0P

Good evening, everyone. I'm Ryuichi Ikeya from Institute of Science Tokyo.

Today, I'd like to introduce a new method for estimating force from deformation.

1P

First, let's start with the technical background.

Recently, it has become possible to easily extract the shape of an object.

In terms of software development, classification of object at pixel level is now possible.

For example, Meta released their latest segmentation model, which makes it easy to extract shape feature from images.

In terms of hardware development, capturing 3D images using inexpensive RGBD cameras is now possible.

(With advances in software and hardware, it is now possible to obtain the 3D data of an object by extracting the object from a depth image.)

These technical advancements open a new way of estimating force just by observing deformations.

2P

In other words, conventionally, force was measured by sensors embedding like this.

3P

In the future, force was measured by deformation observation like this.

4P

Vision sensors that estimate forces from images are attracting attention, and there is some related research.

In particular, by using machine learning based on experimental data, models can be easily created and inference can be performed in real time.

However, existing research simply outputs a scalar value from plain RGB images, so it has some limitation.

First, it is not flexible to changes in the experimental environment.

For example, changes in background, changes in lighting, presence of humans.

Second, just single-point force estimation limits application possibilities. For example, ...

5P

So, in this study, we propose a new force sensor principle that estimates force distribution just by observing deformation and verifying its feasibility.

In the proposed method, we first capture video, then extract shape deformation, and finally output image format force using a convolutional neural network.

The proposed method has two novelties.

First, by extracting deformation feature, it is flexible to environmental changes.

Secondly, by expressing force in image format, multi-point estimation rather than single-point estimation become possible.

6P

This slide shows the details of the proposed method.

In the inference process, deformation feature is extracted from two images, one before and one after deformation, and force distribution is inferred using a trained machine learning model (CNN)

7P

In the training process, deformation feature is extracted from the depth and RGB image by image processing, and the force distribution can be expressed in image format by embedding the sensor values.

By learning the relationship between this deformation feature and force distribution, we can create a model which infer force distribution from an RGBD camera.

8P

Next, we explain how to express deformation feature.

The deformation feature is composed of three elements as images: the shape before deformation, the shape after deformation, and the contact position.

In addition, we prepared contour and depth feature for the shape before deformation and after deformation.

The contour feature is a binary image, and the depth feature is a grayscale image normalized to 0 to 1.

9P

Machine learning model architecture is the typical encoder-decoder model which is used for tasks with image inputs and output.

As a unique innovation, we introduced a mechanism to limits the model output to a reasonable range depending on the contact position.

10P

Next, I will explain how to create a dataset.

This consists of two steps: extracting deformation feature as input data, and creating force distribution images as correct answer data.

First, I will explain the process of extracting deformation feature using image processing.

11P

This slide shows an overview of extraction of deformation feature by image processing.

First, we extract the binary image of the contour by clustering the RGB image.

Next, we extract the depth image of object by multiplying the binary image of contour.

In this process, by extracting the deformation feature of the object, the model becomes robust to background changes.

Finally, for contact position, we use a skeleton estimation method called openpose, and obtain a binary image of contact position.

12P

Next, we will explain how to create teacher data using force sensor.

13P

To create the teacher data, first, we use openpose to identify 11 representative finger skeleton coordinates.

Next, the force sensor values are assigned to the corresponding pixels.

Finally, a force distribution image like this can be obtained for each measurement time.

14P

Until now, force distribution has been expressed in two-dimensional image format, but two-dimensional images have poor visibility for humans to understand.

Therefore, in this research, we developed a system that displays as a vector on a three-dimensional point.

This is achieved by assuming that the force is applied perpendicularly to the object at each position of the force distribution.

Using this system, experimental results are displayed as a 3D point rather than as an image.

15P

This is the numerical results.

As you can see from the table, the mean relative error of the test data was 18% for the balance ball. Also, the mean relative error of the test data was 20% for the cushion. These errors are a bit higher than ultra-high-precision sensors such as load cells. However, they are within the acceptable range for a pressure sensor.

16P

This is an example of with average error for balance ball. The left figure is the model's output, and the right figure is the ground truth.

The colors of the vectors are in the same range for the left and right diagrams, so the same color means the same magnitude of force.

From this figure, although there are small differences, the points where the force is strongest and trends in force distribution are well understood.

17P

Here is an example of an output with a relatively large error.

Errors come from mistaking the points where the force is strongest, or come from estimating forces that do not exist.

18P

This is an example of with average error for cushion.

As with the balance ball, the points where the force is strongest and trends in force distribution are well understood.

19P

Here is an example of an output with a relatively large error.

As with the balance ball, errors come from mistaking the points where the force is strongest, or come from estimating forces that do not exist.

20P

Finally, we discuss some error factors.

The first is image processing noise. As shown in this figure, there is noise around the contours of shapes.

The second is information loss due to the nature of the depth camera. As shown in this figure, areas with a large incidence angle have some defect.

The third is occlusion issues. Occlusion issues occur behind hands and objects.

The fourth is insufficient accuracy of contact judgment. As shown in this figure, there are cases where a non-touched position is recognized as touched.

21P

In conclusion, we proposed a new force sensor based on deformation feature.

For validation, we conducted experiments with an integrated system for measuring deformation feature and estimating force distribution.

As for accuracy, for the balance ball, the system was able to infer the position with an error of about 18%, For cushions, the system was able to infer with an error of about 20%.

Finally, we discuss some error factors.

22P

This research showed the possibility of sensorization by deformation observation.

This type of new force sensor is expected to enable a variety of applications in living environments and sensing robots that were difficult with conventional sensors.