

# SoundROKA: 日常生活における言葉の伝達を視覚的に提示する水槽

佐川俊介<sup>†1</sup> 小川剛史<sup>†2</sup>

日常生活において人間の声はとても弱い存在であり、騒音や環境音などの様々な因子によって、時折意図したものと異なる言葉として伝わる。その様子はさながら言葉が濾過（ろか）されるようであるが、それを現実世界で見ることは不可能である。本稿ではユーザの発した言葉が自然環境を模した水槽内で様々な影響により異なった言葉に変換される一連の遷移を視覚的に提示するシステムを提案する。

## SoundROKA: An Aquarium to Visualize the Propagation of Words in Real-life Situations

SHUNSUKE SAGAWA<sup>†1</sup> TAKEFUMI OGAWA<sup>†2</sup>

Human voice is very fragile in daily life. The voice is subject to change by various factors such as noise, environmental sounds. As a result, meaning of words that we have spoken is also changed. The situation seems that words are filtered, however it is impossible to see the changes visually in the real world. In this paper, we propose a novel system that visualizes the process of converting original words into other phrases depending on various influences in the aquarium representing the natural environment.

### 1. はじめに

我々人間の主なコミュニケーション手段は空気を媒体とした音声言語であり、日常生活において重要な役割を担っている。しかし、音声言語には時間的・空間的な制約も多く、その内容や存在がどのような状態で伝達していくのかを逐うように確認することはできない。そのため、最近では音声認識を使って会話を可視化する研究が盛んに行われており、ユニークな手法や表現方法を用いたものも多く存在する。例として、文献[1]のように食べかすが衣服に付着するメタファを利用したウェアブルインタフェースや、文献[2]のような音声を餌とし生物を育てるインタラクティブアート作品などが挙げられる。

本稿では SoundROKA と名付けた、日常生活において言葉がどのような影響を受け伝達されるのかを視覚的に提示するシステムについて説明する。このシステムではユーザが発した言葉を認識し、バイナリコードに変換、そのバイナリコードに対し、公園や高架下などの騒音の多い自然環境を模した仮想空間内で様々なフィルタリングを施し、入力時とは全く違った言葉となって出力するまでの一連の遷移を視覚的に表示する。

### 2. コンセプト

日常生活において、周囲の音に邪魔をされ意図したように声が届かないという状況は、誰でも一度は経験があるだろう。SoundROKAはその騒音のある日常生活のワンシーンにおいて、自分の声がどのように変換されて聞こえるのかを仮想的に提示する。本システムでは、言葉が変化する様子を“言葉を濾過する”と捉え、全体的に濾過をメタファとして取り入れている。これは言葉のフィルタリングと水の濾過をリンクさせることで、より直感的かつ効果的な表現を狙ったためである。通常、液体の自然濾過では上部から注入された液体を濾材に通すことで、液体と固体を分離している。SoundROKAでは、濾過されるのは入力された音声で、濾材の役割を果たすのはシチュエーションフィルタと呼ぶ、日常生活のワンシーンを仮想的に表現したグラフィックである。シチュエーションフィルタは公園や高架下など、現実世界において騒音の多い様々な環境を模して作られている。このフィルタ部分を通じた言葉は様々な影響を受け、最終的に入力時とは違った言葉となり表示される。もちろん、現実世界でも同じ様に変換され聞こえるわけではないが、もしかしたら現実でも起こりえるかもしれない言葉の変容を楽しむことができる。また、言葉が水中を漂うかのごとく落下していく様子は、言葉の儚さを効果的に表現しており、視覚的にも楽しめるものとなっている。

<sup>†1</sup> 東京大学大学院学際情報学府学際情報学専攻  
Graduate School of Interdisciplinary Information Studies,  
The University of Tokyo.

<sup>†2</sup> 東京大学情報基盤センター  
Information Technology Center, The University of Tokyo.

### 3. システム

#### 3.1 処理の流れ

本システムの処理の流れを図1に示す。まず、マイクから得られたユーザの音声をオープンソースである音声認識エンジン Julius[3]を使い、UTF-8形式のテキストデータとして取得、その後バイナリデータに変換する。入力された言葉がバイナリデータとなると、Flash CS6にて描かれたアニメーションが始まり、水槽上部から得られたバイナリ（0と1の文字）が落下してくる。バイナリは、落下途中にあるシチュエーションフィルタで様々な影響を受け濾過される。すべてのバイナリが落下を終えると、アニメーションがストップし文章の復元が始まる。復元は入力音声の品詞毎に行われ、通過したフィルタに応じた言葉を辞書データから取得、最終的には新しい文章となって表示される。

また、シチュエーションフィルタには様々な効果がある。例えば、日常生活の公園においては鳥や木々、子供など、音を邪魔する様々なものが存在する。本システムの公園フィルタも同様であり、風に揺れる木々は言葉を吸収し、元気に遊ぶ子供達は言葉を掻き乱す。噂好きの主婦達は言葉を付け加えるといった効果を持っている。最終的な復元の際は、バイナリがどの部分のフィルタを通過したのかをカウントしておき、辞書データから復元する単語を選択している。なお、公園の他にも高架下や雨の日など数種類のシチュエーションフィルタを用意した。

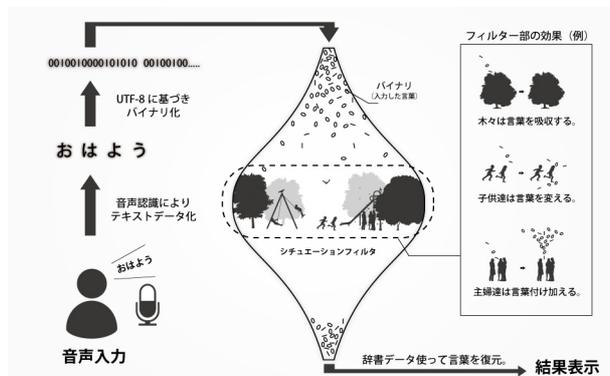


図1 処理の流れ

#### 3.2 展示方法

本システムの展示方法を図2に示す。本システムはそのコンセプトに基づき、水槽に投影した映像を水越しに鑑賞することで、まるで水中で言葉の濾過が行われているような表現を可能にしている。そのため今回、濾過器をイメージしたオリジナル形状のアクリル水槽を制作した。高さは400mmであり、前面と側面は透明アクリル、背面は半透明アクリルで構成している。什器には水槽の他にマイクを設置し、ユーザの音声を取得する。取得した音声をPCで処理した後、背面方向からプロジェクタと指向性スピーカーを用いて、水槽に映像と音声を提示している。

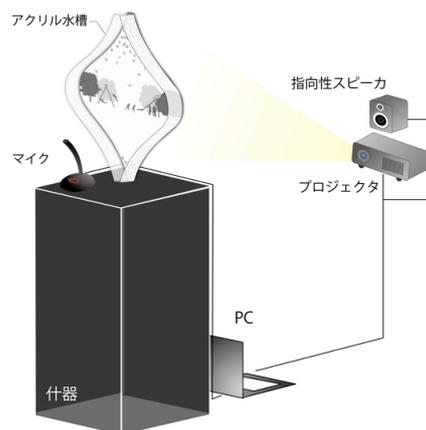


図2 展示方法

### 4. 展示による評価と考察

図3は2012年12月6日から10日にかけて開催された、第14回東京大学制作展にて、SoundROKAのプロトタイプを展示した時の様子である。来場者は5日間で495名にのぼり、大学生や教員、幼児に至るまで様々な方々が体験した。体験者からは映像が美しい、コンセプトと表現方法がマッチしている、フィルタの仕掛けが面白いなど好評を得た。一方で、音声認識精度や文章の復元方法については改善を求めるコメントも得られた。



図3 展示の様子

### 5. まとめ

濾過メタファを用いた日常生活における言葉の変化を視覚的に提示するシステムを実装した。今後は日常生活の実際の環境音を取得し、音量や周波数に応じてフィルタ部分の効果を変化させることで、現実世界の音とリンクした機能を実現したい。

**謝辞** 本研究の一部は日本学術振興会科学研究費補助金基盤研究(C)(25330227)の研究助成によるものである。ここに記して謝意を表す。

#### 参考文献

- 1) 清水大悟, 安村通晃: “しゃべりカス: 発話の視覚化を用いたウェアラブルインタフェース”, 情報処理学会研究報告, Vol.2009-HCI-132, pp.1-8 (2009)
- 2) Theodore Watson - Terrarium  
[http://www.theowatson.com/site\\_docs/work.php?id=44](http://www.theowatson.com/site_docs/work.php?id=44)
- 3) 大語彙連続音声認識エンジン Julius  
<http://julius.sourceforge.jp>