



HAL
open science

Learning to Measure the Static Friction Coefficient in Cloth Contact

Abdullah-Haroon Rasheed, Victor Romero, Florence Bertails-Descoubes,
Stefanie Wuhrer, Jean-Sébastien Franco, Arnaud Lazarus

► **To cite this version:**

Abdullah-Haroon Rasheed, Victor Romero, Florence Bertails-Descoubes, Stefanie Wuhrer, Jean-Sébastien Franco, et al.. Learning to Measure the Static Friction Coefficient in Cloth Contact. CVPR 2020 - IEEE Conference on Computer Vision and Pattern Recognition, Jun 2020, Seattle, United States. pp.9909-9918, 10.1109/CVPR42600.2020.00993 . hal-02511646

HAL Id: hal-02511646

<https://inria.hal.science/hal-02511646v1>

Submitted on 23 Mar 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Learning to Measure the Static Friction Coefficient in Cloth Contact

Abdullah Haroon Rasheed¹, Victor Romero^{1,2}, Florence Bertails-Descoubes¹,
Stefanie Wuhrer¹, Jean-Sebastien Franco¹, Arnaud Lazarus²

¹ Univ. Grenoble Alpes, Inria, CNRS, Grenoble INP, LJK, Grenoble, France

² Sorbonne Université, Institut Jean Le Rond d’Alembert, Paris, France

abdullah-haroon.rasheed@inria.fr

Abstract

Measuring friction coefficients between cloth and an external body is a longstanding issue in mechanical engineering, never yet addressed with a pure vision-based system. The latter offers the prospect of simpler, less invasive friction measurement protocols compared to traditional ones, and can vastly benefit from recent deep learning advances. Such a novel measurement strategy however proves challenging, as no large labelled dataset for cloth contact exists, and creating one would require thousands of physics workbench measurements with broad coverage of cloth-material pairs. Using synthetic data instead is only possible assuming the availability of a soft-body mechanical simulator with true-to-life friction physics accuracy, yet to be verified. We propose a first vision-based measurement network for friction between cloth and a substrate, using a simple and repeatable video acquisition protocol. We train our network on purely synthetic data generated by a state-of-the-art frictional contact simulator, which we carefully calibrate and validate against real experiments under controlled conditions. We show promising results on a large set of contact pairs between real cloth samples and various kinds of substrates, with 93.6% of all measurements predicted within 0.1 range of standard physics bench measurements.

1. Introduction

Computer vision techniques are powerful in analysing the shape and motion of physical objects from mere images. Yet, accurately inferring physical properties from visual data is still a challenging problem [33]. Our work aims at estimating physical parameters from the rich dynamic effects that can be observed in cloth motion. Applications range from non-invasive estimation of cloth properties in textile engineering to realistic clothing synthesis for virtual and augmented reality. So far, most existing methods for cloth acquisition have focused on estimating cloth material parameters, that is, its stiffness and mass [3, 35, 11]. Friction

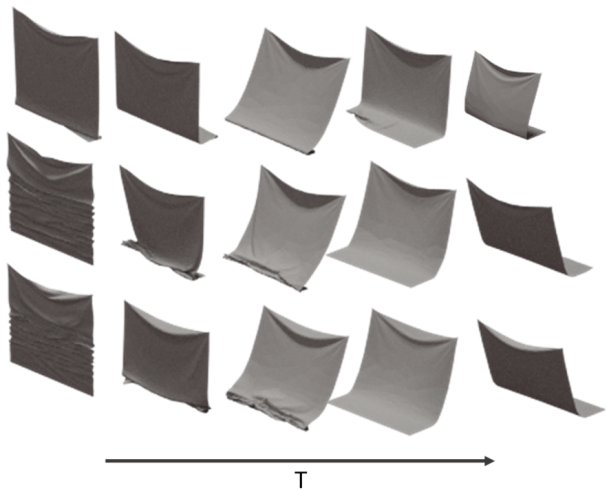


Figure 1. Frames from the same timestep for three cloth motion sequences simulated with the same material but different friction coefficients at contact (μ). Top: $\mu = 0.0$, Centre: $\mu = 0.5$, Bottom: $\mu = 0.6$. Differences are significant between $\mu = 0.0$ and $\mu = 0.5$, but more subtle between $\mu = 0.5$ and $\mu = 0.6$.

tion at contact is a much less studied problem due to the difficulty of modelling, measuring, and simulating dry frictional contact accurately. However, friction has a high impact on the overall cloth dynamics, as illustrated in Figure 1.

The problem of measuring a friction coefficient accurately in cloth is a longstanding problem, mainly studied in textile engineering [29, 24]. Better accuracy often comes at the expense of tedious and generally invasive protocols with specialised mechanical sensors and hardware. Yet, the visual variability induced by friction shown in Figure 1 hints at the possibility of a purely vision-based protocol, discriminating the different motion and folding patterns observed under friction variation. This general idea has recently been leveraged by a few studies in physics to infer friction coefficients from contacting slender structures like hair strands [7] or stiff ribbons [26]. While such methods effectively exploit the connection between observed geom-

entry and friction, they are only valid for very specific classes of materials and conditions – such as setups where gravity plays a negligible role. A major challenge for cloth is that, larger friction coefficient differences are easily distinguished, but smaller differences yield increasingly subtle visual differences (Fig. 1).

In this work, we propose a first step towards a generally applicable vision-based method to estimate the dry friction coefficient between cloth and a contacting surface. To this goal, we use a Long-term Recurrent Convolutional Networks (LRCN) [12] to regress from an input video showing a cloth motion under contact to both material parameters and a dry friction coefficient. Our network is trained on sequences of simulated cloth deformations. We leverage the recent simulator ARGUS [18] for capturing dry frictional contact in cloth, and search parameter settings that lead to sufficiently accurate results for our needs. Our results indicate that the network, trained exclusively on simulated data, does generalise to real videos showing similar cloth motions; and this in spite of the differences between the renderings of the simulations and the real captures.

In scenarios where capturing significant amounts of real data is prohibitive, many works have explored the idea of training on simulated data with the aim to generalise inference on real data. These include approaches for cross domain transfer learning [15, 30] and interpretable low dimensional representation learning [14]. Our approach can be considered orthogonal to such efforts as we present a protocol to calibrate simulated data generation by experimental verification, before learning a model from this data.

In summary, we propose the following contributions:

- We present a deep learning based algorithm to solve the inverse problem of parameter estimation for cloth simulation. Our algorithm measures the material and friction properties of cloth given a video sequence.
- To generate physically correct training data, we validate the ARGUS simulator physically and calibrate its accuracy against a constrained, measurable real-world physics set of experiments. Using this calibration, we produced a dataset containing 3840 simulated videos. We acquire a test dataset of 315 real videos in controlled conditions closely matching the simulation setup and make baseline estimates of friction values within a measurement error range. The combined simulated and real dataset provides a valuable correspondence for learning cloth material and friction properties purely from simulation and testing on real data. We have publicly released this dataset¹.
- We verify the accuracy of our deep learning model on held-out simulation data and achieve an error of < 0.1 on 95.5% of the synthetic data. We further demonstrate the validity of our learned model by achieving an error < 0.1 from the baseline range on 93.6% of real data.

¹<https://hal.inria.fr/hal-02511646v1>

2. Related Work

Inferring material properties of an object from geometric or visual data has been studied in various communities including computer vision, computer graphics, textile engineering and physics. We first introduce the general model of friction we consider and existing techniques to measure it, before focusing on slender elastic structures such as cloth.

Dry friction model Dry friction is a force that opposes the relative motion of two solid objects. The way in which the interface of two compliant objects in contact evolves when subjected to load is complex, and finding the precise law that describes this phenomenon remains an open problem in physics and mechanics. Our work is based on Amontons-Coulomb’s law [2] for friction, which is a commonly used model that successfully approximates this complex scenario at the macroscopic scale. In this model, surfaces in contact interact throughout normal and shear forces, and sliding occurs when the ratio between the shear and normal force reaches a threshold value, called the *static friction coefficient*, which is independent of the area of contact and depends only on the roughness of the interacting surfaces.

Friction measurement Estimating friction and material parameters jointly from visual data has recently become a topic of study in computer vision. Miguel *et al.* [23] consider estimating internal friction in cloth from geometric information. Internal friction is however different from static friction at contact since it models internal dissipation within the cloth, hence this technique does not apply to our case. Wang *et al.* [34] propose to use a generative model to estimate friction and material parameters. Unlike our work, this method is targeted at rigid objects. Zhang *et al.* [39] analyse visible reflections, while Yuan *et al.* [37] combine visual and haptic data to estimate friction information. Both works present a static joint estimation of material and friction based on visual attributes of the material, whereas we focus on the dynamic behaviour of cloth under frictional contact and wish to estimate the friction coefficient directly.

In textile engineering, friction estimation has been studied using invasive techniques [13, 29, 24, 21]. Some studies in other fields connect the perception of friction (*i.e.* feeling) to actual measurements [6, 20] and visual features [5]. These works show that visual features correlate to friction information using perceptual studies.

In physics, a few isolated studies consider inferring dry static friction coefficients from contacting elastic slender structures. In particular, the friction of a hair fibre can be inferred by the geometric configuration of a relaxed knot formed by the hair [7], and the friction of a relatively stiff isotropic ribbon can be inferred by the geometric configuration of the ribbon’s bending and slipping behaviour when pushed onto a substrate [26]. While these studies only hold in specific scenarios with strict boundary conditions, they

demonstrate the relationship between geometric configurations and friction for elastic slender structures. Our work of estimating cloth friction based on videos is inspired by such intriguing studies. Note that in order to obtain reference friction measurements for evaluation, we use a classical inclined-plane protocol [10, 29].

Material parameter estimation The problem of estimating material parameters from geometric or visual data has been studied in different scenarios and for various types of materials, including soft deformable materials [28, 36, 31]. To our knowledge, recovering material information from cloth has always been formulated as a fitting problem, consisting of three steps: first, considering a material model as a function of deformation of cloth, then capturing deformations, and finally fitting the function to captured data. One approach is to control boundary conditions by setting up a physical apparatus for yielding deformations [22, 9, 32]. To avoid capturing deformations under controlled conditions, some works have explored extracting material parameters from casually captured videos of cloth [3, 17]. Bouman *et al.* [4] capture stiffness and density of fabrics from video using handcrafted discriminatory features.

The closest methods to ours identify cloth material parameters from videos [35, 11, 19]. Yang *et al.* [35] train a neural network on simulated cloth deformations and use the resulting architecture to infer bending and stiffness parameters. This approach follows a recent trend in computer vision and machine learning, where deep neural networks are trained with simulated input data and tested on real data. Training with simulated input offers the advantage of yielding sufficiently large training datasets, hence we follow the same approach here. However, unlike previous methods, we carefully validate and finely tune our simulator for physical realism, through an experimental protocol described in Sec. 4.2.1. Liang *et al.* [19] propose a computationally efficient way of estimating material parameters with a differentiable cloth simulator. Davis *et al.* [11] estimate material parameters based on videos that show small vibrating motions, and apply this method to fabrics.

Note that none of the works discussed in this section consider estimation of friction at contact.

3. Background

In this section we describe the elastic model used for our material properties, and the Amontons-Coulomb’s law for friction. We chose these models as they are known to be realistic in the range of macroscopic deformations.

3.1. Material Parameter Space

To obtain physically accurate material parameter settings in the simulator, we leverage the work of Wang *et al.* [32] that encodes the material properties of ten representative classes of cloth ranging from very soft to stiffer materials.

Material parameters are encoded as three parameters of the cloth simulator ARCSIM [25]: a linear mass density, the coefficients of the strain-stress matrix and the coefficients of bending stiffness. This model has been previously used by Yang *et al.* [35] for recovering cloth material parameters.

3.2. Amontons-Coulomb’s Law for Friction

Our work is based on Amontons-Coulomb’s law for dry friction as this model successfully approximates the macroscopic behaviour of two solid surfaces at contact. In a simplified version that we shall use here (no distinction between static and dynamic friction coefficient), this law defines the friction coefficient μ as a threshold value for the stick to slip transition for two contacting surfaces, and as the coefficient relating normal and tangential forces during sliding. More specifically, in our scenario, let \vec{R} denote the reaction force on the surface in contact with a piece of cloth. We can divide \vec{R} into two components: the force component normal to the surface and the one tangential to the surface, denoted by \vec{P} and \vec{Q} , respectively. The force \vec{P} keeps the two contacting surfaces from interpenetrating, and \vec{Q} opposes relative displacements between the two surfaces in contact. Fig. 3 illustrates these forces for a strip of material that is pushed onto a substrate.

Our dry friction law distinguishes three possible states depending on the values of \vec{P} and \vec{Q} . There is no contact for $\vec{P} = \vec{Q} = \vec{0}$, the two surfaces stick if $\vec{Q} \leq \mu\vec{P}$ and they slip if $\vec{Q} > \mu\vec{P}$ (the case $\vec{Q} > \mu\vec{P}$ is not admissible).

4. Data Generation

Our first contribution consists in generating a dataset of closely matching captured videos and simulations with, for each video, corresponding material classes and friction coefficients. We choose a simple motion that can easily be replicated with a real piece of cloth while containing representative material classes and friction parameters. In particular, we consider a *drop and drag* motion, in which a square of cloth of side length 20 cm, suspended by its corners, is dropped vertically on a substrate floor and then dragged back and forth as shown in Fig. 4.

4.1. Real Data Capture

We now present our experimental data capture setup. Since we aim at predicting material and frictional properties from videos, we constrain the setup of generating real data to a controlled environment in order to remove sources of variation other than material properties and friction.

Cloth materials We use nine materials in bright colours, out of which eight are close in composition and density to material classes defined in [32], and one (silk) which is not covered in these material classes. Details on the materials can be found in the supplementary material. From each material we laser cut 20x20 cm pieces with 1x1cm holders at two corners of one edge for clamping purposes.

Substrates We choose seven substrates that, combined with our material samples, allow to cover a wide range of friction behaviours. In particular, chosen substrate materials are aluminium, aluminium-PET, ceramic, rough glass, smooth glass, polyester-mirror and stainless steel.

Experimental setup Using two translational stages (Thorlabs LTS-300M) in front of a black background, we control the drop and drag movement of the cloth for repeatability. The total movement is characterised as follows. First, the material is held 1 cm above the substrate and dropped by 6 cm in a movement, which accelerates from rest at 10 mm/s^2 until it reaches 10 mm/s . Subsequently, the system decelerates to reach 6 cm of total displacement. Immediately, the second motorised stage, for the horizontal displacement, starts the horizontal drag by accelerating at 10 mm/s^2 until it reaches 10 mm/s . This velocity is kept constant until the system starts decelerating to achieve a total displacement of 30 cm. We repeat this process forward and backward twice per experiment to allow the observation of a sufficient number of drag and wrinkling patterns.

Video acquisition With a calibrated camera, we record the cloth motion from the viewpoint shown in Fig. 4, which allows to observe wrinkling patterns. The whole video contains around 300 frames, which corresponds to 2.4 frames per second. For each material / substrate pair, we repeat the experiment 5 times, leading to 5 synchronised videos.

Reference friction measurements A quantitative evaluation of the predicted friction coefficients requires measurements of the friction coefficient μ for each material / substrate pair. Accurately measuring μ is an involved process in physics and mechanics that often follows invasive protocols. For simplicity, we capture merely a reference measurement for μ using the non-invasive inclined plane technique [10, 29]. We believe that the friction in the inclined plane scenario is close to the one in the drag situation, and hence its quantification provides a good reference estimation for our purposes. The inclined plane protocol measures friction by placing an object on an inclined plane, and by increasing the slope of the plane until the object starts slipping. The friction coefficient is then computed based on the slope of the plane at the point where the object slips.

To ensure that the reference measurement is robust, we test for each material / substrate pair different locations and orientations of the cloth on the substrate and do not find a noticeable difference in the slippage angle. Our physical setup of the inclined plane further gives rise to a measurement error described in supplemental material. Fig. 2 shows a histogram of the values of μ that were measured for all material / substrate pairs. Note that our dataset covers a wide range of friction coefficients.

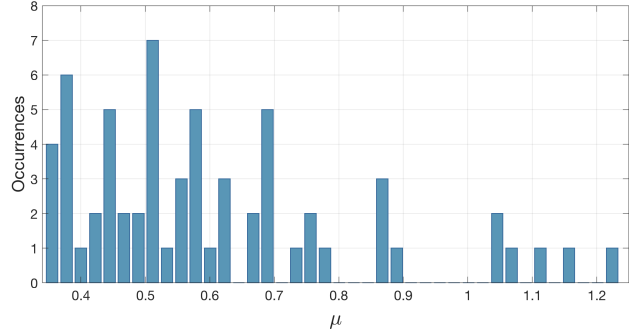


Figure 2. Histogram of μ for all material / substrate pairs.

4.2. Simulated Data Generation

For training and testing, we simulate a physically accurate dataset that closely resembles the videos captured using the experimental protocol. Implementing Amontons-Coulomb’s law poses difficulties in practice because the force response is nonsmooth. This is further complicated by the requirement for discrete representations in both space and time. We use the ARGUS implementation [18] to simulate cloth deformations for two reasons. First, this state-of-the-art simulator uses an efficient nonsmooth solver to model friction behaviour for mesh-based systems. Second, a free implementation of ARGUS is available on [github](#).

4.2.1 Physical Validation of the ARGUS Simulator

To our knowledge, most frictional contact solvers for cloth, including ARGUS, have never been validated against real experiments. To verify that ARGUS produces physically accurate simulations, we compare simulations produced by ARGUS against physical experiments. This comparison is performed in a constrained setting, in which the friction behaviour can be analytically derived. The reason for restraining the experimental setting is that a verification against the ground truth physical behaviour is not possible in more complex scenarios, as measuring static friction accurately remains a challenging problem in physics and mechanics.

The constrained experiment we use has recently been proposed by Sano *et al.* [26] in the physics community, and is depicted in the supplemental video. The experiment considers the deflection of a strip, clamped at its top, that is vertically pushed against a substrate with a vertical strain ϵ_y , as illustrated in Fig. 3-right. Because of friction, this strip remains *pinned* at its bottom for small ϵ_y , which is shown in red in Fig. 3. The more the strip is pushed down, the higher the frictional force holding the strip pinned. The strip is geometrically constrained, hence depending on the value of the friction coefficient, this system has two different outcomes. First, the elastic forces can overcome friction, and the strip abruptly slips. This regime is called *slipped*, and

illustrated in green in Fig. 3-left. Second, for rougher surfaces, the frictional forces are high enough to prevent the strip from slipping and it deforms until there is areal contact between the strip and the substrate. This regime is called *partially slipped*, and illustrated in blue in Fig. 3-left. Sano *et al.* show analytically that in the case where gravity is negligible, these outcomes are purely regulated by the friction coefficient, and independent of the strip’s material parameters. That is, the deformed shape of the strip only depends on the vertical strain ϵ_y . The analytical solution allows to calculate at what vertical strain the strip will slip, which is shown as black curve in Fig. 3-left.

When simulating the experiment of Sano *et al.*, a physically accurate simulator should simulate strips that slip when the vertical strain lies on the black curve. We succeeded in matching this theoretical result with the output of ARGUS, but this required carefully setting different numerical parameters, as explained in the following. First, the temporal resolution of the simulation can be controlled by a timestep parameter between two adjacent frames. Second, the spatial resolution can be controlled by either forcing a static discretisation to be used (whose resolution needs to be provided), or by allowing an adaptive remeshing of the geometry, in which case a rate of refinement needs to be provided. Finally, ARGUS also provides contact solver tolerance values that can be controlled by a residual value and a maximum iteration count of the solver, and damping forces are allowed to stabilise the simulations. In our experiments, we found that all of these numerical parameters heavily influence the geometry of the resulting simulations, and thus should be chosen carefully. We justify the necessity of this procedure by showing examples in the supplemental material of incorrect regimes occurring when these parameters are not adequately set. Our rationale for finding an adequate set of numerical parameters was, first, to set a fixed and high resolution for the cloth mesh, avoiding small energy perturbations due to remeshing; then, mimic a quasi-static experiment by increasing damping forces, thus removing spurious dynamical vibrations of the cloth; finally, find the right level of accuracy by sufficiently decreasing the timestep of the simulation, as well as the solver tolerance, to the point where further refinement would lead to indistinguishable results at our observation scale. In our comparisons, we use a timestep of 0.5 milliseconds, a mesh resolution of 149 vertices for a 20 cm long ribbon, a maximum number of iterations of 50000 and a damping value set to 0.002. We repeat the simulation using 2 materials having different Young’s moduli. Our simulations yield the purple curve shown in Fig. 3-left, which matches the black analytically computed curve almost perfectly. Fig. 3-right shows an overlay of the physical strip with our simulated result shown as a red curve, and they are in agreement. Please watch the supplemental video for animated illustrations.

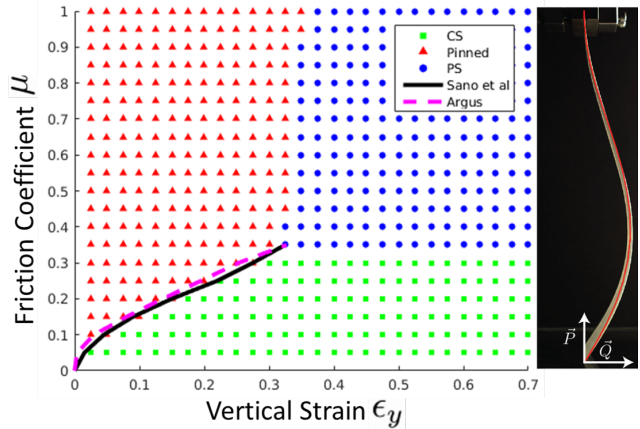


Figure 3. Physical validation of ARGUS [18] under a constrained setting that is well understood in physics [26]. After proper calibration of the simulator, we observe simulations (dotted curve) that are in very good agreement with the theory (black curve).

These results certify that under correct numerical calibration, ARGUS generates simulations that match reality in a constrained setting, under negligible effect of gravity. In the following, we work under the hypothesis that the same numerical calibration will guarantee highly realistic results while simulating cloth in a more general setting, when gravity is no more negligible, and where both friction and material parameters influence the geometry of the cloth.

4.2.2 Dataset Generation

To generate the cloth simulation dataset, we use the calibrated parameter setting of ARGUS and implement the exact path of the motorised stage used for the physical data captures, which is analytically accessible, leading to a high temporal synchronisation between the real data and the simulations. Furthermore, the calibrated camera parameters of the real data captures are used to render a similar view for our simulations. This results in simulations that can be considered physically valid to train a discriminatory model.

The drop and drag motion is simulated for the 10 material classes measured by Wang *et al.* [32]. For each material, 16 friction coefficients evenly distributed between 0.0 and 1.5 are explored, which represents a reasonable range for fabrics according to values tabulated in [13] and is in agreement with our reference measurements. The resulting simulated 3D sequences contain 300 frames each. To generate a dataset of 2D videos, each simulated sequence is rendered using 8 different texture maps and from 3 different viewpoints, using the free Blender software. The addition of texture variation increases generalisation of learning as shown in Sec. 6.2.1. One of the rendered viewpoints is calibrated based on the real data to replicate the experimental settings in our simulated data. Two additional viewpoints, one frontal and one side view, are selected to increase visual

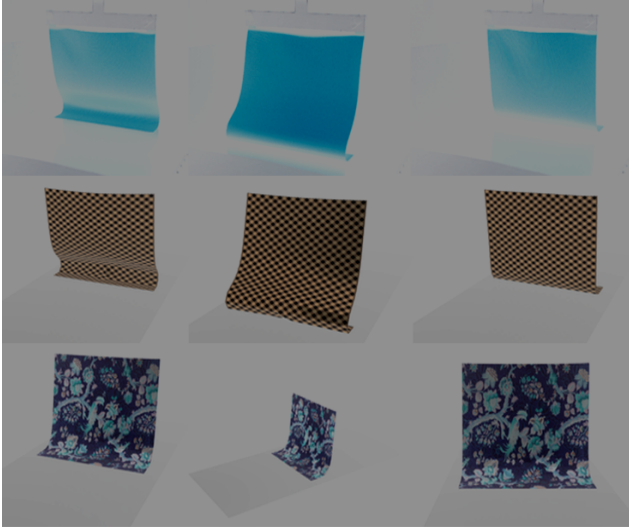


Figure 4. Dataset Examples: First and second row show corresponding frames from real and synthetic data respectively. Third row shows 3 viewpoints rendered in the simulated dataset. Colours inverted for visibility, see supplemental video for originals.

variation. Furthermore, to match the experimental setup and reduce variability due to environmental factors, we render a dark background and a substrate floor. The viewpoints for simulated data, alongside real data are depicted in Fig. 4. Further illustrations of the dataset are provided in supplemental material.

5. Pipeline to predict static friction coefficient

We now specify our pipeline to estimate friction of cloth from a 2D input video. Since the input to our model is a sequence of images, we use a Long-term Recurrent Convolutional Network (LRCN) model similar to those used for action classification [12]. We are inspired by a previous model that recovers a material class label of cloth from an input image sequence [35]. A variant of this model, called *baseline model* in the following, is explained in Sec. 5.1.

While the baseline model is very effective as predicting a material class, it performs significantly worse when trained to predict friction coefficients (see Sec. 6 for details). We therefore propose a novel pipeline for this task that conditions friction on material classes, as outlined in Sec. 5.2.

5.1. Baseline Model

In the baseline model by Yang *et al.* [35], convolution and pooling layers are used to extract image features from each frame. Their architecture of this feature extraction block is a modified version of AlexNet [16]. We replace this architecture by a simplified version of VGG16 [27] with ReLU activations as shown in Table 1. The weights are shared between all feature extraction blocks. Let

$$f_i = CNN_{VGG}(I_i) \quad (1)$$

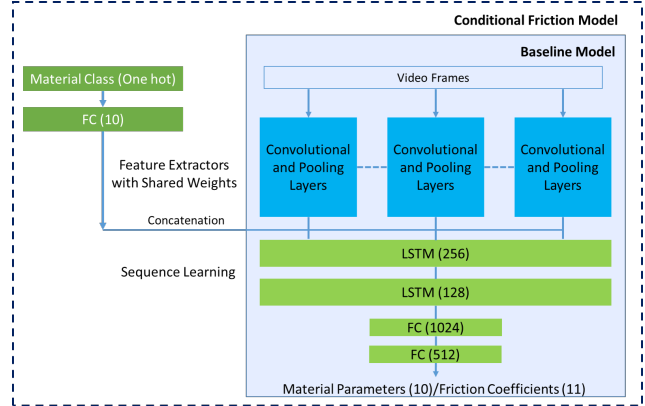


Figure 5. Proposed architecture to estimate friction conditioned on material parameters. The coloured inlay shows the baseline model, which is augmented with material class information to form the conditional friction model.

Name	Description
Input	Image 224, 224, 3
conv1a	Conv 3 x 3, 64, ReLU
conv1b	Conv 3 x 3, 64, ReLU
maxpool1	3 x 3, stride 2 x 2
conv2a	Conv 3 x 3, 128, ReLU
conv2b	Conv 3 x 3, 128, ReLU
maxpool2	3 x 3, stride 2 x 2
conv3a	Conv 3 x 3, 256, ReLU
maxpool3	3 x 3, stride 2 x 2
conv4a	Conv 3 x 3, 512, ReLU

Table 1. Architecture details for a feature extractor block.

denote the image features extracted from frame I_i , where CNN_{VGG} is the simplified VGG16 of Table 1.

A sequence of learned image features f_1, f_2, \dots, f_k is then passed to long short term memory (LSTM) layers, which extract temporal information. The output of these layers is finally passed to fully connected layers to learn a function from the extracted spatial and temporal features to the data labels. This can be written as

$$\hat{h} = FC(LSTM(f_1, f_2, \dots, f_k)), \quad (2)$$

where \hat{h} is the final likelihood computed for each material class label, $LSTM$ is a set of two LSTM layers and FC is a set of two fully connected layers. The architecture is depicted in Fig. 5. This architecture is trained with a standard categorical cross-entropy classification loss, and for prediction, the class label with the highest likelihood is reported.

5.2. Conditional Friction Model

The baseline model is significantly worse at predicting friction than at predicting material class. The reason is that different materials combined with different friction can be visually very similar. For a fixed material, however, differ-

ent friction behaviour is typically visually distinctive. Motivated by this observation, we estimate the friction coefficient using a model that is conditioned on material classes.

Model The model is shown in Fig. 5. In addition to a sequence of 2D video frames $I_i, i = 1, \dots, k$, the model takes as input a material class label. The material class label is represented by a one-hot vector m which is passed to a fully connected layer with softmax activation. For this architecture, the video frames are processed using the same convolution and pooling layers as for the baseline model. The material information is then cloned for each input frame and concatenated with the feature vectors of each frame, before being passed to the LSTM layers and on to the fully connected layers. That is, the vector \hat{y} containing the likelihood for each friction class label is computed as

$$\hat{y} = FC(LSTM(FC_s(m) \frown f_1, \dots, FC_s(m) \frown f_k)), \quad (3)$$

where FC_s denotes a fully connected layer with softmax activation and \frown is the concatenation operator.

Training The training loss can be written as

$$\mathcal{L} = -\log P(y|I_1, I_2, \dots, I_k, m), \quad (4)$$

where y is the friction label provided for the training examples. This loss is implemented as the categorical cross-entropy loss function. In both models, we use dropout layers between fully connected layers for regularisation. We first train the baseline model to predict the material class, which is subsequently used as label to train the conditional friction model. We use predicted material classes instead of ground truth labels as this may allow the conditional model to learn the uncertainty of the baseline model.

Data representation A data point for training or testing consists of 30 frames sampled at regular intervals from a video sequence and corresponding material class and friction coefficient as label. We select an input size of 30 frames as the maximum number of frames from a single video sequence that we can fit during the training cycle on a single GPU, without running into memory constraints. We use a split of 80, 10, 10 percent for training, cross-validation and testing, respectively, which results in training on 92160 images from 3072 video sequences and testing on 11520 images from 384 video sequences.

Model initialisation We observed experimentally that model initialisation is important for training convergence. Intuitively, this might be due to the fact that certain viewpoints provide better discriminatory information than others. To obtain training convergence in practice, we therefore train our models progressively by adding one rendered viewpoint from the dataset in each training cycle to our training data. After training simultaneously on all viewpoints, we finely tune our model on the viewpoint which is calibrated based on our experimental setup.

Prediction At test time, our pipeline first uses the baseline model to infer a material parameter and uses this to infer a friction coefficient with the conditional model.

6. Results

In this section we present and analyse the results of testing our pipeline on both synthetic and real test data. For synthetic data we present a comparison between our conditional friction estimator and baseline models and evaluate the model’s generalisation capability. For real data, we present our predictions and discuss limitations. The supplemental material provides further details.

6.1. Implementation and Evaluation Details

Our implementation uses Keras [8] and Tensorflow [1]. While we experimented with various optimisers, we empirically observed Adadelta [38] to converge faster. We use a learning rate of 1.0 and a decay factor of 0.95. The training time of our model and the baseline is around 8 hours with a single NVIDIA TitanX GPU and the training converges in circa 30 epochs.

Evaluation protocol As our goal is to build a friction measurement protocol for cloth, we evaluate our friction prediction by considering the absolute difference between the predicted value and the ground truth. As the reference measurement r for real data is only known up to a measurement error e , we consider any value within the interval $r \pm e$ as having no error, and report the absolute difference of our prediction to this interval. This provides an optimistic estimate of the error as the error is calculated from the range and not from an absolute value. For material parameter estimation we report the top-1 and top-2 accuracy.

6.2. Results on Simulated Test Data

To compare the performance of our conditional model with the baseline model, we train both on our simulated dataset. We use 10% of our dataset for testing.

Task	Material Estimation		Friction Estimation	
	Acc	Top-2 Acc.	Err <0.1	Err <0.2
Baseline	99.5 %	100.0 %	78.6 %	88.5 %
Conditional	-	-	95.5 %	99.2 %

Table 2. Results on Simulated Test Data

The baseline model performs vastly better at predicting material parameters than at predicting friction coefficients. The conditional estimation model performs better at predicting friction coefficients than the baseline model which indicates that the material and friction parameters are not decoupled in the global behaviour of the cloth, and that adding material information as an input parameter reduces the search space. Furthermore, the error distribution for test data indicates that our model learns coherently.

6.2.1 Generalisation to unseen textures

We render our training data with different textures to make the model agnostic to appearance variations. To test this generalisation ability, we render our simulations with a texture that has not been seen by the model during training. We test our model on 160 sequences, and our model achieves an error of < 0.1 on 87.2%, and an error of < 0.2 on 95% of the samples. This shows that the model’s predictions do not degrade dramatically if an unseen texture is encountered, implying that the model is capturing more nuanced spatio-temporal phenomena to classify friction behaviour.

6.2.2 Generalisation to unseen viewpoints

We demonstrate the model’s generalisation ability to unseen viewpoints. Starting with the camera position from one of our simulated viewpoints, we rotate the camera origin by ± 5 degree intervals on either side to generate 6 viewpoints which are 5, 10 and 15 degrees apart on either side. Afterwards we randomly select 20 material, friction and texture combinations for each viewpoint and render them as our test data. We report the accuracy for each viewpoint in Table 3. While the accuracy degenerates with unseen viewpoints, the decay happens progressively based on the difference in viewpoint from the original one.

	-15°	-10°	-5°	5°	10°	15°
Error < 0.1	65 %	75 %	80 %	75 %	65 %	50 %
Error < 0.2	70 %	85 %	95 %	80 %	80 %	65 %

Table 3. Results on unseen viewpoints.

6.3. Results on Real Test Data

We present results on real data captured through our experimental setup presented in Sec. 4.1. Our dataset contains 5 videos for each material / substrate pair, leading to 5 test datapoints. Our model provides a unique friction coefficient label for 76.2% of the material / substrate pairs, *i.e.* the same predicted label for all 5 videos from that pair. For the remaining 23.8% of the material / substrate pairs, we take the median value among 5 videos as the predicted label. Fig. 6 shows the cumulative error plot for this test data.

Analysis and Limitations On real data, our model achieves an error of < 0.1 on 61.9% of data on all substrates (see red curve of Fig. 6). Our test data includes silk, for which no similar material is present in the simulated training data, and our method achieves an error of < 0.1 of over 60%. The decay in model accuracy compared to simulated test data can be mostly attributed to the presence of reflective substrates in the real data, a factor that is not modelled in our training data. The light blue curve of Fig. 6 shows that the model accuracy is significantly worse on substrates which are reflective. Removing these substrates from the evaluation yields a error of < 0.1 for 86.1% of data, as

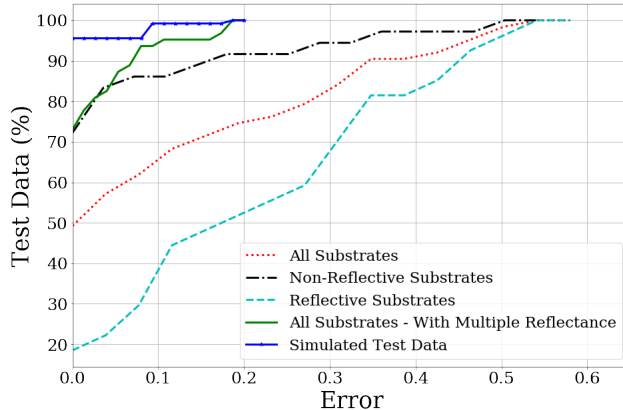


Figure 6. Cumulative error plots for real data (red, black, light blue, green) and simulated test data (dark blue).

shown in the black curve of Fig. 6. This shows that the model has indeed learned the friction behaviour of cloth which is transferable from simulated to real data, barring the confounding factors. One way to mitigate this degeneration on reflective substrates is to vary the reflectance of the substrate floor during rendering of our simulated training data. After adding reflective substrate renderings, we obtain an error of < 0.1 for 93.6% of real test data on all substrates, as shown in the green curve of Fig. 6. Note that adding the variation to the training data improves the results overall, yielding better results than training and testing without reflective materials (black curve). The supplemental material provides a more detailed analysis of these results.

7. Conclusion and Future Work

We propose the first protocol for vision-based measurement of dry friction in cloth contact. It is grounded firstly in the careful analysis and setup of a physics based frictional contact simulator, which we calibrate with a set of real world experiments, and secondly leverages recent CNN-based classification methods. These contributions open interesting future directions. First, the paper provides an architecture that can be further enhanced and compared against, using the dataset we provide to the community. Second, it paves the way towards estimation of friction in-the-wild by progressively relaxing the video-acquisition protocol, with interesting applications for non-invasive physics measurements, finer-grain capture of real surfaces, and physically accurate re-simulations of pre-observed surfaces. Third, the successful calibrated use of synthetic simulator-based training could be transposed to other inverse parameter estimation problems.

Acknowledgements This work was supported in part by the ERC grant GEM (StG-2014-639139).

References

- [1] Martín Abadi, Ashish Agarwal, Paul Barham, Eugene Brevdo, Zhifeng Chen, Craig Citro, Greg S. Corrado, Andy Davis, Jeffrey Dean, Matthieu Devin, Sanjay Ghemawat, Ian Goodfellow, Andrew Harp, Geoffrey Irving, Michael Isard, Yangqing Jia, Rafal Jozefowicz, Lukasz Kaiser, Manjunath Kudlur, Josh Levenberg, Dan Mané, Rajat Monga, Sherry Moore, Derek Murray, Chris Olah, Mike Schuster, Jonathon Shlens, Benoit Steiner, Ilya Sutskever, Kunal Talwar, Paul Tucker, Vincent Vanhoucke, Vijay Vasudevan, Fernanda Viégas, Oriol Vinyals, Pete Warden, Martin Wattenberg, Martin Wicke, Yuan Yu, and Xiaoqiang Zheng. TensorFlow: Large-scale machine learning on heterogeneous systems, 2015. Software available from tensorflow.org. **7**
- [2] Guillaume Amontons. De la resistance causée dans les machines. *Mémoires de l'Académie royale*, 4:257–282, 1699. **2**
- [3] Kiran S Bhat, Christopher D Twigg, Jessica K Hodgins, Pradeep K Khosla, Zoran Popović, and Steven M Seitz. Estimating cloth simulation parameters from video. In *Proceedings of the 2003 ACM SIGGRAPH/Eurographics symposium on Computer animation*, pages 37–51. Eurographics Association, 2003. **1, 3**
- [4] Katherine L Bouman, Bei Xiao, Peter Battaglia, and William T Freeman. Estimating the material properties of fabric from video. In *2013 IEEE International Conference on Computer Vision*, pages 1984–1991, 2013. **3**
- [5] Martim Brando, Kenji Hashimoto, and Atsuo Takanishi. Friction from vision: A study of algorithmic and human performance with consequences for robot perception and teleoperation. In *2016 IEEE-RAS 16th International Conference on Humanoid Robots (Humanoids)*, pages 428–435, 2016. **2**
- [6] Raphaël Candelier, Alexis Prevost, and Georges Debrégeas. The role of exploratory conditions in bio-inspired tactile sensing of single topological features. *Sensors*, 11(8):7934–7953, 2011. **2**
- [7] Nicolas R. Chevalier. Hair-on-hair static friction coefficient can be determined by tying a knot. *Colloids and Surfaces B: Biointerfaces*, 159:924–928, 2017. **1, 2**
- [8] François Chollet et al. Keras. <https://keras.io>, 2015. **7**
- [9] David Clyde, Joseph Teran, and Rasmus Tamstorf. Modeling and data-driven parameter estimation for woven fabrics. In *Proceedings of the ACM SIGGRAPH/Eurographics Symposium on Computer Animation*, page 17. ACM, 2017. **3**
- [10] R. Courtel and L. Tichvinsky. A brief history of friction. *Naval Engineers Journal*, 76(3):451–460, 1964. **3, 4**
- [11] Abe Davis, Katherine L Bouman, Justin G Chen, Michael Rubinstein, Oral Buyukozturk, Fredo Durand, and William T Freeman. Visual vibrometry: Estimating material properties from small motions in video. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(4):732–745, 2017. **1, 3**
- [12] Jeffrey Donahue, Lisa Anne Hendricks, Sergio Guadarrama, Marcus Rohrbach, Subhashini Venugopalan, Kate Saenko, and Trevor Darrell. Long-term recurrent convolutional networks for visual recognition and description. In *Proceedings of the IEEE conference on Computer Vision and Pattern Recognition*, pages 2625–2634, 2015. **2, 6**
- [13] E.C. Dreby. A friction meter for determining the coefficient of kinetic friction of fabrics. *Journal of Research of the National Bureau of Standards*, 31(4):237, 1943. **2, 5**
- [14] Muhammad Waleed Gondal, Manuel Wüthrich, Dorde Miladinovic, Francesco Locatello, Martin Breidt, Valentin Volchkov, Joel Akpo, Olivier Bachem, Bernhard Schölkopf, and Stefan Bauer. On the transfer of inductive bias from simulation to the real world: a new disentanglement dataset. Technical report, arXiv, 2019. **2**
- [15] Stephen James, Paul Wohlhart, Mrinal Kalakrishnan, Dmitry Kalashnikov, Alex Irpan, Julian Ibarz, Sergey Levine, Raia Hadsell, and Konstantinos Bousmalis. Sim-to-real via sim-to-sim: Data-efficient robotic grasping via randomized-to-canonical adaptation networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 12627–12637, 2019. **2**
- [16] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. In *Advances in Neural Information Processing Systems*, pages 1097–1105, 2012. **6**
- [17] Shoji Kunitomo, Shinsuke Nakamura, and Shigeo Morishima. Optimization of cloth simulation parameters by considering static and dynamic features. In *ACM SIGGRAPH 2010 Posters*, SIGGRAPH '10, pages 15:1–15:1, New York, NY, USA, 2010. ACM. **3**
- [18] Jie Li, Gilles Daviet, Rahul Narain, Florence Bertails-Descoubes, Matthew Overby, George E Brown, and Laurence Boissieux. An implicit frictional contact solver for adaptive cloth simulation. *ACM Transactions on Graphics (TOG)*, 37(4):52, 2018. **2, 4, 5**
- [19] Junbang Liang, Ming C. Lin, and Vladlen Koltun. Differentiable cloth simulation for inverse problems. In *Conference on Neural Information Processing Systems*, 2019. **3**
- [20] Xianping Liu, Z Yue, Zhiping Cai, D G. Chetwynd, and Stuart Smith. Quantifying touch-feel perception: Tribological aspects. *Measurement Science & Technology - MEAS SCI TECHNOL*, 19, 2008. **2**
- [21] Ningtao Mao, Yiyi Wang, and Jianguo Qu. Smoothness and roughness: Characteristics of fabric-to-fabric self-friction properties. In *The Proceedings of 90th Textile Institute World Conference*. The Textile Institute, 2016. **2**
- [22] Eder Miguel, Derek Bradley, Bernhard Thomaszewski, Bernd Bickel, Wojciech Matusik, Miguel A Otaduy, and Steve Marschner. Data-driven estimation of cloth simulation models. In *Computer Graphics Forum*, volume 31, pages 519–528. Wiley Online Library, 2012. **3**
- [23] Eder Miguel, Rasmus Tamstorf, Derek Bradley, Sara C. Schvartzman, Bernhard Thomaszewski, Bernd Bickel, Wojciech Matusik, Steve Marschner, and Miguel A. Otaduy. Modeling and estimation of internal friction in cloth. *Transactions on Graphics*, 32(6):212:1–212:10, 2013. **2**
- [24] R Rathina Moorthy and P Kandhavadi. Surface friction characteristics of woven fabrics with nonconventional fibers and their blends. *Journal of Textile and Apparel, Technology and Management*, 9(3), 2015. **1, 2**

- [25] Rahul Narain, Armin Samii, and James F. O'Brien. Adaptive anisotropic remeshing for cloth simulation. *ACM Trans. Graph.*, 31(6):152:1–152:10, Nov. 2012. [3](#)
- [26] Tomohiko G. Sano, Tetsuo Yamaguchi, and Hirofumi Wada. Slip morphology of elastic strips on frictional rigid substrates. *Physical Review Letters*, 118, 2017. [1](#), [2](#), [4](#), [5](#)
- [27] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014. [6](#)
- [28] Cédric Syllebranque and Samuel Boivin. Estimation of mechanical parameters of deformable solids from videos. *The Visual Computer*, 24(11):963–972, 2008. [3](#)
- [29] GH Thorndike and L Varley. Measurement of the coefficient of friction between samples of the same cloth. *Journal of the Textile Institute Proceedings*, 52(6):P255–P271, 1961. [1](#), [2](#), [3](#), [4](#)
- [30] Josh Tobin, Rachel Fong, Alex Ray, Jonas Schneider, Wojciech Zaremba, and Pieter Abbeel. Domain randomization for transferring deep neural networks from simulation to the real world. In *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 23–30. IEEE, 2017. [2](#)
- [31] Bin Wang, Longhua Wu, KangKang Yin, Uri M Ascher, Libin Liu, and Hui Huang. Deformation capture and modeling of soft objects. *ACM Transactions on Graphics (TOG)*, 34(4):94–1, 2015. [3](#)
- [32] Huamin Wang, James F O'Brien, and Ravi Ramamoorthi. Data-driven elastic models for cloth: modeling and measurement. In *ACM Transactions on Graphics (TOG)*, volume 30, page 71. ACM, 2011. [3](#), [5](#)
- [33] Jiajun Wu, Joseph J Lim, Hongyi Zhang, Joshua B Tenenbaum, and William T Freeman. Physics 101: Learning physical object properties from unlabeled videos. In *British Machine Vision Conference*, 2016. [1](#)
- [34] Jiajun Wu, Ilker Yildirim, Joseph J Lim, Bill Freeman, and Josh Tenenbaum. Galileo: Perceiving physical object properties by integrating a physics engine with deep learning. In *Advances in Neural Information Processing Systems*, pages 127–135, 2015. [2](#)
- [35] Shan Yang, Junbang Liang, and Ming C Lin. Learning-based cloth material recovery from video. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 4383–4393, 2017. [1](#), [3](#), [6](#)
- [36] Shan Yang and Ming C Lin. Materialcloning: Acquiring elasticity parameters from images for medical applications. *IEEE Transactions on Visualization and Computer Graphics*, 22(9):2122–2135, 2016. [3](#)
- [37] Wenzhen Yuan, Shaoxiong Wang, Siyuan Dong, and Edward Adelson. Connecting look and feel: Associating the visual and tactile properties of physical materials. In *Conference on Computer Vision and Pattern Recognition*, 2017. [2](#)
- [38] Matthew D Zeiler. Adadelta: an adaptive learning rate method. *arXiv preprint arXiv:1212.5701*, 2012. [7](#)
- [39] Hang Zhang, Kristin Dana, and Ko Nishino. Friction from reflectance: Deep reflectance codes for predicting physical surface properties from one-shot in-field reflectance. In *European Conference on Computer Vision*, 2016. [2](#)