に30、60、90度であった。回転角は、料理のある方向を基準として、その方向から右回りに30度刻みに330度までの角度であった。

3.1.3 魅力度の付与

高橋らは、Thurstoneの対比較法 [10] により、料理画像群の各料理画像に魅力度を付与した。Thurstoneの対比較法は官能検査の1つであり、対比較結果に基づいて複数の試料の感覚値を間隔尺度化するものである。

具体的には、まず、各料理画像群の料理画像36枚に対して、異なる2枚の組み合わせ ${}_{36}C_2 = 630$ 通りの料理画像対を生成した。次に、全ての組み合わせに対して各々3人以上の被験者から回答が得られるように対比較による選好実験を行った。その際に、被験者は「美味しく見える方はどっち?」という設問に対して、「左」または「右」と回答し、判断できない場合には「分からない」と回答した。本データセットを構築した際には、20代の男女延べ20名

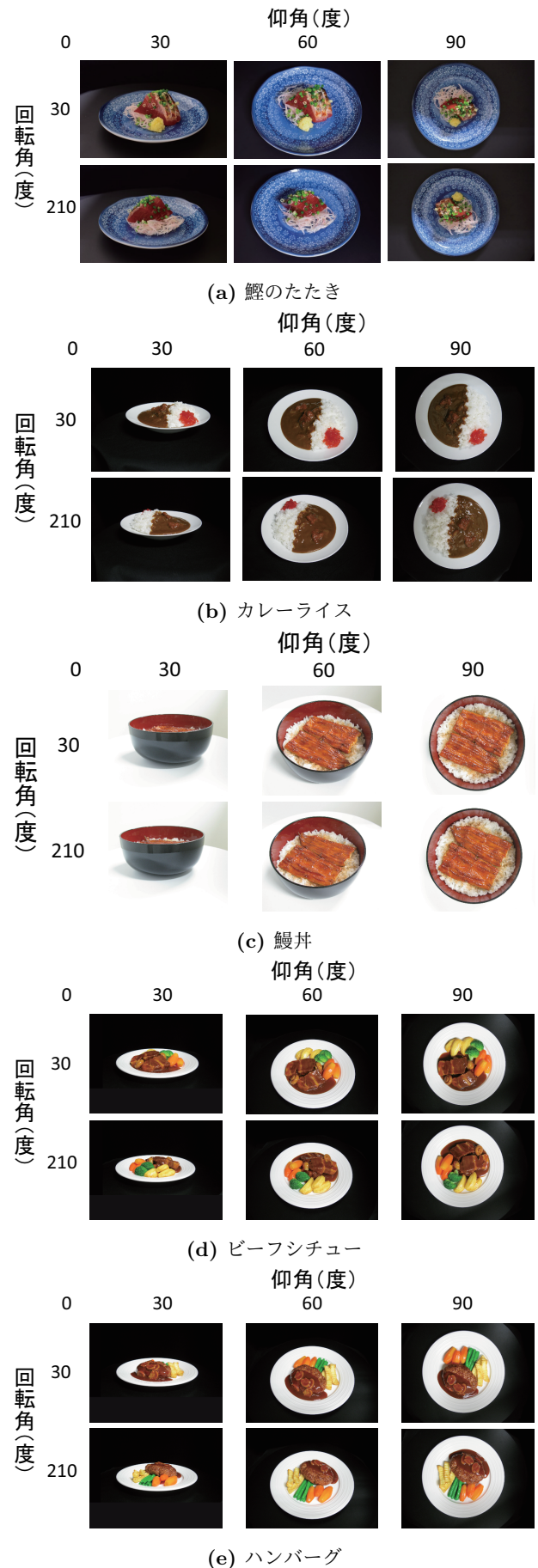


図 6: データセット [1] 内の料理画像群の抜粋。

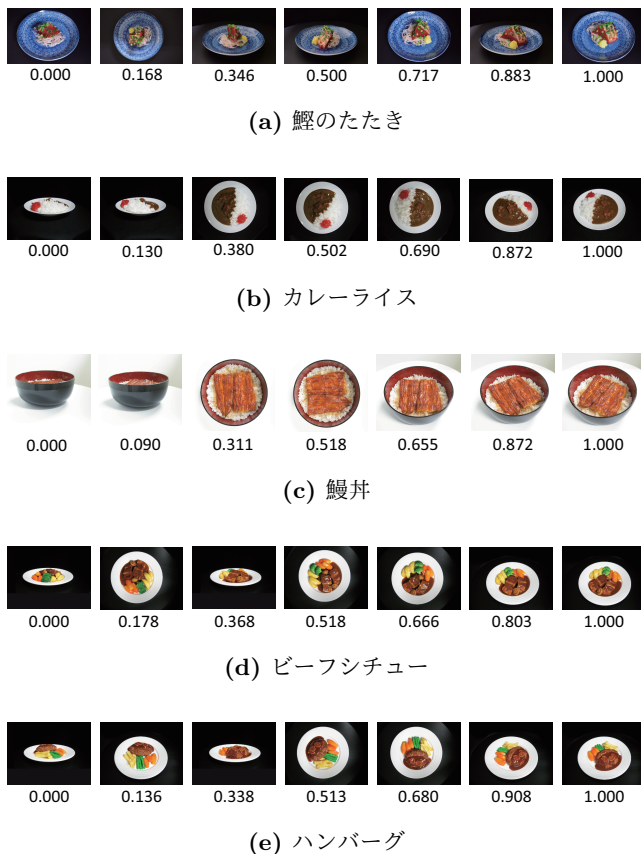


図 7: 選好実験により付与された魅力度付き料理画像群の抜粋 [1].

の被験者から料理ごとに 2,015 件の選好結果を得た。そして、得られた選好結果に対して、間隔尺度値を求めた。最後に、間隔尺度値の最小値が 0, 最大値が 1 となるように正規化し、料理画像の魅力度とした。魅力度が 1 に近いほど魅力度が相対的に高い画像である。この処理を各料理に適用することで、魅力度付き料理画像データセットを構築した。

3.1.4 一対比較により付与された魅力度

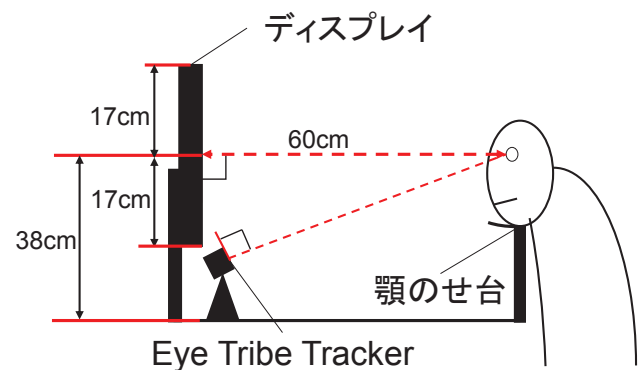
図 7 に、付与された魅力度により順位付けされた各料理画像群の抜粋を示す。鰹のたたきは、焼き目や生姜が手前に見える構図の魅力度が高い。カレーライスは、福神漬けが奥に見える構図の魅力度が高い。鰻丼は、仰角 60 度から撮影され、鰻の向きが手前に向かって垂直ではない構図の魅力度が高い。ビーフシチューとハンバーグは、肉が手前に見える構図の魅力度が高く、共通した傾向が見られる。魅力度が低い画像では、5 種類の料理に共通して、料理全体や主となる食材が見えている部分が少ない構図や、立体感が少ない構図が多い。

3.2 料理画像選好時における視線計測

本節では、3.1 節で紹介したデータセットを用いて被験者に選好課題を課した時の視線計測方法について述べる。



(a) 実験風景



(b) 見取り図

図 8: 視線計測実験の環境。

3.2.1 実験環境

視線計測を行うために、ディスプレイ、視線計測装置、顎のせ台を用意した。図 8 に、視線計測の実験環境を示す。視線計測装置には Eye Tribe Tracker^{*1} を使い、そのサンプリング周波数を 30 Hz に設定した。ディスプレイの大きさは 27 インチ、解像度は 1,920×1,080 画素であった。図 8(b) に示すように、ディスプレイと顎のせ台の距離を約 60 cm に固定し、被験者の目の高さをディスプレイ中央に合わせて実験を行った。この場合、被験者の視覚の中心視野 1 度はディスプレイ上で直径約 34 画素の円に相当する。

被験者に提示する料理画像は、3.1 で紹介したデータセットから、見た目が異なるハンバーグとカレーライスの料理画像群を使用した。被験者の負担を軽減するために、各料理画像群の画像 36 枚のうち、魅力度推定が比較的困難と思われる 0.5 以上の魅力度が付与された画像 10 枚を選び、これらの間の異なる 2 枚の順列 ${}_{10}P_2 = 90$ 通りの画像対を生成した。

^{*1} The Eye Tribe Aps, "Eye Tribe Tracker," <http://theeyetribe.com/>

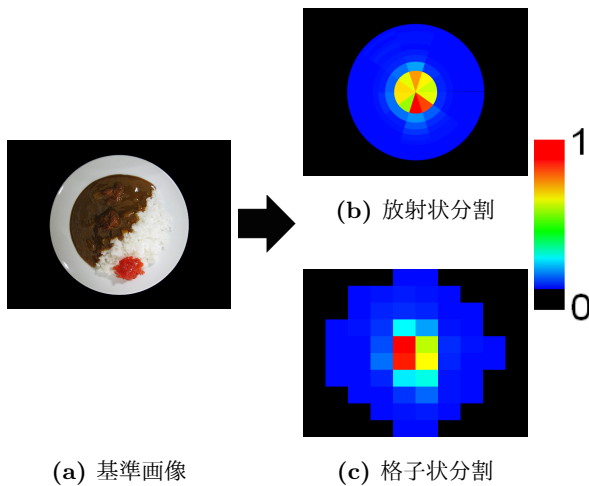


図 9: 各領域分割方法における視線の累積停留時間のヒートマップ (カレーライス).

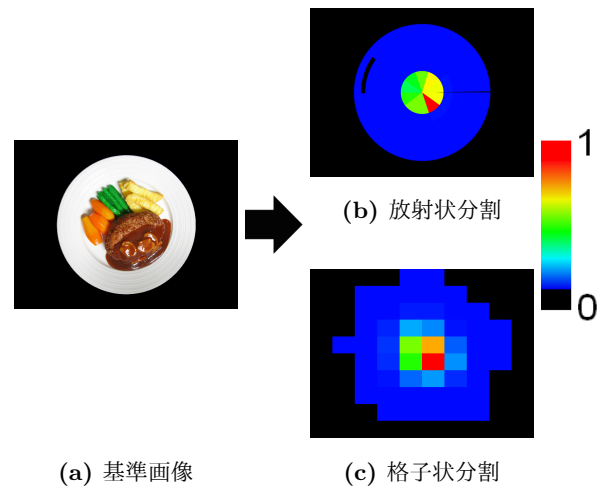


図 10: 各領域分割方法における視線の累積停留時間のヒートマップ (ハンバーグ).

3.2.2 実験方法

まず、視線を固定させるための画像を提示した。被験者には、その画像の中央を注視した時にキーボードのキーを押下するように指示した。その直後に、3.2.1 で生成した一対の料理画像を左右に提示した。3.1.3 で紹介した選好実験と同様に、被験者は、「美味しそうに見える方はどちらか」という質問に対して、「左」または「右」という回答を入力するキーを押下し、判断できない場合には「分からない」という回答を入力するキーを押下した。一対の料理画像が表示されてから質問に対する回答を入力するまでの間、視線を計測した。その後、再び視線を固定させるための画像を提示した。1 人の被験者につき、90 通り全ての料理画像対に対してこの課題を課した。なお、被験者は 20 代の男女 9 名であった。

閲覧者一般の料理上の注目領域を分析するために、計測した注視点座標を各料理の基準画像に合わせるように座標変換することで統合した。そして、統合した注視点群を個々の元画像へ再び座標変換することで各撮影角度の料理画像に対する疑似的な視線情報を作成した。

3.2.3 視線計測結果

図 9 と図 10 に、統合した視線情報に対して、2.1.3 で述べた方法で画像を $10 \times 10 = 100$ 個に分割した局所領域ごとに累積停留時間を求め、ヒートマップとして表したものを示す。

図 9 と図 10 で料理間についてヒートマップを比較すると、どちらの料理でも、中心付近の局所領域で累積時間が長くなったことが確認できる。つまり、料理や角度によらず、被験者は周辺に比べて中心を注視する時間が長い。

4. 評価実験

本節では、視線計測実験で得られた視線情報に基づいて画像特徴抽出領域を設定して料理写真の魅力度を推定する

提案手法の評価実験について報告する。

4.1 実験条件

3.1 節で構築したハンバーグとカレーライスのデータセットそれぞれについて、視線計測実験で得られた視線情報に基づいて画像特徴抽出領域を設定して料理写真の魅力度を推定することの効果の評価した。なお、視線情報を考慮せずに料理領域全体から画像特徴を抽出する従来手法 [1] を比較手法とした。

画像特徴抽出領域の選択条件として、各局所領域中の視線の累積停留時間の範囲は、3.2.3 項の結果から注視領域とそれ以外の面積差が最小となるように、0.1 をしきい値としてそれ以上か未満かの 2 種類を設定した。ここで、0.1 以上を条件とする領域選択については注視領域内から画像特徴を抽出、0.1 未満を条件とする領域選択については注視領域外から画像特徴を抽出することになる。

推定器の構築および評価は、データセット内の各料理の画像群で学習した推定器を用いる Leave-One-Out 法を用いた。ただし、テスト画像の撮影角度を既知として、各料理の基準画像における画像特徴抽出領域をテスト画像へと座標変換した。Random Regression Forest の学習には、scikit-learn ライブラリ [11] 中の RandomForestRegressor を利用し、パラメータは `random_state = 2`, `n_estimators = 150` とした。事前処理として特徴量の各次元を $[0, 1]$ に正規化した。評価指標は、Thurstone の一対比較法により算出したデータセット内の画像の魅力度と、各手法により推定された魅力度の平均絶対誤差 (MAE: Mean Absolute Error) とした。

なお、 $L^*a^*b^*$ 色空間の色差の明度 L^* 、色度 a^* 、色度 b^* からなる色特徴は各々 8 段階に量子化し、形状特徴である最大エッジ強度の勾配方向は 36 段階に離散化した。

表 1: 2つの画像特徴を統合利用した場合の平均絶対誤差. 下線は従来手法の誤差以下の値, 太字は最小値を示す.

		色特徴		
		全体 (従来)	注視領域内 (0.1以上)	注視領域外 (0.1未満)
形状特徴	全体 (従来)	<u>0.121</u>	0.122	<u>0.114</u>
	注視領域内 (0.1以上)	0.165	0.131	0.131
	注視領域外 (0.1未満)	<u>0.120</u>	0.124	0.113

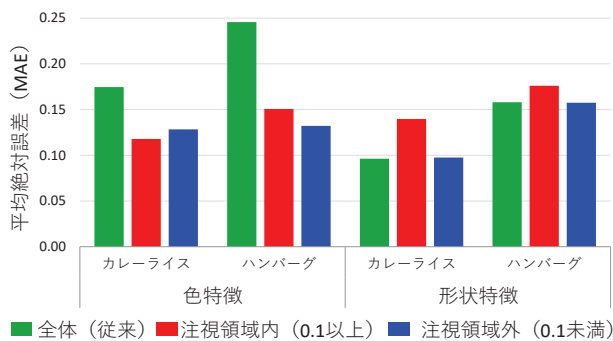


図 11: 画像特徴単体で利用した場合の平均絶対誤差.

4.2 実験結果

表 1 に 2 つの画像特徴を統合利用した場合の結果を示す. この表から, 両画像特徴ともに注視領域外から抽出した場合で推定誤差が最小となったことが確認できる. このことから, 両画像特徴ともに注視領域外から抽出することが有効であると言える. しかし, どちらかの画像特徴で注視領域内から抽出した場合は, 従来手法 [1] よりも推定誤差が増大した. このことから, 両画像特徴ともに注視領域内から抽出することは有効ではないと言える. また, 図 11 に画像特徴を単体で使用した場合の料理ごとの結果を示す. この図より, どちらの料理でも色特徴を抽出する領域を選択することが有効であると考えられる.

4.3 考察

画像特徴抽出領域の選択について, 注視領域外からの抽出により推定誤差が減少したのに対して, 注視領域内からの抽出では推定誤差が増大した. 尾崎らの研究 [5] において, 画像を選好する際に人間は注視領域内の情報だけでなく注視領域外の情報を積極的に利用していると指摘されているように, 注視領域外の画像特徴が料理写真の魅力度推定に大きく貢献したことは妥当な結果であると考えられる. 本研究で使用した料理画像群は, 撮影角度のみの変化に着目したものであるため, 回転中心である料理画像中心に近いほど, 撮影角度間で画像特徴の差異が小さい. 図 9 と図 10 に示すように, 視線の累積停留時間が長い場合, 画

像中心付近が注視領域として検出されやすい. そのため, 注視領域内の画像特徴は撮影角度間で差異が小さい傾向がある. したがって, 注視領域を狭めて設定し, その内部から抽出した画像特徴だけで料理写真の魅力度を推定するのは困難であることから, 推定誤差が増大したと考えられる.

5. むすび

本報告では, 料理を魅力的に撮影するための支援を目的として, 料理写真の魅力度を推定する手法を提案した. 提案手法は, 視線情報に基づき設定した局所領域から特徴量を抽出し, 回帰パラメータを学習することにより料理写真の魅力度を推定するものである. 具体的には, 画像を分割した局所領域における視線の累積停留時間に基づいて注視領域を検出し, それに応じて学習に利用する画像特徴抽出領域を設定した. そして, 設定した領域から色特徴と形状特徴を抽出し, Random Regression Forest [7] を用いて学習した.

被験者に一対比較による料理画像の選好課題を課して, 視線を計測した. 料理画像は高橋らの魅力度付き料理画像データセット [1] のものを利用した. 具体的には, 2 種類の料理カテゴリ (カレーライス, ハンバーグ) に対して, 各料理を 36 方向から撮影した画像群のうち 10 方向分を使用した.

実験により, 提案手法の効果を評価した. 比較手法として, 注視領域を考慮せずに料理領域全体から画像特徴を抽出する従来手法を用いた. その結果, 注視領域内ではなく, 注視領域外から画像特徴を抽出することの有効性が示唆された. これは, 尾崎らの従来知見 [5] とも一致した. また, 形状特徴よりも, 色特徴のほうが画像特徴抽出領域を設定した効果が表れた.

今後の課題として, 料理の種類を追加, 画像特徴抽出領域の分割方法の検討, 他の画像特徴の導入, 視線情報の利用方法の改良, 撮影角度が未知である状況への対応が挙げられる.

本報告で評価した料理は 2 種類だけであり, 提案手法の有効性を結論づけるには十分な結果を得ていない. そのため, より多くの種類の料理画像に対して提案手法を評価する必要がある.

本研究では, 抽出する画像特徴として色特徴と形状特徴を採用した. 色特徴は放射状に, 形状特徴は格子状に領域分割して抽出したが, 色特徴を格子状領域分割, 形状特徴を放射状領域分割によって抽出する手法が考えられる. また, 局所領域の分割個数を変化させて検証する必要がある. さらに, 更なる精度向上のためには, 色や形状に関する特徴以外との統合利用を検討する必要がある.

提案手法では, 視線の累積停留時間が長い局所領域を注視領域として, それに応じて画像特徴抽出領域を設定した. 本報告では, 画像特徴抽出領域の選択条件として, 視

線の累積停留時間の範囲を1つのしきい値のみでしか設定していないので、他の選択条件で評価する必要がある。また、菅野らの従来研究 [4] によれば、視線の停留以外にも視線の遷移や瞬きといった情報が有効である可能性があるため、その導入を検討することができる。さらに、提案手法では画像特徴を抽出する領域としない領域を分けたが、より柔軟に局所領域に重みづけを行う手法も検討することができる。加えて、提案手法では撮影角度が既知である料理画像に対して推定器を構築し、その評価を行った。しかし、応用を考えると、撮影角度が未知である状況に対応する必要がある。例えば、食材の配置や大きさが異なるときに共通する視線停留の傾向を分析し、その結果を魅力度推定に利用する手法が考えられる。

謝辞 本研究の一部は、科研費及び MSR Core-12 プログラムによる。

参考文献

- [1] 高橋和馬, 道満恵介, 川西康友, 平山高嗣, 井手一郎, 出口大輔, 村瀬 洋: 料理写真の魅力度推定に用いる特徴量の検討による精度向上, 電子情報通信学会技術研究報告, MVE2016-12 (2016).
- [2] Donahue, J., Jia, Y., Vinval, O., Hoffman, J., Zhang, N., Tzeng, E. and Darrell, T.: DeCAF: A Deep Convolutional Activation Feature for generic visual recognition, *Proc. 31st Intl. Conf. on Machine Learning*, pp. 647–655 (2014).
- [3] Shimojo, S., Simion, C., Shimojo, E. and Scheier, C.: Gaze bias both reflects and influences preference, *Nat. Neurosci.*, Vol. 6, No. 12, pp. 1317–1322 (2003).
- [4] Sugano, Y., Ozaki, Y., Kasai, H. and Sato, Y.: Image preference estimation with a data-driven approach: A comparative study between gaze and image features, *J. of Eye Movement Research*, Vol. 7, No. 3, pp. 5:1–5:9 (2014).
- [5] 尾崎安範, 菅野裕介, 佐藤洋一: 視線情報と画像特徴に基づく画像の嗜好推定, 電子情報通信学会技術研究報告, PRMU2013-168 (2014).
- [6] Rother, C., Kolmogorov, V. and Blake, A.: GrabCut: Interactive foreground extraction using iterated graph cuts, *ACM Trans. on Graphics*, Vol. 23, No. 3, pp. 309–314 (2004).
- [7] Andy, L. and Wiener, M.: Classification and regression by random forest, *R News*, Vol. 2, No. 3, pp. 18–20 (2002).
- [8] Nishiyama, M., Okabe, T., Sato, I. and Sato, Y.: Aesthetic quality classification of photographs based on color harmony, *Proc. 2011 IEEE Conf. on Computer Vision and Pattern Recognition*, pp. 33–40 (2011).
- [9] Michel, C., Woods, A. T., Neuhauser, M., Landgraf, A. and Spence, C.: Rotating plates: Online study demonstrates the importance of orientation in the plating of food, *Food Quality and Preference*, Vol. 44, pp. 194–202 (2015).
- [10] Thurstone, L. L.: Psychophysical analysis, *American J. of Psychology*, Vol. 38, No. 3, pp. 368–389 (1927).
- [11] Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M. and Duches-

nay, E.: Aesthetic quality classification of photographs based on color harmonics, *Machine learning in Python, J. of Machine Learning Research 12*, pp. 2825–2830 (2011).