

美味しさを伝えるための食事音に応じた視覚的表現提示システム

中川 慎哉 岩淵 志学 益子 宗 田中 二郎*

概要. 食べ物の美味しさを誰かに説明したい時がある。美味しさを感じる要素の1つである食感という言葉で説明することが難しいが、「パリパリ」「サクサク」などの咀嚼音から食感を連想することが可能である。また、ラーメンをすすする「ズズーッ」という音や炭酸ジュースを開封した時の「プシュッ」という音のような、咀嚼音以外にも食べ物の美味しさを連想させる音は多く存在する。しかし咀嚼音は基本的に食べている本人にしか聴こえないことや、人間が情報の多くを聴覚よりも視覚から得ていることから、音だけで美味しさを伝えることは不十分である。そこで本研究では咀嚼音などの食事動作によって出る音(以下、食事音)を取得、識別し、その音に応じた擬音文字などの視覚的表現を提示するシステムを開発した。本システムは食事音をフーリエ変換し、サポートベクタマシン(SVM)によって音の種類を識別し、対応する擬音文字を提示する。性能評価として、スナック菓子の咀嚼音とラーメンをすすった時の音を SVM に学習させ、それぞれを食べた時の識別精度を検証した。

1 序論

食べ物の美味しさを誰かに説明したい時がある。例えば、最近のテレビ番組によくある食レポ番組では、レポーターが飲食店に出向き、その店の料理の美味しさを紹介している。食べ物の美味しさに関わる重要な1つの要素に食感があるが[1]、食感は感覚的なものなので言葉にすることが難しい。そこで、「パリパリ」「サクサク」といった咀嚼音を聴くことで、その食べ物の食感を連想し、理解することができる。また咀嚼音に限らず、ラーメンをすすする「ズズーッ」という音や炭酸ジュースを開けた「プシュッ」という音など、食べ物の美味しさを連想させる音は多く存在する。

このような音を用いて他者に美味しさを伝える方法があるが、この方法にはいくつか問題がある。まず、食感を理解するための咀嚼音は基本的に食べている本人にしか聴こえず、リアルタイムに周囲の人にその食べ物の食感を説明できない。さらに、人間は情報の多くを聴覚よりも視覚から得ているため[2]、美味しさを伝えるためには音だけでは不十分である。そこで本研究では咀嚼音のような食事動作によって発生する音(以下、食事音)を識別し、その音に合った擬音文字などの視覚的表現を提示することで、他者にその食べ物の美味しさをより強く伝えることを目的とする。このような表現は映像を手動で編集すれば実現可能だが、本研究ではリアルタイムかつ自動的に音の識別と視覚的表現の提示を行う。

2 関連研究

食事に関する研究は幅広く存在し、本研究のように音に着目した研究も存在する。小泉らは、下顎部に貼り着けたマイクから咀嚼音を取得し、その音を加工してフィードバックすることで食感を変化させるシステムを開発した[3]。本研究では咀嚼音の取得のためにこの下顎部にマイクを貼り付ける手法を利用した。この研究は咀嚼音に対して一定の加工を施すだけで、咀嚼音から食べ物の識別はしていない。

食べ物を識別する関連研究として、Kadomura らは食べ物の色と食べ物をフォークで刺したときのフォークの先端間の電気抵抗値から食べ物を識別する手法を用いている[4]。だが、スナック菓子などの食べ物は噛んでいると柔らかくなって咀嚼音が変化するため、この手法ではその変化に対応できない。そのため、本研究では音のみを判断材料とした。

3 システムの構成

開発したシステムの構成を図1に示す。ユーザは下顎部にマイクをテープで固定し、webカメラと向かい合う。マイクとwebカメラはPCに接続されている。PC上でアプリケーションを起動すると、ウィンドウにwebカメラの映像が表示される。その状態でユーザが食事動作を行うと、その動作で生じる食事音に応じた視覚的表現がウィンドウ上に提示される。本システムでは視覚的表現として擬音を文字として表示し、一定時間経過後に消すことにした。

4 システムの実装

本システムは Processing 2.2.1 を用いて実装された。マイクからの音の取得、および高速フーリエ変換に

Copyright is held by the author(s).

*Shinya Nakagawa, 筑波大学情報学群情報科学類,
Shigaku Iwabuchi and Soh Masuko, 楽天株式会社 楽
天技術研究所, Jiro Tanaka, 筑波大学システム情報系

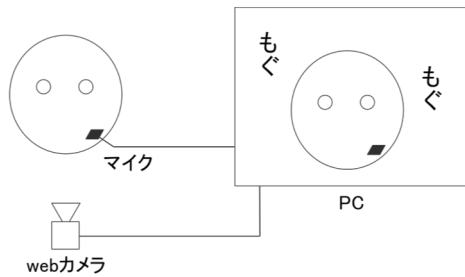


図1 システム構成図

は Processing の標準ライブラリである Minim ライブラリの機能を使用した。SVM は PSVM ライブラリ [5] の機能を使用した。開発、実行は Windows 7 搭載の PC で行った。

擬音がウィンドウ上に表示されていない時、システムはマイクからの音を受け取り、その音の大きさが閾値を超えているかを監視する。音の大きさが閾値を超えた時、その瞬間を食事行動があった瞬間と判断し、そこから約 250ms の間の音を入力音とする。入力音の長さは長いほど識別精度が向上すると考えられるが、この長さは食事行動と擬音表示の時間差でもあるため、擬音表示に違和感を覚えないう程度の長さとした。

入力音は高速フーリエ変換によって 1~1024Hz の周波数成分に変換される。1~100Hz の低周波成分はノイズを多く含んでいるためにカットし、101~1024Hz の周波数成分を正規化したものを SVM への入力とした。SVM からは各クラスに対するクラス所属確率が得られるが、この確率が閾値を超えた場合にのみ特有の視覚的表現を出すようにした。例えば、スナック菓子を食べた時は「ガリッ」と表示され、ラーメンをすすった時は「ズズウ」と表示される。どのクラス所属確率も閾値を超えない場合、曖昧な音と判断して「もぐ」と表示させた。

開発したシステムのスクリーンショットを図2に示す。識別結果に応じた擬音を web カメラの映像に重畳表示している。

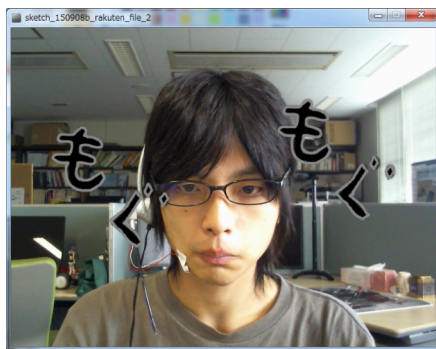


図2 実行画面

5 識別精度の評価

本システムの識別精度を検証するために、美味しさが連想される食事音であるスナック菓子の咀嚼音とラーメンをすする音のデータを SVM に学習させ、正しいクラスへの所属確率について検証した。スナック菓子は噛むごとに柔らかくなり咀嚼音が変化するため、最初の一噛みの音のみ学習させた。

検証の結果、スナック菓子を食べた時の正しいクラス所属確率の平均値が 84%、ラーメンをすすった時の正しいクラス所属確率の平均値が 61%となった。ラーメンの精度が低い原因として、すする速度によって音が変わることが考えられる。様々な速度のすする音を学習データとすることで解決できると予想している。

6 結論

本研究は食べ物の美味しさを他者に強く伝えるために、食事動作によって生じる音を認識し、それに応じた視覚的表現を提示するシステムを開発した。

今後の展望として、より美味しさを伝えられる視覚的表現への変更が挙げられる。現状の擬音文字を表示する方法では言語が異なる国の人が理解できないという問題がある。そこで、例えば炭酸ジュースを開けた時の「プシュッ」という音に合わせて画面に泡が浮き上がるアニメーションを表示するなど、言語に依存しない食感表現へと変更することで世界中に美味しさを伝えることが可能になる。また、喉にマイクを取り付けて喉越し音を取ることで、センサを増やすことで更なる応用が可能である。

参考文献

- [1] 稲見昌彦, 橋本悠希. 五感工学の最先端 5.食感の技術. 映像情報メディア学会誌 Vol.61, pp.1424-1426, 2007.
- [2] 教育機器編集委員会. 産業教育機器システム便覧. 日科技連出版社, p.4, 1972.
- [3] 小泉直也, 田中秀和, 上間裕二, 稲見昌彦. Chewing JOCKEY: 咀嚼音提示を利用した食感拡張装置の検討. 日本バーチャルリアリティ学会論文誌 18(2), pp.141-150, 2013.
- [4] A. Kadomura et al. Persuasive Technology to Improve Eating Behavior using a Sensor-Embedded Fork. *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, pp.319-329, 2014.
- [5] PSVM: Support Vector Machines for Processing - Makematics(2015年10月13日確認)
<http://makematics.com/code/psvm/>