Recognition of food-texture attributes using an in-ear microphone

Vasileios Papapanagiotou[1], Christos Diou[1,2], Janet van den Boer[3], Monica Mars[4], Anastasios Delopoulos[1]

[1] Multimedia Understanding Group, Dpt. Electrical and Computer Engineering, Faculty of Engineering, Aristotle University of Thessaloniki, Greece
vassilis@mug.ee.auth.gr, adelo@eng.auth.gr

[2] Department of Informatics and Telematics, Harokopio University of Athens
cdiou@hua.gr

[3] Dpt. of Biomedical Signals and Systems, Faculty of Electrical Engineering, Mathematics and Computer Science, University of Twente, The Netherlands
j.h.w.vandenboer@utwente.nl

[4] Division of Human Nutrition and Health, Wageningen University
monica.mars@wur.nl

25th International Conference on Pattern Recognition
6th International Workshop on Multimedia Assisted Dietary Management

- Chewing is one of the main ways of how we perceive food texture
- Some textures are generally perceived as more pleasant and desirable than others [1]
- Several studies show food texture and structure are becoming more important in understanding eating behavior, especially in food intake regulation and weight management [2]

[1] Aguayo-Mendoza, M.G., Ketel, E.C., van der Linden, E., Forde, C.G., Piqueras-Fiszman, B., Stieger, M.: Oral processing behavior of drinkable, spoonable and chewable foods is primarily determined by rheological and mechanical food properties. Food Quality and Preference 71, 87 – 95 (2019). https://doi.org/https://doi.org/10.1016/j.foodqual.2018.06.006

[2] StribiŢcaia, E., Evans, C.E.L., Gibbons, C., Blundell, J., Sarkar, A.: Food texture influences on satiety: systematic review and meta-analysis. Scientific Reports 10(1), 12929 (7 2020). https://doi.org/10.1038/s41598-020-69504-y

Introduction and motivation

- Chewing is one of the main ways of how we perceive food texture
- Some textures are generally perceived as more pleasant and desirable than others [1]
- Several studies show food texture and structure are becoming more important in understanding eating behavior, especially in food intake regulation and weight management [2]

As a result, automatically recognizing food texture can help

- understand human preference on food selection and help with weight management, as well as improve the nutrition content of diets
- understand consumer preference on food products and design more desirable products

[1] Aguayo-Mendoza, M.G., Ketel, E.C., van der Linden, E., Forde, C.G., Piqueras-Fiszman, B., Stieger, M.: Oral processing behavior of drinkable, spoonable and chewable foods is primarily determined by rheological and mechanical food properties. Food Quality and Preference 71, 87 – 95 (2019). https://doi.org/https://doi.org/10.1016/j.foodqual.2018.06.006

[2] StribiŢcaia, E., Evans, C.E.L., Gibbons, C., Blundell, J., Sarkar, A.: Food texture influences on satiety: systematic review and meta-analysis. Scientific Reports 10(1), 12929 (7 2020). https://doi.org/10.1038/s41598-020-69504-y

Food texture definition

Food-texture attributes as presented and organized in [3], and their correspondence with the food-attributes used in this work.

Attributes	**Crisp**	**Wet**	**Chewy**
Attributes related to surface attributes and springiness			
wetness	✗	✓	✗
adhesiveness to lips	✗	✗	✗
roughness	✓	✗	✗
self-adhesiveness	✗	✗	✓
springiness	✓	✗	✗
Attributes assessed during mastication			
cohesiveness of mass	✓	✗	✗
moisture absorption	✗	✓	✗
adhesiveness to teeth	✗	✗	✓
Attributes assessed during manual manipulation			
manual adhesiveness	✗	✗	✗

[3] Muñoz, A.M.: Development and application of texture reference scales. Journal of Sensory Studies 1(1), 55–83 (1986). https://doi.org/10.1111/j.1745459X.1986.tb00159.x

Food texture definition

Food-texture attributes as presented and organized in [3], and their correspondence with the food-attributes used in this work.

Attributes	Crisp	Wet	Chewy
Attributes related to surface attributes and springiness			
wetness	✗	✓	✗
adhesiveness to lips	✗	✗	✗
roughness	✓	✗	✗
self-adhesiveness	✗	✗	✓
springiness	✓	✗	✗
Attributes assessed during mastication			
cohesiveness of mass	✓	✗	✗
moisture absorption	✗	✓	✗
adhesiveness to teeth	✗	✗	✓
Attributes assessed during manual manipulation			
manual adhesiveness	✗	✗	✗

[3] Muñoz, A.M.: Development and application of texture reference scales. Journal of Sensory Studies 1(1), 55–83 (1986). https://doi.org/10.1111/j.1745459X.1986.tb00159.x

Food texture definition

Food-texture attributes as presented and organized in [3], and their correspondence with the food-attributes used in this work.

Attributes	Crisp	Wet	Chewy
Attributes assessed during mastication			
cohesiveness of mass	✓	✗	✗
moisture absorption	✗	✓	✗
adhesiveness to teeth	✗	✗	✓

Multi-label formulation of the food-texture attribute recognition problem.

Label value	Crispiness label	Wetness label	Chewiness label
1	crispy	wet	chewy
0	non-crispy	dry	non-chewy

[3] Muñoz, A.M.: Development and application of texture reference scales. Journal of Sensory Studies 1(1), 55–83 (1986). https://doi.org/10.1111/j.1745459X.1986.tb00159.x

- In-ear microphone (Knowles FG-23329-D65), captures at 48 kHz
- A PPG sensor (New Balance NB439B), not used in this work

Figure: The ear-worn device with the microphone and PPG sensor.

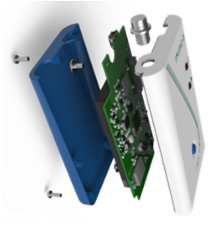


Figure: The belt-mounted device with 3D accelerometer and data-logger.

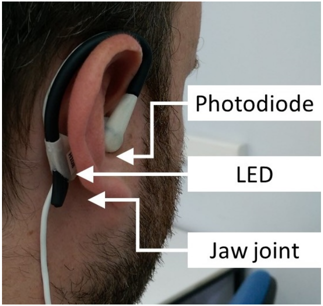


Figure: PPG placement in ear.

Created within the context of the EU funded SPLENDID project.
`https://splendid-program.eu/`

Chew-level recognition algorithm

Algorithm outline

Chew-level recognition algorithm

Algorithm outline

1. Pre-processing
   1. Down-sampling (tested 2, 4, 8, 16, and 32 kHz, selected 8 kHz)
   2. High-pass filtering at 20 Hz

Chew-level recognition algorithm

Algorithm outline

1. Pre-processing
   1. Down-sampling (tested 2, 4, 8, 16, and 32 kHz, selected 8 kHz)
   2. High-pass filtering at 20 Hz
2. Feature extraction
   1. Features per chew audio segment, as in Table (0.56 ± 0.15 s duration)
   2. Feature standardization: $\boldsymbol{f}_{\text{norm}}[i] = \left(\boldsymbol{f}[i] - \mu_i\right)/\sigma_i$

Table: Audio features.

	Feature	Dimension	Window
1	Energy of log sectral band	9	0.2 s
2	Fractal dimension	1	0.1 s
3	Condition number	1	0.1 s
4	Skewness $m_3((0,0))$	1	0.1 s
5	Kurtosis $m_4((0,0,0))$	1	0.1 s
6	Moment $m_4((0,1,1))$	1	0.1 s
7	Moment $m_4((0,2,2))$	1	0.1 s

Algorithm outline

1. Pre-processing
   1. Down-sampling (tested 2, 4, 8, 16, and 32 kHz, selected 8 kHz)
   2. High-pass filtering at 20 Hz
2. Feature extraction
   1. Features per chew audio segment, as in Table (0.56 ± 0.15 s duration)
   2. Feature standardization: $\boldsymbol{f}_{\text{norm}}[i] = \left(\boldsymbol{f}[i] - \mu_i\right)/\sigma_i$
3. Classification (per chew)
   1. Binary SVMs, one SVM per food attribute: crispiness, wetness, and chewiness
   2. RBF kernel
   3. Parameters C and γ are selected automatically using Bayesian optimization in a 5-fold cross-validation

Table: Audio features.

	Feature	**Dimension**	**Window**
1	Energy of log sectral band	9	0.2 s
2	Fractal dimension	1	0.1 s
3	Condition number	1	0.1 s
4	Skewness $m_3((0,0))$	1	0.1 s
5	Kurtosis $m_4((0,0,0))$	1	0.1 s
6	Moment $m_4((0,1,1))$	1	0.1 s
7	Moment $m_4((0,2,2))$	1	0.1 s

Algorithm outline

Bout-level recognition algorithm

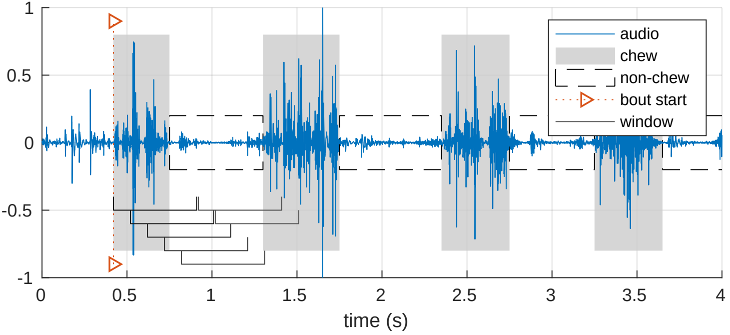
Algorithm outline

1. Pre-processing
2. Feature extraction
   1. Same chew-based features
      - ground truth bouts: 15.22 ± 10.7 s duration
      - windows: 0.5 s size, 0.1 s step
   2. Feature standardization
3. Bag-of-words
   1. On the chew-based features
   2. Normalized histogram across the chewing bout as the final feature vector
4. Classification (per bout)

Bout-level recognition algorithm

Algorithm outline

1. Pre-processing
2. Feature extraction
   1. Same chew-based features
      - ground truth bouts: 15.22 ± 10.7 s duration
      - windows: 0.5 s size, 0.1 s step
   2. Feature standardization
3. Bag-of-words
   1. On the chew-based features
   2. Normalized histogram across the chewing bout as the final feature vector
4. Classification (per bout)



Evaluation methodology

Evaluation methodology

Evaluation levels

- Chew-level evaluation
  - for chew-level recognition algorithm
- Bout-level evaluation
  - chew-level recognition algorithm with majority voting
  - chew-level recognition algorithm with majority voting over the n first chews (of each bout)
  - bout-level recognition algorithm

Evaluation methodology

Evaluation levels

- Chew-level evaluation
  - for chew-level recognition algorithm
- Bout-level evaluation
  - chew-level recognition algorithm with majority voting
  - chew-level recognition algorithm with majority voting over the n first chews (of each bout)
  - bout-level recognition algorithm

Evaluation methods

- LOSO: leave one subject out
- LOFTO: leave one food-type out

Evaluation methodology

Evaluation levels

- Chew-level evaluation
  - for chew-level recognition algorithm
- Bout-level evaluation
  - chew-level recognition algorithm with majority voting
  - chew-level recognition algorithm with majority voting over the n first chews (of each bout)
  - bout-level recognition algorithm

Evaluation methods

- LOSO: leave one subject out
- LOFTO: leave one food-type out

Evaluation metrics:

- Accuracy (per food attribute)
- Weight accuracy (per food attribute)

Dataset

- Collected at Wageningen University, Netherlands, in the context of the EU-funded SPLENDID project
- Recording apparatus: in-ear microphone (Knowles FG-23329-D65) connected via wire to a computer audio interface
- Sensor housing and recording by CSEM S.A.

Dataset

- Collected at Wageningen University, Netherlands, in the context of the EU-funded SPLENDID project
- Recording apparatus: in-ear microphone (Knowles FG-23329-D65) connected via wire to a computer audio interface
- Sensor housing and recording by CSEM S.A.
- In total, 21 subjects were enrolled for the data collection trials, however, signals from only 9 could be used in this work due to problems with data acquisition (such as incorrect sensor placement or corrupted audio due to hardware/software malfunction)
- Each subject consumed a variety of food types

Dataset

- Collected at Wageningen University, Netherlands, in the context of the EU-funded SPLENDID project
- Recording apparatus: in-ear microphone (Knowles FG-23329-D65) connected via wire to a computer audio interface
- Sensor housing and recording by CSEM S.A.
- In total, 21 subjects were enrolled for the data collection trials, however, signals from only 9 could be used in this work due to problems with data acquisition (such as incorrect sensor placement or corrupted audio due to hardware/software malfunction)
- Each subject consumed a variety of food types
- Ground truth was manually created based on visual inspection of the audio signals and experimental logs

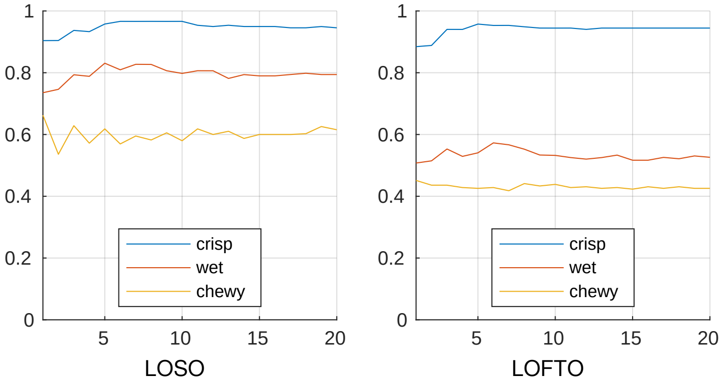
Table: LOSO results for chew-level recognition.

	Prior	**Weighted accuracy**
Chew level		
crispy (avg)	0.4707	0.9068
crispy (sum)	0.4666	0.9017
wet (avg)	0.4280	0.7516
wet (sum)	0.4235	0.7503
chewy (avg)	0.1741	0.5994
chewy (sum)	0.1746	0.6212
Majority voting per bout		
crispy (avg)	0.4943	0.9519
crispy (sum)	0.5063	0.9496
wet (avg)	0.4850	0.7978
wet (sum)	0.4937	0.7900
chewy (avg)	0.1666	0.6296
chewy (sum)	0.1632	0.6154

Results

Table: LOFTO results for chew-level recognition.

	Prior	**Weighted accuracy**
Chew-level		
crispy (sum)	0.4666	0.8987
wet (sum)	0.4235	0.5481
chewy (sum)	0.1746	0.3957
Majority voting per bout		
crispy (sum)	0.4957	0.9446
wet (sum)	0.4829	0.5046
chewy (sum)	0.1667	0.4179



Weighted accuracy for each attribute for the LOSO and the LOFTO experiments.

Table: Results for bout-level recognition.

	Prior	Weighted accuracy
LOSO		
crispy (avg)	0.4967	0.9541
crispy (sum)	0.5084	0.9534
wet (avg)	0.4869	0.7865
wet (sum)	0.4958	0.7900
chewy (avg)	0.1625	0.5200
chewy (sum)	0.1597	0.5238
LOFTO		
crispy (sum)	0.5084	0.9288
wet (sum)	0.4958	0.6422
chewy (sum)	0.1597	0.4970

Conclusions

- Algorithms for automatic recognition of 3 food-texture attributes, namely crispiness, wetness, and chewiness
  - per chew recognition
  - per bout recognition
- Evaluation in LOSO and LOFTO approach
- High recognition for crispiness, promising results for wetness and chewiness

Thank you