社団法人　電子情報通信学会
THE INSTITUTE OF ELECTRONICS,
INFORMATION AND COMMUNICATION ENGINEERS

信学技報
TECHNICAL REPORT OF IEICE.

料理写真の魅力度を推定する際の画像特徴に関する定量分析

食材構成の理解が魅力度に及ぼす影響

服部　竜実†　道満　恵介†　井手　一郎††　目加田慶人†

† 中京大学 工学部　〒 470–0393 愛知県豊田市貝津町床立 101
†† 名古屋大学 大学院情報学研究科　〒 464–8601 愛知県名古屋市千種区不老町
E-mail: †hattori.t@md.sist.chukyo-u.ac.jp, {kdoman,y-mekada}@sist.chukyo-u.ac.jp
††ide@i.nagoya-u.ac.jp

あらまし　本発表では，食材構成の理解の有無が魅力度および魅力度推定精度に及ぼす影響を定量的に分析した結果について述べる．これまで我々は，魅力度に影響を及ぼす要因として撮影角度，ぼかし，スケールに注目し，種々の画像特徴を統合利用して料理写真の魅力度を推定する手法を提案した．具体的には，料理全体の印象や主食材の見えに基づく画像特徴量と魅力度をそれぞれ説明変数，目的変数とする回帰分析により料理写真の魅力度を推定する手法を提案し，有効性を確認した．しかし，高精度な魅力度推定の実現のためには，色，形状，テクスチャという低レベルの画像特徴の組み合わせだけでなく，食材構成という高レベルな意味的特徴を考慮する必要があると考えられる．そこで，食材構成の理解の有無が魅力度に及ぼす影響を被験者実験を通して定量的に分析した．また，それらが魅力度推定手法の精度においてどの程度影響を及ぼすかについて定量的に調査し，必要な画像特徴について検討した．
キーワード　料理写真，魅力度，画像特徴，回帰分析

Quantitative analysis on image features for the estimation of the attractiveness of food photography

How the composition of ingredients affects the attractiveness

Tatsumi HATTORI†, Keisuke DOMAN†, Ichiro IDE††, and Yoshito MEKADA†

† School of Engineering, Chukyo University
Tokodachi 101, Kaizu-cho, Toyota, Aichi, 470–0393 Japan
†† Graduate School of Informatics, Nagoya University
Furo–cho, Chikusa-ku, Nagoya, Aichi, 464–8601 Japan
E-mail: †hattori.t@md.sist.chukyo-u.ac.jp, {kdoman,y-mekada}@sist.chukyo-u.ac.jp
††ide@i.nagoya-u.ac.jp

Abstract　We report the results of quantitative analysis on how the composition of food ingredients a ects the attractiveness of food photography. We have been focusing on some photography parameters and have proposed a method for attractiveness estimation for food photography by integrating multiple image features. It estimated the attractiveness based on regression analysis using image features as explanatory variables and attractiveness as an objective variable. Although we confirmed its e ectiveness for attractiveness estimation, we considered that it is necessary to introduce image features that evaluate the composition of ingredients for more accurate estimation, because they should significantly a ect the attractiveness of food photography. Thus, we analyzed how the attractiveness is a ected by understanding the composition of food ingredients through subjective experiments. We also investigated how it a ects the performance of our method in the estimation accuracy, and studied the image features for attractiveness estimation.
Key words　Food photography, attractiveness, image features, regression analysis

(a) 魅力的でない料理写真　　(b) 魅力的な料理写真

図 1: 魅力度が異なる料理写真の例

Fig. 1　Example of food photos with different attractiveness.

1. はじめに

料理レシピサイトや SNS への料理写真の投稿機会が増加している．このような 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．例えば，図 1 は同一の料理を撮影した写真であるが，図 1a よりも図 1b の方が，被写体の大きさ，ぼかし，撮影角度等の観点で料理が美味しそうに撮影されている．料理写真を美味しそうに撮影することは必ずしも容易ではなく，魅力的な料理写真の撮影を支援するシステムがあれば有用である．本研究では，料理写真が美味しそうに見える度合いを「魅力度」と定義し，これを定量的に分析・推定する技術の開発を目的とする．

これまで我々は，魅力度に影響を及ぼす要因の一つとして撮影角度に注目し，種々の画像特徴を統合利用して料理写真の魅力度を推定する手法を提案した [1]．具体的には，料理全体の印象や主食材の見えに基づく画像特徴量と魅力度をそれぞれ説明変数，目的変数とする回帰分析により，料理写真の魅力度を推定する手法を提案した．この手法の評価のために，10 種類の料理を様々な角度から撮影して得られた画像データセット(注1)を作成し，被験者実験により各画像に魅力度を付与した．このデータセットを用いて手法の有効性を確認したが，撮影角度以外の要因が魅力度に及ぼす影響については分析できていなかった．さらに我々は，撮影角度にぼかしとスケールを加えた計 3 種類の要因が魅力度に及ぼす影響を定量的に分析した [2]．その結果，料理カテゴリによって最適なぼかしや拡大の方法が異なること，および，それらの要因の考慮が高精度な魅力度推定に重要であることが示唆された．

これまでの手法では，料理全体の印象や主食材の見えに関する色，形状，テクスチャ基づく画像特徴を利用してきたが，料理というコンテキストにおいては，単に画像特徴のみではなく食材構成も魅力度に影響を及ぼすと考えられる．これに関して，料理写真の例を図 2 に示す．図 2b では副菜や添え物を含めた料理の食材構成が写真から読み取れるが，図 2a ではそうではない．そのため，図 2a では色彩調和の観点から左の写真の方が

(注1)：NU FOOD 360x10: http://www.murase.is.i.nagoya-u.ac.jp/nufood/

(a) 食材構成が判らない場合：左の写真の方が魅力的

(b) 食材構成が判る場合：右の写真の方が魅力的

図 2: 食材構成の理解の有無による魅力度の違い

Fig. 2　Difference in the attractiveness of food photos depending on understanding the composition of ingredients.

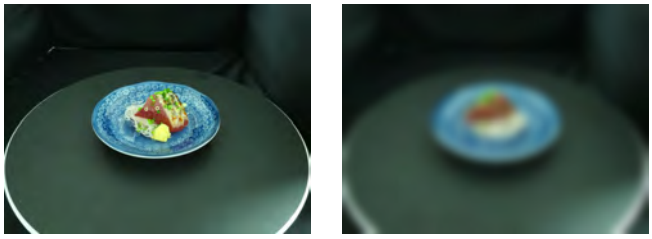
魅力度が高いが，図 2b では食材構成が判別可能なことから料理の主食材が強調されている右の写真の方が魅力度が高いと考えられる．つまり，料理というコンテキストにおいては，料理の食材構成の理解の有無が料理写真の魅力度に影響を及ぼす．よって，高精度な魅力度推定の実現のためには，色，形状，テクスチャという低レベルの画像特徴の組み合わせだけでなく，食材構成という高レベルな意味的特徴を考慮する必要があると考えられる．

そこで発表では，食材構成の理解の有無が魅力度および魅力度推定精度に及ぼす影響を定量的に分析し，食材構成を評価するための画像特徴に関して検討する．以降，2. では料理写真の魅力度推定手法について概説する．続く 3. および 4. では，食材構成の理解の有無が魅力度および魅力度推定精度に及ぼす影響を調査した実験の結果についてそれぞれ述べる．最後に 5. でまとめる．

2. 料理写真の魅力度推定手法

我々がこれまで提案してきた魅力度推定手法 [1], [2] の大まかな処理手順は，図 3 に示すように学習段階と推定段階の 2 つから構成される．学習段階では，予め魅力度が付与された料理画像を用いて，回帰の枠組みにより推定器を構築する．回帰モデルには Random Regression Forests [3] を利用し，料理画像の魅力度とその画像特徴量をそれぞれ目的変数，説明変数とする．推定段階では，学習段階で構築された推定器を用いて，入力された料理画像に対する魅力度を推定する．

各段階における画像特徴抽出では，図 4 に示すように，まず入力画像を料理領域 R_d と主食材領域 R_m に分割する．これには，GrubCut [4] の利用およびスマートフォン等のデバイスを用いたユーザインタラクションを想定する．その後，料理領域 R_d から料理全体の印象に関する画像特徴を抽出し，主食材領



(a) 食材構成が判る場合　(b) 食材構成が判らない場合

図 8: 鰹のたたきにおいて魅力度最大となった画像

Fig. 8 Food image with the highest attractiveness value in the Sashimi category.

際に D_2 と D_1 間の推定精度の差は僅かであった．回帰分析において画像特徴と目標値の関係が複雑になるほど未知の入力に対する予測は難しくなり，そのことが上記の結果の一因であると考える．なお，鰻丼については，3. で述べたように食材構成の理解の有無により撮影角度と魅力度の関係に若干の差異はあるものの，推定タスクが劇的に難化するような差異ではなかったものと考える．

次に，食材構成が判らない D_2 で学習した推定器を用いて食材構成が判る D_1 でテストした場合に，比較的大きな推定誤差となったことについて考察する．前述の通り，D_2 においては，料理写真より情景等を対象とした一般的な写真に近い観点から魅力度が付与されていると考えられる．一方，D_1 においては，単純な色の分布や形状だけでは測れない観点から魅力度が付与されていると考えられる．よって，D_2 で学習した画像特徴と魅力度の関係を以て D_1 のそれを推定することは困難であり，そのことが上記の結果の一因であると考える．

4.3.2 高精度な魅力度推定に向けた画像特徴の導入

前述の結果および考察から，高精度な魅力度推定の実現には，まず料理カテゴリおよび食材を認識した上で各食材の見えと配置を評価する画像特徴が必要であるといえる．今後は，本実験で得られた知見を踏まえて最適な画像特徴を検討するが，その一方で，畳み込みニューラルネット（CNN）を用いた特徴抽出も検討の余地がある．CNN は料理認識をはじめ様々な認識のタスクにおいて高い精度を達成しており [8]，大規模な学習データがあれば有効な画像特徴を自動で獲得できる．ただし，本研究で扱う魅力度は人の感覚量であり，学習の際の目標値を設定する際には被験者実験が必要不可欠となる．そこで，大規模な料理画像データセットを用いて識別問題として CNN を構築し，その後，本実験で構築したデータセットを用いてファインチューニングすることも併せて検討する．

5. まとめ

本報告では，食材構成の有無の理解が料理写真の魅力度に及ぼす影響を定量的に分析した結果について述べた．被験者実験の結果から，食材構成の理解の有無が魅力度に影響を及ぼすこと，および，その影響度合いは料理の種類によって異なることが確認された．さらに，評価実験を通して，食材構成の理解の有無が魅力度推定精度に及ぼす影響も料理の種類によって異なることが確認され，高精度な魅力度推定の実現には食材構成を考慮した画像特徴が必要であることも確かめられた．今後は，本稿で述べた実験により得られた知見を踏まえて，料理における食材構成を評価するために最適な画像特徴および魅力度推定の枠組みを検討していく．

表 1: 実験結果：各データセットの組み合わせに対する魅力度の推定精度（MAE：平均絶対誤差）

Table 1 Experimental results: Estimation accuracy (MAE: Mean Absolute Error) for each dataset combination.

学習用セット	推定用セット	MAE			
		鰹のたたき	ビーフシチュー	鰻丼	平均
D_1	D_1	0.101	0.089	0.089	0.093
D_2	D_2	0.084	0.092	0.085	0.087
D_2	D_1	0.327	0.349	0.641	0.439

謝辞　本研究の一部は JSPS 科研費 JP17K12719，Microsoft CORE-12 による．

文　献

[1] K. Takahashi, K. Doman, Y. Kawanishi, T. Hirayama, I. Ide, D. Deguchi, and H. Murase, “Estimation of the attractiveness of food photography focusing on main ingredients,” Proc. 9th Workshop on Multimedia for Cooking and Eating Activities, pp.1–6, Aug. 2017.

[2] 道満恵介，服部竜実，井手一郎，目加田慶人，“複数の撮影パラメータを考慮した料理写真の魅力度推定に向けた検討，” 第 20 回画像の認識・理解シンポジウム，no.PS1-53，pp.1–4，Aug. 2017.

[3] A. Liaw and M. Wiener, “Classification and regression by randomForest,” R News, vol.2, no.3, pp.18–22, Dec. 2002.

[4] C. Rother, V. Kolmogorov, and A. Blake, “GrabCut —Interactive foreground extraction using iterated graph cuts,” ACM Trans. on Graphics —Proc. ACM SIGGRAPH 2004, vol.23, no.3, pp.309–314, Aug. 2004.

[5] J. Donahue, Y. Jia, O. Vinyals, J. Hoffman, N. Zhang, E. Tzeng, and T. Darrell, “DeCAF: A deep convolutional activation feature for generic visual recognition,” Proc. 31st International Conference on Machine Learning, pp.647–655, June 2014.

[6] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, “ImageNet: A large-scale hierarchical image database,” Proc. 2009 IEEE Conference on Computer Vision and Pattern Recognition, pp.248–255, June 2009.

[7] L.L. Thurstone, “Psychophysical analysis,” American Journal of Psychology, vol.38, no.3, pp.368–389, July 1927.

[8] H. Hassannejad, G. Matrella, P. Ciampolini, I.D. Munari, M. Mordonini, and S. Cagnoni, “Food image recognition using very deep convolutional networks,” Proc. 2nd International Workshop on Multimedia Assisted Dietary Management, pp.41–49, Oct. 2016.