

EdgeSonic：視覚障がい者のための画像特徴の可聴化

EdgeSonic: Sonification of Image Features for the Visually Impaired

吉田 翼 木谷 クリス 真実 ベロンジー サージ シュレイ ケビン*

Summary 本論文では視覚障がい者のための物体形状の可聴化手法として、画像中のエッジとエッジまでの距離の可聴化を提案する。視覚障がい者は外界の物体形状を理解するために触覚を用いる。しかし、印刷された模様や風景など触覚によって認識できない物体も存在する。そこで本研究では、視覚情報を可聴化することで聴覚による物体形状の認識を目指す。提案手法では、画像情報を表示したタッチパネルに指で触れた位置に応じて、画像から抽出したエッジに対する2種類の可聴化を行う。エッジ形状を知らせるモードでは、矩形領域内のエッジを、音高と時間的変化にマッピングする。もう一つのモードでは、エッジまでの最短距離を、パルス波の間隔で表現する。本提案手法をタッチパネル式のモバイル端末上に実装し、視覚障がい者による実験を行った結果、エッジの勾配による音の変化によってエッジの形状が容易に認識できることが確認できた。また、距離に応じて変化する音によってエッジの検索が容易になることも確認できた。

1 はじめに

視覚障がい者は実世界を認識するために、視覚の代替として触覚を用いている。しかし、平面に印刷された模様や風景など、触覚によって認識することが出来ない物体も多く存在する。本研究では触覚で認識できない視覚情報を可聴化することにより、実世界の認識を目指す。提案手法では、画像情報を表示したタッチパネルに指で触れた位置に応じて、エッジ勾配とエッジへの距離情報を可聴化する。

2 関連研究

可聴化の研究は一般的に健常者の視点で行われているが、本論文では視覚障がい者のための可聴化の観点から関連研究を示す。従来手法では概念的な音（言葉）による可聴化が行われている。自動化された物体認識と言葉によるフィードバックを行うシステム LookTel[1]では、認識したい物体の画像をデータベースに保存しておくことで、物体の名称を読み上げることができる。しかし、各物体の識別器の学習と音声ラベルの準備を予め行う必要がある。また VoiceOver(Apple 社)や JAWS(Freedom Scientific 社)では text-to-speech(TTS)システムを利用し、コンピュータ上の文字を音声化する。しかし言語による可聴化には、非言語的な情報を音声化して提示する事が難しいという問題がある。

画像特徴そのものを提示する Ivanら [2]の研究では色情報を2つの基本色に分割し、音色のパラメー

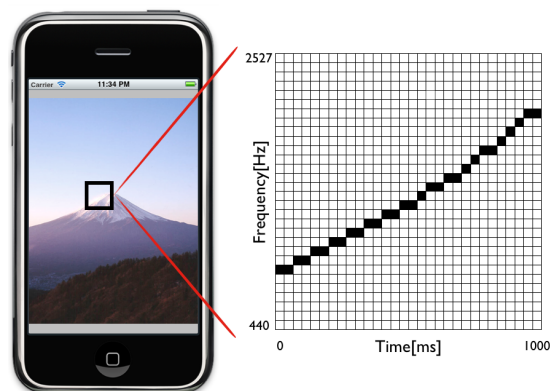


図 1. エッジから音へのマッピング

タ（エンベロープ、波形、周波数）に当てはめて可聴化を行う。また、vOICe[3]は画像全体の輝度情報をサイン波の組み合わせで可聴化する。具体的には画像の輝度値を音の振幅、垂直軸を音高とし、画像全体を水平軸の左から右に1列ずつ走査する音を流し続ける。しかしvOICeはグローバルな表現に限定されているため、画像の詳細な情報提示ができない。これに対して、本研究では指で指定した局所領域の可聴化を行う。局所領域を扱う研究としてタッチパネル上の地図情報を可聴化する Timbremap[4]が挙げられる。Timbremapでは道を表す線に対する指の接触位置に応じた音を鳴らし、道の発見を補助する。具体的には線の上、左右を示す3種類の音の割合によって道の位置を知ることができる。しかし、線上での音は一定であり、線がどの方向に伸びているか理解する事は難しい。一方本研究では局所領域におけるエッジの傾きを可聴化することで、エッジの伸びる方向を知ることができる。

Copyright is held by the author(s).

* Tsubasa Yoshida and Kris Kitani, University of Electro-Communications; Serge Belongie, University of California at San Diego; Kevin Shlei, University of Wisconsin

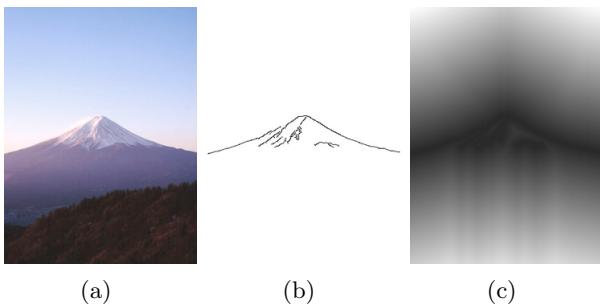


図 2. (a) 元画像, (b) エッジ画像, (c) 距離画像

3 アプローチ

本研究では局所領域の可聴化を行う。視覚障がい者は点字や点図から情報を得る際、指先を利用して詳細情報を得る。同様本研究では指先が触れている画像領域に対して、直感的な局所領域の情報提示を目指す。具体的には撮影物体を認識するためにエッジ勾配を可聴化し、近傍のエッジへの最短距離も音に対応付ける。

3.1 エッジと距離画像の抽出

人は外界から視覚情報を得る際、物体の仮想的な輪郭線を取得し、物体の形状を認識するための重要な情報源としている [5] [6]。そこで本研究では画像からエッジを抽出し、その輪郭線を音で表現する。輪郭線の取得には、画像ノイズにロバストなキャニのエッジ検出 [7] を用いた。また、輪郭線の発見を補助するために、各ピクセルの近傍エッジまでの最短距離を計算した。距離画像の作成には Felzenszwalb の手法 [8] を用いた (図 2)。

3.2 局所特徴の可聴化

指の接触点がエッジ上か否かによる 2 種類のモードで可聴化を行った。エッジ形状を知らせるモードでは、接触点を中心とした矩形領域を水平軸に沿って左から右に走査する音を生成した。この可聴化によって、矩形内でのエッジの形状を表現できる。

もうひとつのモードでは、指の接触点におけるエッジまでの最短距離に応じたパルス波を鳴らす。パルス波の波数は距離を示し、エッジに近づくほどパルスの間隔が短くなる。この可聴化によって、指からエッジまでの距離を表現する。

4 実装

本研究では初歩的な取り組みとして、タッチパネル式のモバイル端末上 (iPhone) で画像の可聴化システムを実装した。音色にはサイン波形を用い、エッジの局所的な可聴化には、周波数 440[Hz] から 2527[Hz] の音源を準備した (30 段階、ステップサイズは 1.06 の累乗)。矩形サイズは人差し指で一度

に触れることが出来る 30×30 ピクセルとした。また、エッジは非常に細く指で補足することが難しいため、エッジの近傍に触れたときには、接触点を最も近いエッジ上の点に近似させた。近似半径は 10 ピクセルとした。

5 考察とまとめ

視覚障がい者による予備実験により、音の変化によるエッジの勾配の認識が容易であることが確認できた。また、学習によって音の変化に従って指をエッジ上に動かすことができることも確認できた。さらに、エッジまでの最短距離の可聴化により、輪郭線の発見を補助することが確認できた。エッジ情報に限らず、距離画像は他の情報 (例えばボタンへの距離) を検索する際にも役立つと考える。一方、エッジを辿る際、指の絶対位置が不明になるという問題があった。そのため、現在指の絶対位置を確認できる可聴化モードの開発に取り組んでいる。エッジを辿りつつ、全体の中での位置を可聴化することにより、輪郭線の認識がより高速に行えると考える。

本論文では視覚障がい者向けの画像情報の可聴化手法として、エッジとエッジまでの距離の可聴化を提案した。モバイル端末上で実験を行った結果、可聴化によってエッジの形状が容易に認識できることが確認できた。今後はより多くの視覚障がい者の方に本システムを使ってもらい改良を進めるため、今回開発したアプリケーションは Apple 社の App Store にて無料配布する予定である。

参考文献

- [1] J. Sudol, O. Dialameh, C. Blanchard and T. Dorcey. Look-Tel: A Comprehensive Platform for Computer-Aided Visual Assistance. In *Proceedings of the Workshop on Computer Vision Applications for the Visually Impaired*, 2008.
- [2] K. Ivan and O. Radek. Hybrid Approach to Sonification of Color Images. In *Proceedings of the International Conference on Convergence and Hybrid Information Technology*, 2008.
- [3] P.B.L. Meijer. An Experimental System for Auditory Image Representations. In *IEEE Transactions on Biomedical Engineering*, 1993.
- [4] J. Su. Enabling the Visually-Impaired to Use Maps on Touch-Enabled Devices. In *Proceedings of the MOBILECHI*, 2010.
- [5] R. L. Gregory. Cognitive Contour. In *Nature*, vol 238, pp.51-52, 1972.
- [6] I. Rock and R. Anson. Illusory Contours as the Solution to a Problem. In *Perception*, vol.8, pp.665-681, 1979.
- [7] J. Canny. A Computational Approach to Edge Detection. In *IEEE Transactions on Pattern Analysis and Machine Intelligence.*, 1986.
- [8] P.F. Felzenszwalb and D.P. Huttenlocher. Distance Transforms of Sampled Functions. In *Cornell Computing and Information Science Technical Report*, 2004.