

機械学習を用いた動画像からのハーシャル・バルクレイ流体のパラメータ推定の検討

鈴木貴如^{1,a)} 永澤謙太郎^{1,b)} 岡田真人^{1,c)} 楽詠瀬^{2,d)}

概要: 我々の身近にある多くの興味深い流体の流動特性はハーシャル・バルクレイモデルで表現できる。コンピュータアニメーションでこれらの実用的な流体を再現するには、そのパラメータが必要になるので、これらの流体を対象として、その動画映像からモデルのパラメータを推定する手法の開発について検討する。通常、流動特性の計測には、レオメータと呼ばれる粘弾性測定装置が用いられる。しかし、レオメータは高価で一般ユーザーには入手が難しく、またレオメータで測定できる流体は、ゴマなどの混入物を含まないものに限られるなど、一定の制約がある。一方で、液体の粘性を対象としたヒトの視覚認識についての研究では、我々は液体が流動する際の動き情報のみから、粘性を認識できるという知見がある。そこで、レオメータのような特別な計測装置を用いない、よりアクセスのしやすい流動性推定手法として、流体の動画映像を入力とし、ヒトの視覚認識をモデル化した畳み込みニューラルネットワークを用いた手法について検討する。

Parameter Estimation of Herschel-Bulkley Fluids from Videos using Machine Learning

Abstract: The flowing properties of many interesting fluids around us can be well represented by the Herschel-Bulkley model. To handle such practical fluids in computer animation, we need their model parameters for simulation. In this work, we develop a method for estimating such parameters from the videos of the fluids. A standard way to obtain the flowing property of a material is to use a laboratory device called rheometer, which, however, comes with drawbacks: it is not easily available for a standard user because it is expensive, and fluids that can be handled by a rheometer is usually restricted to those without any inclusions (like sesame). On the other hand, recent advancements in human visual perception has revealed that humans can recognize fluid viscosity from (only) the motion information of the fluid. Based on these observations, we develop an easy-to-access estimation method for the flowing properties without the need of a rheometer, namely, a method that uses a convolutional neural network to model human visual perception with the video of the fluid being the only input.

1. はじめに

近年コンピュータグラフィックス分野では、非ニュートン流体の映像表現のためのシミュレーション手法 [1] が盛んに研究されている。非ニュートン流体は、流体状食品をはじめとして我々の身の回りに普遍的に存在しており、興味深く特徴的な流動性を持つ。例えば、クリームなどは流体と固体の双方の性質を持ち、一定の形を保って静止する

ことができ、また、粘性がかけられた力に応じて非線形に変化して流動性が変化する (より流れやすくなったり、逆に流れにくくなったりする) 特性を持つ。これらの流体の動きを写實的に再現したい場合、その流動特性を決定づける物性パラメータが必要となる。

通常、流体の流動特性を定量的に知るためには、レオメータと呼ばれる粘弾性測定装置を用いた計測を行う。しかし、一般的にレオメータは高額装置であるため、一般ユーザーには入手しにくく、また、計測できる流体が空間的に一様なものに限られ、ゴマなどの混入物を含む試料を扱えないなどの問題点がある [2]。一方で、液体の粘性を対象としたヒトの視覚認識についての研究 [3] から、我々

¹ 東京大学

² 青山学院大学

a) suzuki@mns.k.u-tokyo.ac.jp

b) nagasawa@mns.k.u-tokyo.ac.jp

c) okada@edu.k.u-tokyo.ac.jp

d) yonghao@it.aoyama.ac.jp

は液体が流動する際の動きの情報のみから、粘性を認識できることが分かっている。そこで、本研究ではレオメータを用いずに、非ニュートン流体(特に流体状食品)の流れる様子を撮影した動画のみから、その粘性に関する物性パラメータを推定する手法の開発を目指す。これにより、特別な装置の利用を必要としない、アクセスのしやすい流動性推定手法を実現したい。

従来研究 [2] から、様々な流体状食品の流動性は、ハーシャル・バルクレイモデルでモデル化できることがわかっている。従って流動性推定では、ハーシャル・バルクレイモデルの3つのパラメータを出力とする。また、動き情報の抽出には、視覚認識についての研究 [3] で用いられているオプティカルフロー [4] (連続フレーム間での各画素の移動ベクトルを求める画像処理の手法) を利用する。さらに、ヒトの視覚認識をモデル化した畳み込みニューラルネットワーク (CNN) を利用することによって、動き情報からハーシャル・バルクレイモデルの3つのパラメータへと回帰を行う。

本研究の最終目標は、流動する流体を撮影した動画一本を入力とし、その流体のモデルパラメータ組を推定するシステムの開発であるが、本発表はその予備段階の実験結果を報告するものである。具体的には、後述するように、計算されたオプティカルフローのフレームを切り分けたパッチを入力として、各々のパッチに対してモデルパラメータの推定を試みるというものである。このプロトタイプを構築することで、CNN によるオプティカルフローからの特徴抽出の性能を確認するとともに、今後の開発で予想される問題点等を整理する狙いがある。

2. 関連研究

2.1 粘弾塑性体の再現

近年、コンピュータグラフィックス分野では、物質点法 [1] を用いた粘弾塑性体の力学シミュレーション手法が盛んに研究されている。物質点法は格子と粒子の双方を用いるハイブリッド手法であり、物体の流動性や変形する挙動を特徴付ける構成則を容易に組み込むことができ、雪 [5] やクリーム [6]、流体状食品 [2]、粉体 [7], [8]、布 [9] 等の多種多様な物体の再現が実現されてきた。

2.2 流動特性のモデル化

物体の力学特性を記述する構成則のうち、さらさらやどろどろといった流動特性は、流体要素にかかる剪断速度 $\dot{\gamma}$ と剪断応力 σ の関係、すなわち $\sigma(\dot{\gamma})$ で表される流動曲線によりモデル化される。よく知られた流動曲線のモデルとしてニュートン粘性モデル $\sigma = \eta\dot{\gamma}$ 、指数モデル $\sigma = \eta\dot{\gamma}^h$ 、ビンガムモデル $\sigma = \sigma_Y + \eta\dot{\gamma}$ 、カソンモデル $\sqrt{\sigma} = \sigma_Y + \eta\dot{\gamma}$ 、ハーシャル・バルクレイモデル $\sigma = \sigma_Y + \eta\dot{\gamma}^h$ がある [10]。

ニュートン粘性モデルでは、剪断応力と剪断速度の比率である実効粘性が一定であり、粘性係数 η で表される。この性質を持つ流体はニュートン流体と呼ばれ、代表的な例は水や空気である。

ニュートン粘性モデルを拡張する形で、実効粘性が一定とならない非ニュートン流体を記述できる。降伏応力 σ_Y を導入することで、ビンガムモデルでは、小さい応力では流動が生じずに弾性的に振る舞い、降伏応力以上で初めて流動する降伏現象を記述できる。また、剪断速度に指数を導入することで、指数モデルは曲線的な応力変化を記述できる。降伏応力と曲線性の双方を表現するモデルとして、カソンモデルとハーシャル・バルクレイモデルが位置付けられる。ハーシャル・バルクレイモデルはより自由度があり、ニュートンや指数、ビンガムモデルを包括した形式となっており、様々な流体状食品に対する測定結果に対して精度よくフィットすることがわかっている [2], [11]。なお、指数モデルとハーシャル・バルクレイモデルに現れる指数 h は、剪断速度の増加に応じて、 $h < 1$ では実効粘性が減少するシアーシンニング流体、 $h > 1$ では実効粘性が増加するシアーシックニング流体を表現する。

2.3 動画映像からのモデルパラメータ推定

表現したい対象を記述するモデルが確定できると、次に必要なのはモデルパラメータの決定である。そのための手法として、力学的振る舞いに限らず、再現対象の実写映像などを入力として、逆問題を解く形でモデルパラメータを推定する研究が行われている。弾性体を対象とした研究 [12] では、物体の動きを撮影し、その情報を物体の支配方程式の解として当てはめて逆問題を解き、弾性パラメータを決定する。さらに文献 [13] では、ニューラルネットワークを利用した機械学習によるアプローチも提案されている。また、木を対象とした研究 [14] では、高速フーリエ変換を用いて、写真や映像から枝や葉の動きを決定する物理パラメータを推定している。さらに、レンダリングの文脈で行われた雲を対象とした研究 [15] では、現実の雲の画像情報を用いて逆レンダリング問題を解くことで、雲のレンダリングパラメータを決定する。

本研究では粘性流体を対象として、実際の映像から、流体を再現するためのパラメータ推定を行う。特に、我々の身近に多く存在する流体状食品を含む非ニュートン流体を対象とし、機械学習を利用して、流体を撮影した映像からパラメータを推定するアプローチを提案する。

2.4 ヒトの視覚認識

ヒトが映像から抽出できる質感情報に関する研究は、視覚認知の分野で行われており、特に定まった形状を持たない液体を対象にした研究としては、文献 [3], [16], [17] など

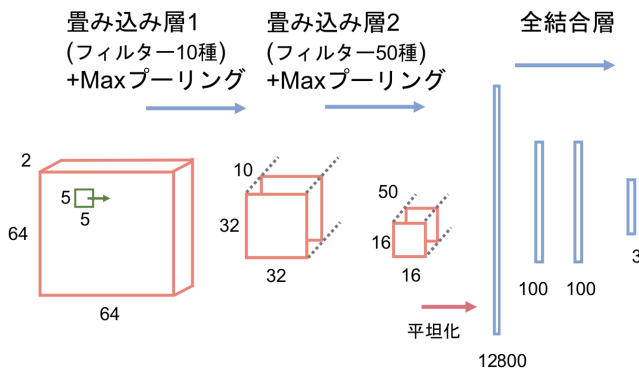


図1 使用したCNNは、畳み込み層1,2共にカーネルサイズを 5×5 、ストライド幅を $[1, 1]$ 、かつ、Max プーリングのサイズを 2×2 、ストライド幅を $[2, 2]$ 、と設計した。

が挙げられる。文献[16]では、単一の静止画に対するヒトの液体粘性認識を調査している一方で、文献[3]では動画を対象とし、ヒトが流体映像から抽出される動き情報だけから、粘性の違いを区別できることを明らかにした。本研究ではこの知見に着目し、ヒトの視覚認識をモデル化したCNN[18],[19]を活用して、動画の動き情報を抽出したオプティカルフロー[4]から、その流体の流動特性をハERSHEY・バルクレイモデルで表した場合のモデルパラメータを推定する、というアプローチをとる。

3. 提案手法

3.1 本研究の目標と本稿の内容

本研究では、最終目標として、流動する流体状食品を撮影した動画1本を入力とし、その流動性をハERSHEY・バルクレイモデルにより表現した際のパラメータ(σ_Y, η, h)を1組推定するシステムを実現したい。その方法として、映像から抽出したオプティカルフローを入力データ、映像内の流体の粘性情報を教師データとして、CNNを利用した学習による回帰を行う。

本稿では、その前段階として、オプティカルフローにおける局所的な情報からの粘性推定の可能性について調べた。具体的には、簡単なネットワーク構造を持つCNNを設計し、オプティカルフローのフレームを切り分けた小さなパッチ群について、粘性情報(σ_Y, η, h)への回帰を試みた。

3.2 CNNの設計

本稿では、CNNのネットワーク構造を図1のように設計した。入力データは、 64×64 ピクセルのオプティカルフローのパッチ(x, y の速度を表す2チャンネルの画像)である。入力された各パッチは、2回の畳み込みおよびプーリング処理の後に平坦化され、その後さらに2回の全結合層を経る。畳み込み層および全結合層における中間層の活性化関数は全てReLU関数とした。最終層では、ハERSHEY・バルクレイモデルのパラメータ3成分が1組出力される。

3.3 学習データ

サンプルをマスタード、とんかつソースおよびマヨネーズの3種類に限定し、それぞれについて一本ずつ動画を撮影した。試料の選定にあたっては、オプティカルフローの抽出を容易にするため、透明度の低いものを対象とし、流体が映る領域に背景情報が透けて映らないように配慮した。また、流動性が大きく離れたものと近いものを含むようにした。マヨネーズはどろどろとして流れにくいのにに対し、マスタードととんかつソースはさらさらと流れることから、これらの流動性の違いは直感的にも分かりやすい。映像撮影においては、図3左列のように、各試料をノズルを用いて鉛直下向きに射出し、30度傾けた金属板の上に流しながら、その様子をスマートフォン(iPhone SE, 1080p, 30fps)により録画した。これは、撮影毎に再現しやすいシンプルな設定下で行うことを考慮している。

データセットとしては、各試料について、ノズルから射出された直後からの約3秒(101フレーム)分の映像を抽出し、隣り合うフレーム同士を用いて図3右列のようなオプティカルフローのフレームを100フレーム分計算した。実写動画およびオプティカルフローのフレームサイズは 400×710 となっている。これらのオプティカルフローのフレームを時系列を無視してランダムに(重複なく)振り分け、訓練用に64枚、検証用に16枚、学習済みモデルのテスト用に20枚とした。

CNNへの入力には、これらの各フレームから 64×64 のサイズで切り出した小さなパッチとした。映像の位置不変性を担保するため、この切り出し操作は、パッチ間で一部重複するようにパッチ幅よりも狭い一定の間隔(縦・横いずれも32)で行なった。こうして得たパッチ群の中には、流体が写っていない背景成分のみのパッチも存在する。これらのパッチは流体に関する情報を何ら持たず、3種類の試料で共通となるため、学習データから取り除く方が学習性能が高いと考えられる。そこで、各パッチ内のオプティカルフローのベクトルの平均ノルムを求め、閾値(0.05)以下のものを背景パッチとして除外した。

一方、CNNの出力は、選定した3種のサンプルのモデルパラメータ(σ_Y, η, h)である。それらの真値は具体的には、マスタード: (12.5, 5.08, 0.461)、とんかつソース: (3.74, 3.75, 0.518)、マヨネーズ: (70.8, 30.5, 0.442)である。すなわち、あるサンプルのオプティカルフローのフレームから切り出された全てのパッチに対して、上記のパラメータ1組が真値として対応する。

3.4 学習ステップ

学習では、各フレームから切り出されたパッチ群を1パッチ分として学習を進める。また、次式のように、粘性パラメータ3成分の真値(σ_Y, η, h)とCNNによる推定値

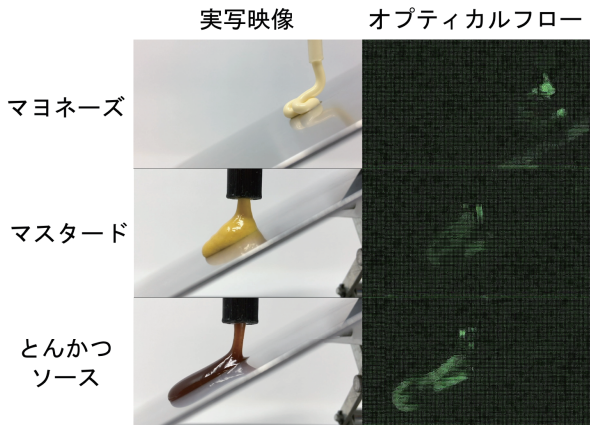


図2 本発表で利用した3種の試料の流動の様子を撮影した実写動画(左列)と、そこから抽出したオプティカルフロー(右列)の1フレームを示す。

$(\hat{\sigma}_{Yi}, \hat{\eta}_i, \hat{h}_i)$ の相対誤差の二乗の総和を誤差関数 E として最適化を行う。

$$E = \sum_i \left[\left(\frac{\sigma_Y - \hat{\sigma}_{Yi}}{\sigma_Y} \right)^2 + \left(\frac{\eta - \hat{\eta}_i}{\eta} \right)^2 + \left(\frac{h - \hat{h}_i}{h} \right)^2 \right] \quad (1)$$

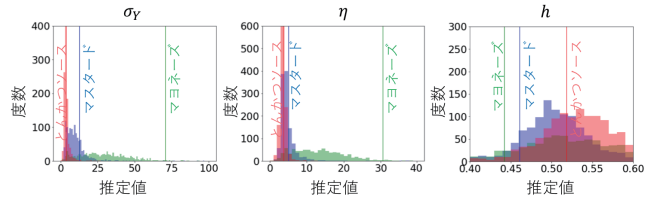
ここで、 i はオプティカルフローの1フレームから切り出されたパッチの番号を表す。なお、最適化には、TensorFlow の AdamOptimizer 関数を利用した。これを訓練用の64フレーム×3種の試料分について1度処理することを1エポックとして、500エポックまで学習を繰り返した。

4. 実験結果

学習と検証には、Intel(R) Xeon(R) CPU E5-2697A v4 2.60GHz, 256 GB メモリを搭載したワークステーションを利用した。学習の効果をテストするため、各試料のテスト用の全フレームから切り出した各パッチを学習済みのCNNに入力し、推定されたパラメータを成分ごとにヒストグラム化し、真値との比較を行った(図4(b))。図中の緑、青、赤の分布はそれぞれ、マヨネーズ、マスタード、とんかつソースの推定値分布であり、縦線はパラメータ各成分の真値を表す。 σ_Y, η については、各成分の推定値の分布が真値を中心に広がっており、性質が大きく異なる試料であるとんかつソース(赤色)とマヨネーズ(緑色)だけでなく、性質が近いとんかつソース(赤色)とマスタード(青色)の違いもはっきりと区別できている。一方、 h に関してはマヨネーズとマスタードの分布が重なっており、真値からも若干ずれている。 h の真の値は三つの試料において0.442 ~ 0.518の間に収まっており、他の成分に比べて変化が小さいことが原因と考えられる。

また、背景パッチ除去の効果を評価するため、背景パッチを除去しない場合についての結果(図3(a))と比較を行った。なお、学習時に背景パッチを除去しない場合について

(a) 背景パッチを除去しない場合 (epoch = 500)



(b) 背景パッチを除去する場合 (epoch = 500)

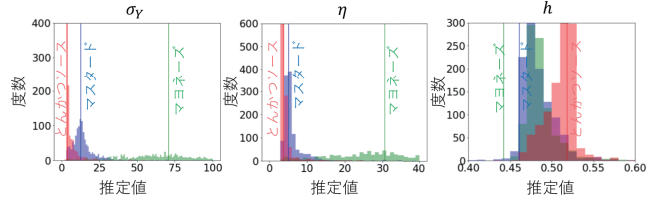
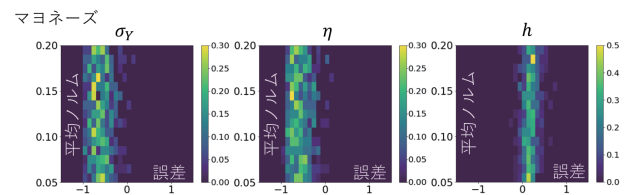


図3 3種の試料間で、テスト用のパッチによるパラメータ三成分の推定値のヒストグラムと真値を比較した。

も、テスト用のパッチからは背景パッチを除外した。これは学習時における背景パッチの効果を検証するためである。学習時に背景パッチを除去しない場合、 σ_Y についてはマスタードとマヨネーズの分布、 η についてはマヨネーズの分布が真値からずれており、また h については推定値が全体的に広く分布している。学習時に背景パッチを除去する場合には、こうした不具合が改善されている。また、図4に、マヨネーズのテスト用のオプティカルフロー全フレームから切り出した各パッチのノルム(各画素の速度ベクトルのノルムを求め、それらを平均したもの)と各成分値の推定誤差(真値と推定値との符号付き誤差)との関係を示す。カラーマップの色は、ある誤差、平均ノルムを持つパッチの頻度を表しており、平均ノルムが同一の値の各行について頻度の合計が1となるように正規化してある。背景を除去した場合(図4(b))は、除去を行っていない場合(図4(a))と比較して、誤差0のラインを中心に誤差の頻度の山が分布している傾向が見られる。他の2つの試料のテ

(a) 背景パッチを除去しない場合 (epoch = 500)



(b) 背景パッチを除去する場合 (epoch = 500)

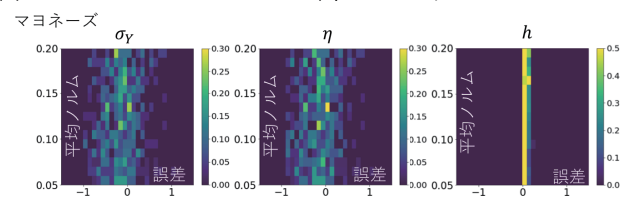
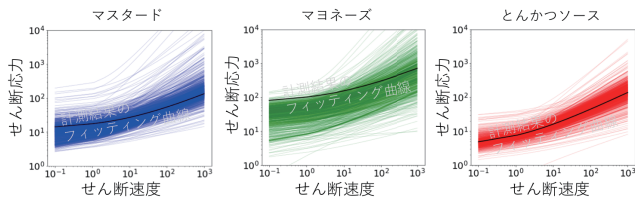


図4 マヨネーズのテスト用のフレームから切り出した各パッチについて、パッチ内の平均ノルム(縦軸)とパラメータの真値との相対誤差(横軸)を軸にとった2次元ヒストグラムを示す。

(a) 背景パッチを除去しない場合 (epoch = 500)



(b) 背景パッチを除去する場合 (epoch = 500)

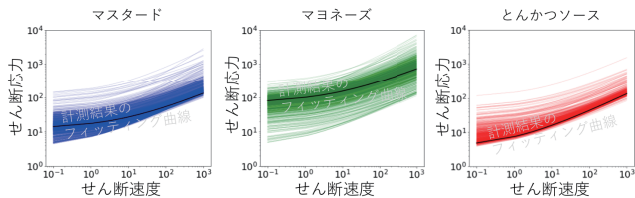


図 5 学習済みモデルに入力したテスト用のパッチから推定された流動曲線群と、真の流動曲線 (計測結果をハーシャルバルクレイモデルでフィッティングした曲線) とを比較した。

スト用パッチからも、同様の傾向が見られた。これらのことから、ハードコーディングした閾値による背景除去だけでも、推定精度向上の効果があると言える。

最後に、図 5 は、試料ごとにテスト用の各パッチについて、推定されたパラメータから流動曲線を復元し、計測で得られた流動曲線と比較している。背景パッチを除外した場合には、推定曲線は真の曲線 (黒線) まわりにより狭く分布している。

5. 結論と今後の展望

流動する流体の様子を撮影した動画から、流体の粘性を推定するための枠組みを提案した。本稿では、最終的に目指すシステムを実現するための前段階として、CNN により流体の粘性情報を推定するための特徴抽出が可能かどうか確認するための予備実験を行った。実験では、3 種類に限定した試料の実写動画のオプティカルフローから切り出したパッチを入力として、単純な CNN にその粘性情報を学習させた。また、学習、及び学習済みモデルの評価を行うためのデータセットとして、オプティカルフローの時系列の連続性を無視したフレーム単位での実験を行った。予備実験の結果から、テスト用の各パッチに対する流動性の推定曲線が真の流動曲線を中心に狭い範囲内に分布することがわかり、本枠組みが粘性情報の推定において有効であることが示唆された。

提案した枠組みにより、実際に実写映像から粘性情報を推定するには、いくつか解決すべきことがある。まず、本予備実験では、背景パッチの除去にハードコーディングした閾値を用いたが、attention network [20], [21] 等の手法を利用した自動的な背景除去を検討したい。また、各々のパッチから粘性情報を推定していたが、こうして得られた全パッチの粘性情報を統合して、対象流体としての粘性情報を出力する方法や、オプティカルフローの時系列の連続

性を担保できるような学習方法が必要である。次に、学習に用いる試料の数を増やし、粘性推定可能な流体の対象を増やしていきたい。そのための方法として、ハーシャル・バルクレイ流体のモデルパラメータ (σ_Y, η, h) の空間内において密にサンプリングを行い、得られたパラメータセットから CG により生成した動画群を学習用のデータセットとして利用する事も検討している。さらに、実験条件や撮影条件に関する仮定を緩和し、斜面の角度が正確に分らない状況や、一般の室内のような複雑な照明環境下で撮影された動画などについても適用できるようにする方法を検討していきたい。

謝辞

本研究は、JSPS 科研費 若手研究 (A) (17H04682), 新学術領域研究 (研究領域提案型) (18H05001) の助成によるものです。

参考文献

- [1] Sulsky, D., Zhou, S.-J. and Schreyer, H. L.: Application of a particle-in-cell method to solid mechanics, *Computer Physics Communications*, Vol. 87, pp. 236–252 (1995).
- [2] Nagasawa, K., Suzuki, T., Seto, R., Okada, M. and Yue, Y.: Mixing sauces: a viscosity blending model for shear thinning fluids., *ACM Trans. Graph. (Proc. of SIGGRAPH 2019)*, Vol. 38, No. 4, pp. 95:1–17 (2019).
- [3] Kawabe, T., Maruya, K., Fleming, R. W. and Nishida, S.: Seeing liquids from visual motion, *Vision Research*, Vol. 109, pp. 125–138 (2015).
- [4] Lucas, B. D. and Kanade, T.: An iterative image registration technique with an application to stereo vision, *IJCAI'81*, Vol. 2, pp. 674–679 (1981).
- [5] Stomakhin, A., Schroeder, C., Chai, L., Teran, J. and Selle, A.: A material point method for snow simulation, *ACM Trans. Graph. (Proc. of SIGGRAPH 2013)*, Vol. 32, No. 4, pp. 102:1–102:10 (2013).
- [6] Yue, Y., Smith, B., Batty, C., Zheng, C. and Grinspun, E.: Continuum foam: a material point method for shear-dependent flows, *ACM Trans. Graph.*, Vol. 34, No. 5, pp. 160:1–160:20 (2015).
- [7] Klár, G., Gast, T., Pradhana, A., Fu, C., Schroeder, C., Jiang, C. and Teran, J.: Drucker-prager elastoplasticity for sand animation, *ACM Trans. Graph. (Proc. of SIGGRAPH 2016)*, Vol. 35, No. 4, pp. 103:1–12 (2016).
- [8] Yue, Y., Smith, B., Chen, P. Y., Chantharayukhonthorn, M., Kamrin, K. and Grinspun, E.: Hybrid grains: adaptive coupling of discrete and continuum simulations of granular media, *ACM Trans. on Graph. (Proc. of SIGGRAPH Asia 2018)*, Vol. 37, No. 6, pp. 283:1–19 (2018).
- [9] Jiang, C., Gast, T. and Teran, J.: Anisotropic elastoplasticity for cloth, knit and hair frictional contact, *ACM Trans. Graph. (Proc. of SIGGRAPH 2017)*, Vol. 36, No. 4, pp. 152:1–14 (2017).
- [10] Rao, M. A.: *Rheology of fluid, semisolid, and solid foods*, Springer US (2014).
- [11] Dervisoglu, M. and Kokini, J.: Steady shear rheology and fluid mechanics of four semi solid foods, *Journal of Food Science*, Vol. 51, pp. 541–546 (2006).
- [12] Wang, B., Wu, L., Yin, K., Liu, L. and Huang, H.: De-

- formation capture and modeling of soft objects, *ACM Trans. Graph. (Proc. of SIGGRAPH 2015)*, Vol. 34, No. 4, pp. 94:1–94:12 (2015).
- [13] Wang, B., Kry, P. G., Deng, Y., Ascher, U. M., Huang, H. and Chen, B.: Neural material: learning elastic constitutive material and damping models from sparse data, *arXiv:1808.04931* (2018).
- [14] Hu, S., Zhang, Z., Xie, H. and Igarashi, T.: Data-driven modeling and animation of outdoor trees through interactive approach, *The Visual Computer*, Vol. 33, No. 6, pp. 1017–1027 (2017).
- [15] Dobashi, Y., Iwasaki, W., Ono, A., Yamamoto, T., Yue, Y. and Nishita, T.: An inverse problem approach for automatically adjusting the parameters for rendering clouds using photographs, *ACM Trans. Graph. (Proc. of SIGGRAPH Asia 2012)*, Vol. 31, No. 6, pp. 145:1–145:10 (2012).
- [16] Paulun, V. C., Kawabe, T., Nishida, S. and Fleming, R. W.: Seeing liquids from static snapshots, *Vision Research*, Vol. 115, pp. 163–174 (2015).
- [17] van Assen, J. J. R., Barla, P. and Fleming, R. W.: Visual features in the perception of liquids, *Current Biology*, Vol. 28, No. 3, pp. 452 – 458.e4 (2018).
- [18] Girshick, R. B., Donahue, J., Darrell, T. and Malik, J.: Rich feature hierarchies for accurate object detection and semantic segmentation, *2014 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 580–587 (2014).
- [19] Lecun, Y., Bottou, L., Bengio, Y. and Haffner, P.: Gradient-based learning applied to document recognition, *Proceedings of the IEEE*, Vol. 86, No. 11, pp. 2278–2324 (1998).
- [20] Fu, J., Zheng, H. and Mei, T.: Look closer to see better: recurrent attention convolutional neural network for fine-grained image recognition, *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 4476–4484 (2017).
- [21] Yoo, D., Park, S., Lee, J.-Y., Paek, A. S. and Kweon, I. S.: AttentionNet: aggregating weak directions for accurate object detection, *Proceedings of the 2015 IEEE International Conference on Computer Vision (ICCV)*, pp. 2659–2667 (2015).