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, it is becoming easier to classify objects at the pixel level.

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

In terms of hardware development, it is now possible to easily capture 3D images using inexpensive RGBD cameras.

(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 only by observing deformations

2P

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

3P

In the future, we will be able to measure forces by observing deformation like this.

(そしてこの、変形観察による力計測は、生活環境のモニタリンによる人間のセンシングや埋め込みの必要がないロボットの触覚センサとしての利用など、従来の力セン

サでは実現が難しい様々な応用が期待されます。)

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 has some limitations because it outputs a scalar from RGB images.

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

Therefore, in this study, we propose a new force sensor principle that estimates force distribution from observable shape deformation information 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 novelty.

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

Secondly, expressing force in image format allows for multi-point estimation rather than single-point estimation.

6P

This slide shows the details of the proposed method.

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

7P

In the training process, shape deformation information is extracted from the depth image and RGB image by image processing, and the force distribution is processed in an image format to embed the sensor values into an RGB image.

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

8P

Next, we explain how to represent shape deformation information.

The shape transformation information is composed of three elements in the form of images: the shape before deformation, the shape after deformation, and the contact position.

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

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

9P

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

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