Image-Based Estimation of Real Food Size for Accurate Food Calorie Estimation

Takumi Ege, Yoshikazu Ando, Ryosuke Tanno, Wataru Shimoda and Keiji Yanai Department of Informatics, The University of Electro-Communications, Tokyo

ⓒ 2017 UEC Tokyo.

Introduction

- ***Food calorie estimation***
- Existing works
  - Associated with the estimated food categories
  - Require users to enter information such as size or volume
- Our works
  - Fully-automatic food calorie estimation

ⓒ 2017 UEC Tokyo.

Review of our works

- (1) “CalorieCam”
  - reference-object-based food calorie estimation system.
- (2) Weakly-supervised segmentation based food calorie estimation
- (3) “AR DeepCalorieCam V2”
  - A real food size and calorie estimation system based on iOS ARKit.
- (4) “DepthCalorieCam” (new system)
  - food calorie estimation system exploiting iPhone stereo cameras.
- (5) Rice grain based size estimation (new method)
  - uses rice grained the size of which are usually almost the same as a reference object.

ⓒ 2017 UEC Tokyo.

CalorieCam

- Image-based calorie estimation system
- User needs to register a size-known reference object
  - Wallet
  - Creditcard-size card

ⓒ 2017 UEC

CalorieCam

- Segmentation
  - Food items
  - Pre-registered reference object.
- Real size of each detected food items
  - Comparing the number of pixels
- Process
  - Estimate rough position based on edge detection
  - Apply color-pixel-based k-means clustering
  - Apply GrabCut with the detected bounding box

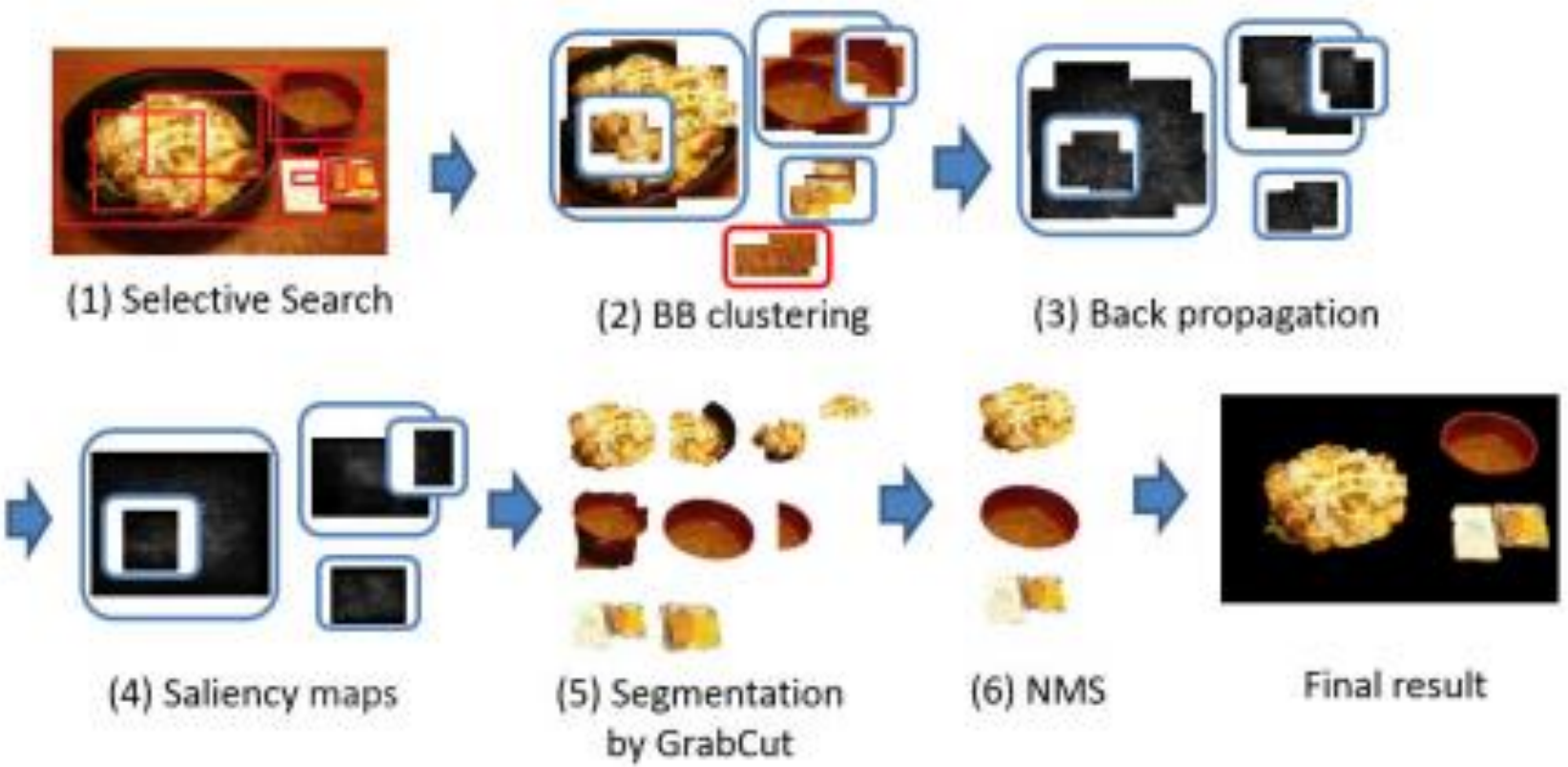
ⓒ 2017 UEC Tokyo.

Weakly-supervised Segmentation Based Calorie Estimation

- Calorie Cam
  - One meal photo contains only one dish
- This work
  - Estimate calories from segmentation results.
  - Without multiple-view photos
  - Without specific reference objects such as wallets and cards.

ⓒ 2017 UEC Tokyo.

Weakly-supervised Segmentation Based Calorie Estimation

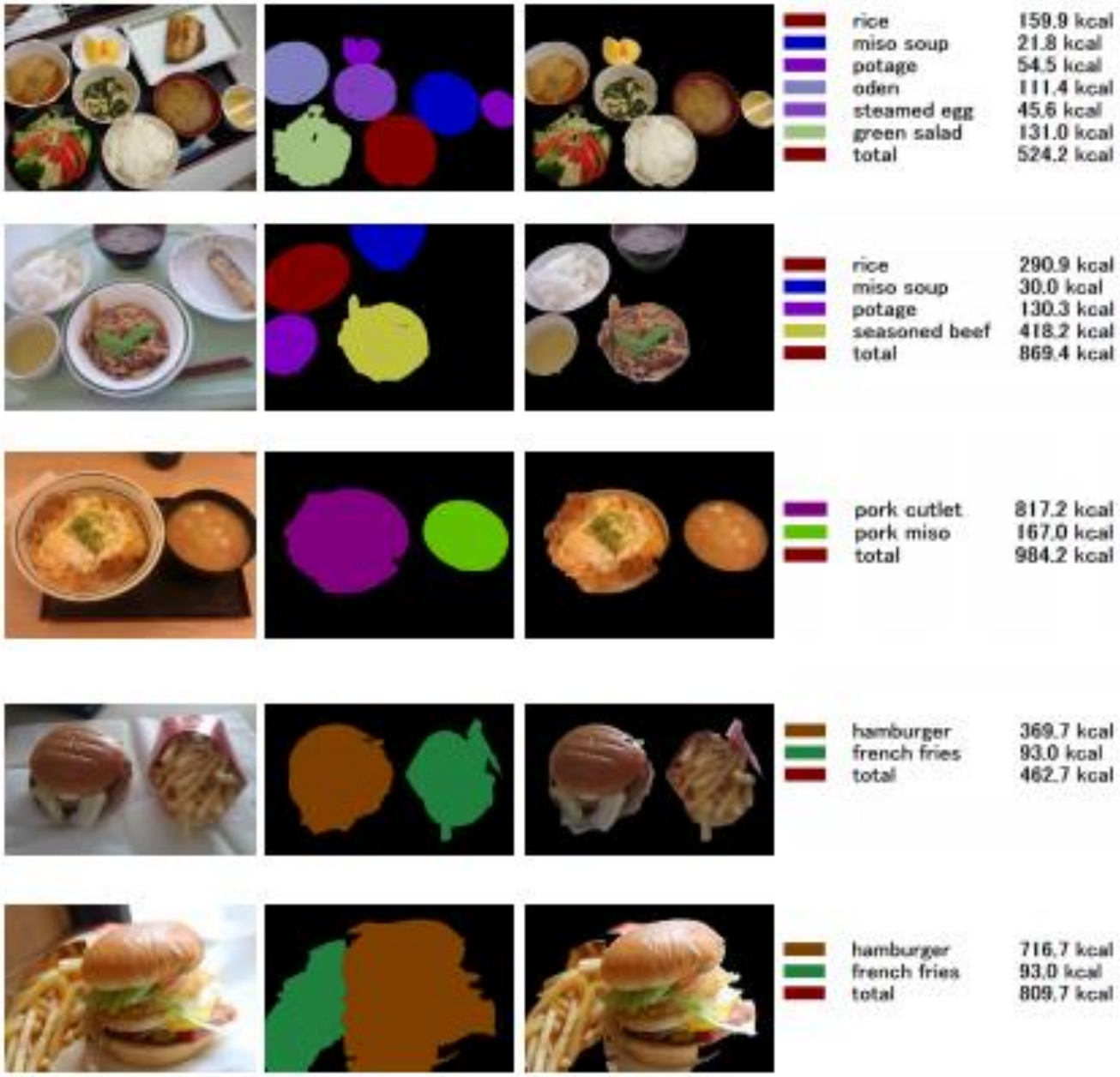


ⓒ 2017 UEC Tokyo.

Weakly-supervised Segmentation Based Calorie Estimation

- Choose base food region
- We decide priorities based on a tendency of unchanging food volumes.
- Some food volumes change frequently, while some foods volume rarely change.
- In “Teishoku”
  - Japanese traditional food combo menu
  - we can often change the volume of “rice”
  - we cannot change “miso-soup” volume in general.

ⓒ 2017 UEC Tokyo.



© 2017 UEC Tokyo.

AR DeepCalorieCam

- Uses an inertial sensor built in a standard smartphone in addition to photos

- Recognize a category of each food item.
- Directly calculate the size using AR
- Calculate food calories based on their

ⓒ 2017 UEC Tokyo.

AR DeepCalorieCam

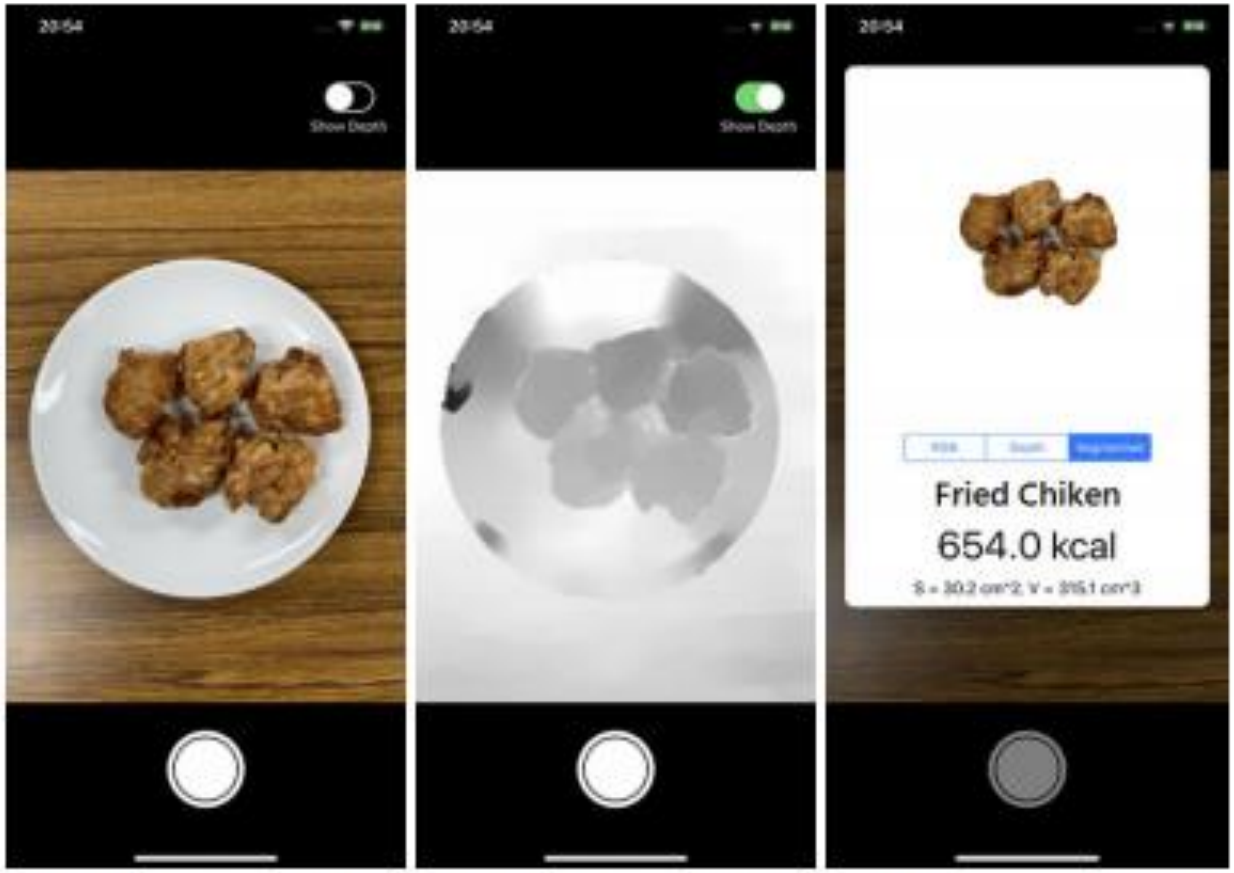
AR DeepCalorieCam

with iOS ARKit

ARKit estimates real size of the bounding box.

ⓒ 2019 UEC Tokyo.

DepthCalorieCam



An obtained RGB image.

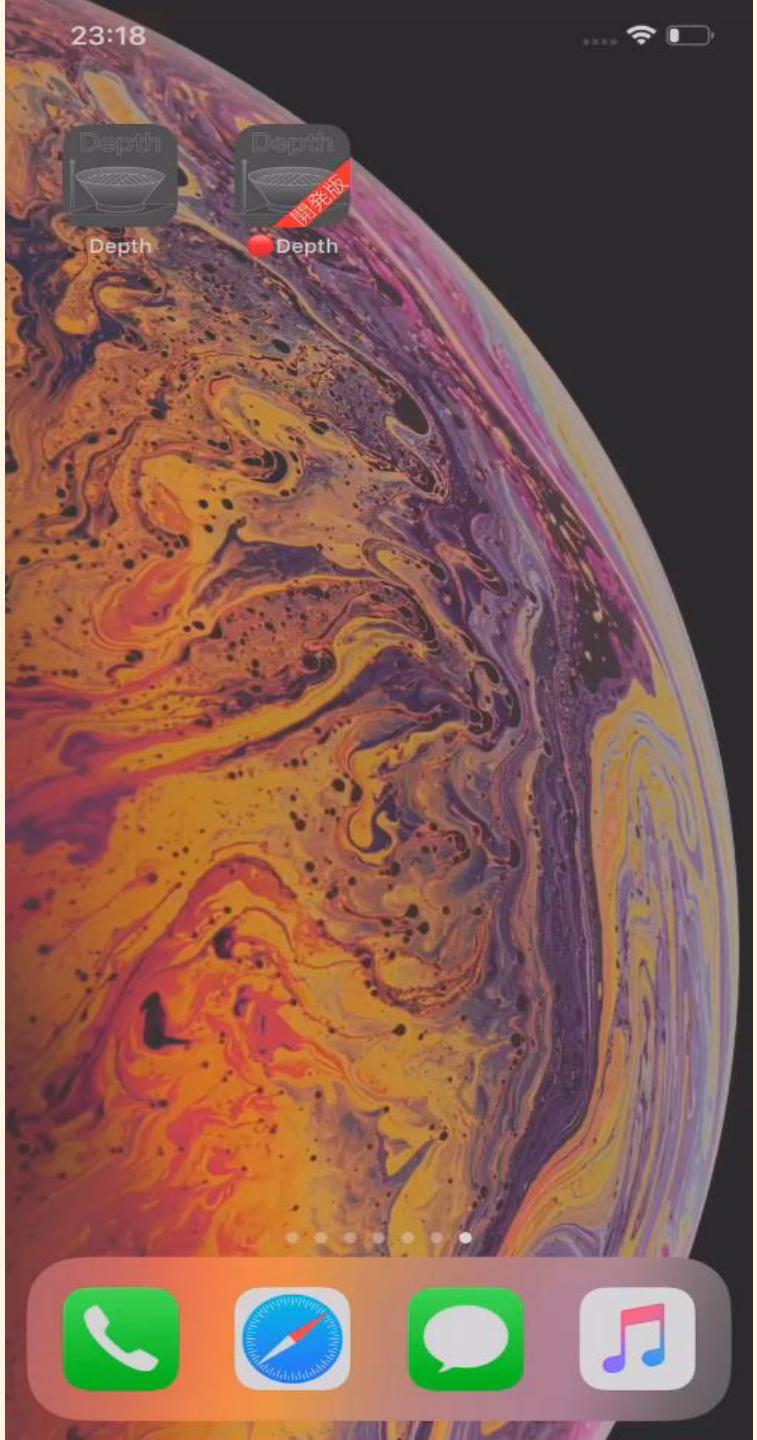
An obtained depth image.

Estimated result with calorie intake and segmented food region.

Croquette calorie estimation demo:

1 croquette
=160kcal

2 croquette
=320kcal



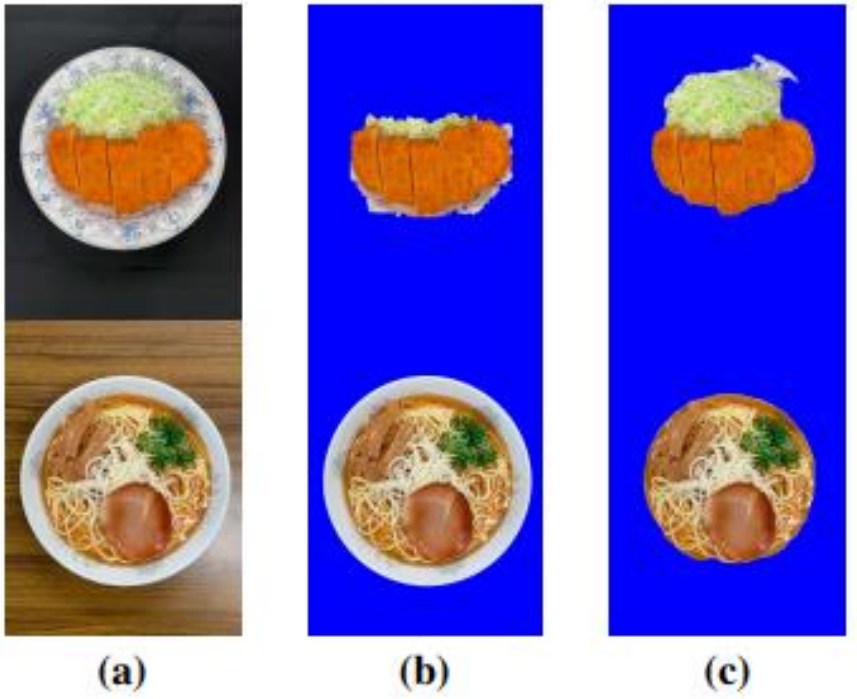
DepthCalorieCam

- Food regions are extracted by U-NET trained with UECFood-100 and segmentation mask sets.
- Estimate the depth of each of the pixels and their actual volumes.
- Calculate food calories from the actual volume and the regression curves

ⓒ 2017 UEC Tokyo.

DepthCalorieCam

- CNN-based food region segmentation
  - Network
    - U-Net
  - Training
    - 5,301 masks
    - UECFOOD-100



ⓒ 2017 UEC Tokyo.

Comparative Experiments with CalorieCam and AR DeepCalorieCam V2

- Target

Pork with sweet and sour source.

Fried chicken.

Croquette.

category	calorie [kcal]
pork with source	500
fried chicken	655
croquette	246

Comparative Experiments with CalorieCam and AR DeepCalorieCam V2

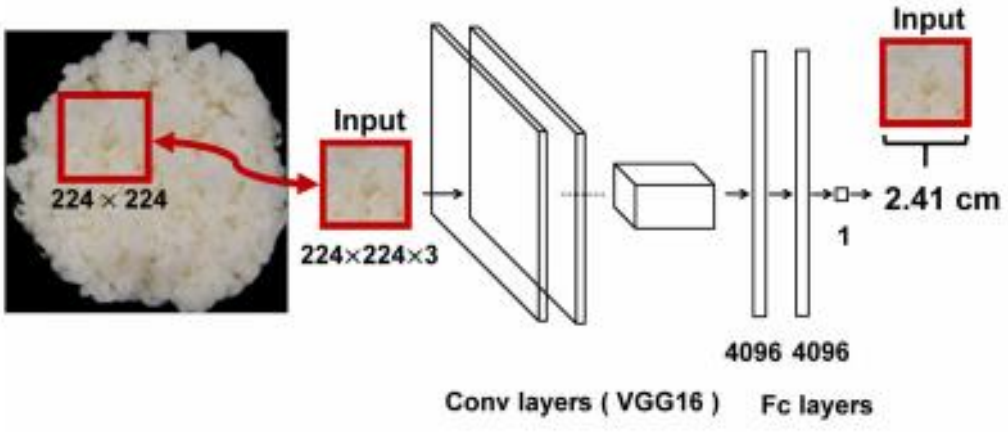
Table 2. Comparison on calorie estimation error (Avg.±SD [kcal]) among CalorieCam [1], AR CalorieCam V2 [3] and DepthCalorieCam.

category	CalorieCam	AR CalorieCam	DepthCalorieCam
pork with source	364±552	-112±163	2±52
fried chicken	-123±171	343±51	-5±64
croquette	-48±16	-104±12	-35±22

ⓒ 2017 UEC Tokyo.

Rice grain based size estimation

- Proposed network



- Examples of dataset

ⓒ 2017 UEC

Dataset

- 360 images

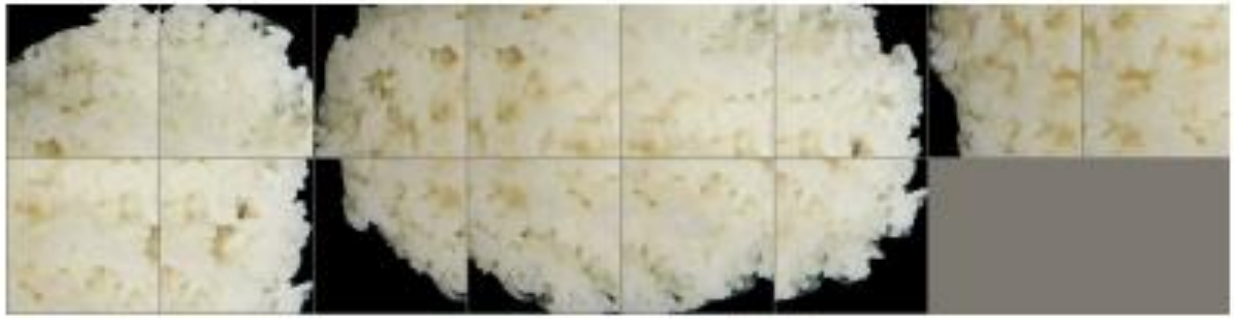
Amount of water / Camera	Small amount 180ml/150g	Medium amount 200ml/150g	Large amount 220ml/150g
COOLPIX AW120 3264x2448			
iPhone8 Plus 4032x3024			

ⓒ 2017 UEC Tokyo.

Experiments

- Training

- Evaluation



ⓒ 2017 UEC Tokyo.

Experiments

Evaluation data.	abs. err.(cm/224pixels)
Camera:COOLPIX, Small amount of water	0.212
Camera:COOLPIX, Medium amount of water	0.178
Camera:COOLPIX, Large amount of water	0.197
Camera:iPhone8 Plus, Small amount of water	0.127
Camera:iPhone8 Plus, Medium amount of water	0.170
Camera:iPnone8 Plus, Large amount of water	0.105

rel. err.(%)	corr.	$\leq$ 5% err.(%)	$\leq$ 10% err.(%)	$\leq$ 20% err.(%)
7.182	0.958	41.667	75.000	91.667
6.550	0.973	43.333	76.667	93.333
6.668	0.962	48.333	78.333	90.000
5.652	0.945	50.000	75.000	98.333
7.512	0.903	43.333	68.333	88.333
4.800	0.967	58.333	88.333	98.333

ⓒ 2017 UEC Tokyo.

Conclusion

- DepthCalorieCam is the most promising approach.
- Large-scale calorieannotated 3D food volume data is needed to extend the system
  - very costly and time-consuming.
- The rice grain based method is also promising
  - Appropriate for Japanese foods

ⓒ 2017 UEC Tokyo.