

医用画像診断における画像特徴の可聴化に関する検討

賀川 経夫^{1,a)} 田上 秀一^{†1} 清末 一路^{†1} 森 宣^{†1} 西野 浩明¹

概要: 近年, 医用画像処理に関する技術の急速な進歩に伴い, CT や MRI など放射線等を利用した身体内部の撮影においては, ミリ以下の単位での断層を高解像度の画像で撮影することが可能となった. しかしながら, そのような画像を用いた診断の際には, 診断医師は, 一つの事例あたり膨大な数の断面画像を精査しなければならない. したがって, 何例もの診断を行う場合には, 1 日に数万枚に及ぶ画像を検査することもある. 本研究では, このような負担のかかる画像診断において, 効果音を利用することにより, 診断者の負担を軽減することを目的とする. 音情報を効果的に用いるためには, 画像における様々な情報を適切に音として表現する必要がある. そこで, 本論文では, 画像内の特徴に基づいた効果音を生成するための可聴化処理に関して検討を行う.

A Sonification Method of Image Features for Medical Diagnosis Supporting System.

Abstract: Recently, Computer Aided Diagnosis (CAD) has become one of the most important for medical activity. The more exact and various CAD become, the larger amount of medical images are provided. Furthermore, these images become high definition. Radiologists have to cost their time and efforts to investigate these medical images. It is strongly required to reduce their burden without debasing the quality of imaging diagnosis. In this paper, we propose the technique to generate sound information based on the image features and discuss their effects for diagnosis. Generating sound effects helps for attention rousing and the fatigue reduction for medical imaging diagnosis.

1. はじめに

医療分野において, 画像処理環境の発達とソフトウェアの充実により, 高品質の画像を利用して CT や MRI などの画像診断が行えるようになった. CT や MRI などの連続断面画像による診断では, 高精細ディスプレイ上に診断用の画像を出力し, 3 次元化や強調処理などの画像処理を駆使して, 診断者が患者の疾患やその前兆を発見する.

画像診断については, 異常個所における既知の特徴を強調し, 診断者による位置や病状の推定を容易にする手法が多く提案されており, 一部の事例に対しては, 大きな成果を挙げている [1], [2]. しかしながら, 病気の種類や進行状況, 疾患の状態によって, 抽出すべき画像特徴が大きく

異なるため, 特徴を絞り込むための画像処理の汎用化が非常に困難であり, 完全に自動的な判断を行うには至っていない. また, 画像を強調する多くの手法においては, 安定した処理結果を得るために, 正規化等の前処理を適切に行う必要があるため, 画像処理の精度が大きな問題となる. そのために, 異常でない部分も異常個所の候補として抽出されることがある. したがって, 画像処理技術や処理環境がいくら整備されても, 最終的には診断者の視診によって異常の有無も含めて異常個所に対する診断を下す必要がある. このような加工された画像を用いて診断を行う場合, 正確な診断を行うためには, 加工のためになされた処理内容とその性質を診断者が正確に知る必要があるため, 画像処理に関する知識も要求される.

このように, 画像診断の精度が向上した分, 専門医師は異常個所を見つけるために, 大量の高精細画像を全て集中して精査しなければならない. 現在, 断面画像は, ミリ以下の単位で撮影することが可能であるため, 一人の患者の分でも数千枚に及ぶことがある. 1 日に多くの症例を扱う

¹ 大分大学工学部知能情報システム工学科
Department of Computer Science and Intelligent Systems,
Oita University, Oita, Japan
^{†1} 現在, 大分大学医学部放射線科
Presently with Department of Radiology, Oita University
Faculty of Medicine, Oita, Japan
^{a)} t-kagawa@oita-u.ac.jp

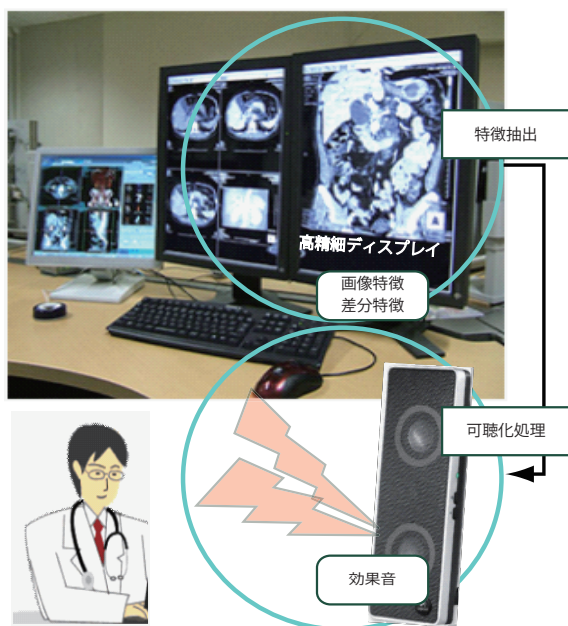


図 1 効果音を用いた医用画像診断支援手法

場合には、調査対象となる画像が何万枚にも及ぶため、相当の負担を背負うことになり、疲労による見落としなどのヒューマンエラーにつながる恐れがある。

本研究では，画像診断における診断者の負担を軽減するために，精細度の高い画像を見続ける医用画像診断システムの構築において，複数モダリティの利用，すなわち視覚モダリティである画像に加えて聴覚モダリティである音情報の利用を考える [3], [4], [5], [6]．すなわち，図 1 のように診断画像に相応した効果音を適切に選択し，画像と同時に提示することによって診断の支援を行う手法を提案する．このように効果音の変化によって画像の変化を知らせることができれば，診断者に対する注意喚起となり，画像を見続けることに関する疲労を低減することができると考えられる．また，診断に必要な画像特徴を音で表現することにより，必要以上の画像強調処理を行わず，見慣れた画像にできるだけ近い画像を利用することが可能となる．本論文では，医用画像内の特徴に基づいた効果音を生成し，医用画像診断時における診断者の負担を軽減するための可聴化処理に関して検討を行う．

2. 画像の可聴化

可聴化とは、表現する情報を音や振動に変換し、聴覚モダリティとして提示する技術である。多くの研究が、視覚障害者を対象として画像を可聴化して提示することにより、生活の支援を行うことを目的としてすすめられている。画像の可聴化については、音声合成 (Text to Speech, TTS) を用いて画像内容を自然言語で説明する手法が提案されている [7], [8], [9] が、言語音声化については、用途が限定される、度合いなどの言語化が困難な情報が含まれると音声化が困難となるといった問題がある。また、カメラから得

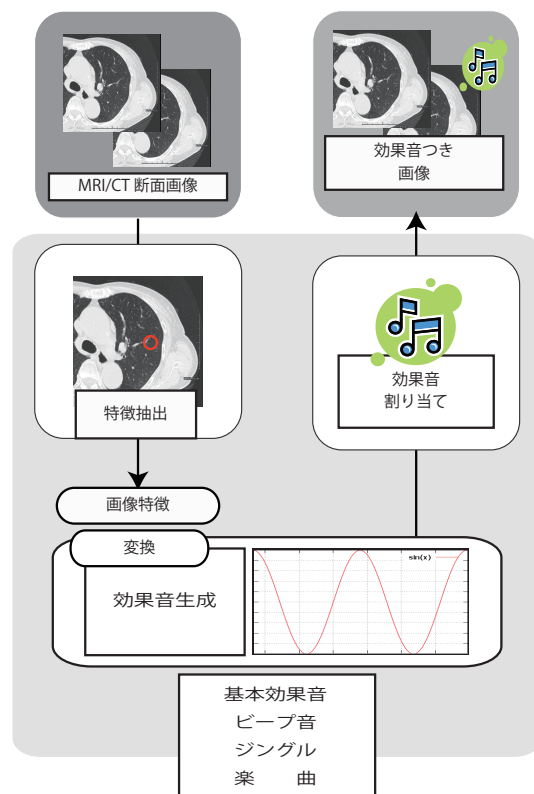


図 2 医用画像診断支援システムにおける効果音の生成処理

られた画像に対して，色分布や輝度情報，エッジ情報などの画像の特徴を，一定の規則にしたがってリアルタイムに音に変換することにより，その音によって視覚障害者に環境を認識してもらう手法が提案されている [10], [11], [12]．しかしながら，音は画像に比べて表現できるチャンネルが狭いために，詳細な情報の表現が困難である．したがって，このような手法は，大まかな情報の提示には有用であるが，詳細な情報を提示するのに多くの音情報が必要となり，それらを聞き分けることが困難となる．また，利用者は，カメラで撮影された画像の特徴と出力された音との対応関係をあらかじめ習得し，音による情報の提示にあらかじめ慣れる必要がある．

本研究では、これらの技術を参考に、画像内容から得られた音声や効果音を画像診断に併用するマルチモーダルな手法について検討する。医用画像を閲覧する際に効果音など他のモダリティを利用することにより、大量の断面画像を見続けることによる視覚の疲労を軽減し、効率よく診断を行うことができる。また、診断に必要な画像特徴を音で表現するため、必要以上の強調処理などの画像処理を行わず、見慣れたオリジナルの画像を精査することができる。

3. 医用画像診断支援手法

本手法の処理の概要を図 2 に示す．図 3 に診断対象となる肺の断面画像シーケンスの一部を示す．これらの画像では，白い大きな領域が筋肉や脂肪，骨を表現している．内

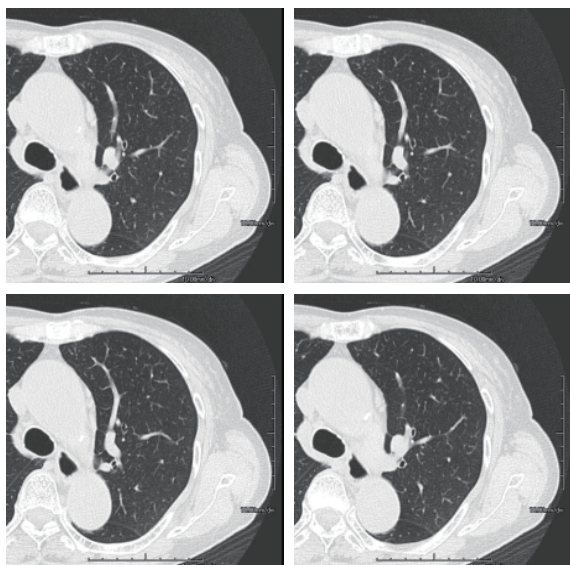


図 3 CT による断面画像シーケンス

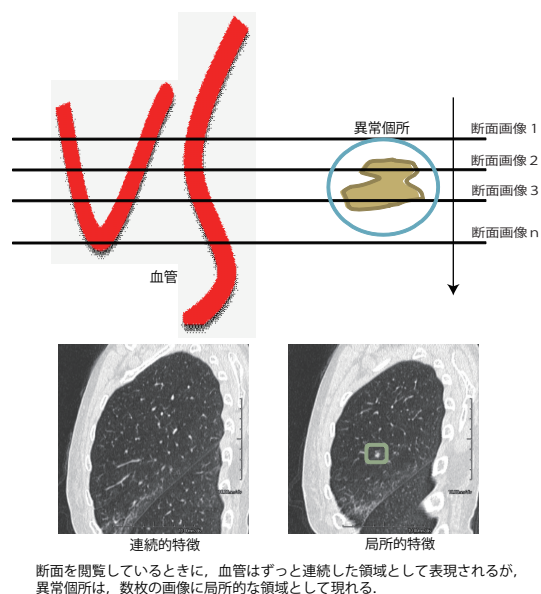


図 5 異常個所の構造

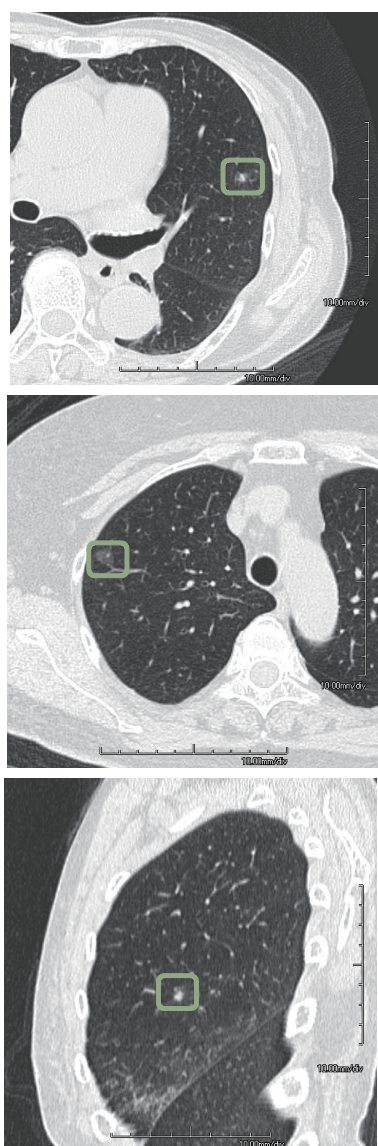


図 4 異常箇所

部の白い細かい領域が血管であり、この中に異常箇所が存在している。発見すべき異常箇所の例を図 4 に示す。これらの病変は、基本的に図 5 のようになっており、画像シーケンスの中で断片的な特徴として表出する。対照的に血管については、連続的な構造として現れる。診断の際には、局所的な情報だけでなく、構造的な情報を把握する必要がある。

これらの処理対象である CT や MRI 断面画像のシーケンスが与えられると、各画像は、256 階調の濃淡画像に変換され、以下に示す処理により、画像シーケンスに効果音が付与される。

- (1) 画像の特徴抽出
病状や画像の種類に応じて各画像から様々な特徴が抽出される。
- (2) 特徴から効果音の生成
音を生成するためのパラメータに対応づけられる。
- (3) 効果音の割り当て
生成された効果音を各画像に割り当てる。
- (4) 診断用画像データの生成
効果音が付与された画像のシーケンスが再生成される。
診断者は、基本的に先頭から逐次的に断面画像を検査していく。また、必要に応じて、前後の画像を比較したり、後ろから順番に検査することがある。どのような場合でも効果音を出すことができるように、全ての画像に対して効果音を付与し、その画像を表示する際に対応する音を出力するようにする。

画像診断の支援を目的とする医用画像の可聴化手法の構築に関して、以下の点を検討する必要がある。

- (1) 画像特徴と効果音の対応づけ
 - 画像特徴のもつイメージと音の効果が一致しない場合、診断者に混乱を招く恐れがある。例えば、着目す

る領域が大きくなっていくように見えるときに、効果音が小さくなったりすると、違和感が生じてしまい、画像診断の妨げとなる。音と画像とに互いに整合性を持たせ、その関係を診断者が直感的に理解できるように提示しなければならない。

- 画像では、識別しにくい特徴に対してより明瞭な効果音を付与する。視認しやすい画像特徴については、モダリティ間で冗長になるのであるべく音に変換しない方がよい。
- 断面画像のシーケンスの内であまり顕著に表れない特徴に関して、反応するような効果音をデザインする必要がある。顕著に表れる特徴へ相応の効果音を付与しても、定常的に音が提示されるため、あまり効率的ではない。
- 複雑な特徴を音に変換する場合、診断時に音を精細に聞き分けなければならない、視覚的にも聴覚的にも、余計な負担をかけることになる。したがって、効果音にはある程度の冗長性を持たせる。

(2) ユーザインタフェース

医用画像の用途や撮影箇所、予想される症状などによって、画像内に表れる特徴が変化し、それに応じて診断者の興味も変化する。したがって、診断者の注意する部分や画像特徴に関して、画像処理に関する知識が乏しくても、その意図を正しく反映し、思い通りに操作できるインタフェースを設計する必要がある。

4. 画像特徴と効果音の対応づけ

4.1 画像特徴

本来、CT や MRI のような断面画像について、診断者が注視すべき特徴は、以下に挙げることに大きく依存する。

- 人体の部位
- 患部の生化学的特性
- 画像内における視認性
- 画像の解像度
- 断層の解像度

これらの状態により得られる画像の種類や画像の内容は大きく変化する。したがって、一意に音や画像特徴を決定することができない。

診断画像の特徴抽出を行うにあたり、以下のように画像特徴を分類する。

(1) 局所画像特徴

医用画像シーケンスの各画像から得られる明度分布や幾何構造などの特徴。シーケンス内の局所的な特徴であり視認もしやすい。しかし、前後の画像間における微細な変化が分かりづらい。

(2) 連続画像特徴

連続した複数の画像によって構成される特徴量を指す。医用画像のシーケンスにおける前後の画像間の差

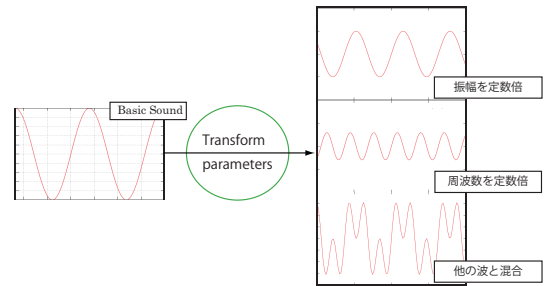


図 6 パラメータの変化によるサイン波の変化の例

分や画像全体の和などの特徴が挙げられる。一枚ずつ画像を検査する場合には、構造や差分情報などは視覚的に表現するのが難しい。

本手法では、双方の特徴を利用しながら、診断者が利用しやすい音を提供できるようにする。

4.2 効果音の生成

4.2.1 効果音の種類

本研究では、初期のテストベッドを利用して効果音の与える影響について検討してきた [13]。現段階では、生成・出力する効果音は、以下の 3 種類の音源を利用する。

- (1) ビープ音：単音による効果音。単純なサイン波を利用する。周波数や振幅など、スカラー量で得られる特徴を反映しやすいため、詳細な制御が可能である。
- (2) ジングル：チャイムなどの 1 小節程度のまとまった音楽。短い音で様々な意図を表現することができる。
- (3) 楽曲：ある程度のまとまった曲。

ジングルと楽曲については、著作権フリーの音源を利用する。また、生成される音は単一ではなく、曲を聴きながらジングルやビープ音を鳴らすというように複数の音源を利用する。

4.2.2 効果音の特徴

効果音は、以下に挙げる特徴を考慮して作成する。

- 快適さ：診断者は、診断を行う間ずっと音を聞き続けることになる。シンプルな単音のビープ音などは、連続して聞くと耳障りな音となる。また、ジングルなども同じものを何度も利用すると次第に不快な音に変化する。
- 聞きやすさ：特徴量の変化に対応して、音を変化させる場合には、診断者がその変化に気づく必要がある。そこで、変化に気づきやすい音を選択できる必要がある。聞きやすさについては、個人差や利用環境による差が大きいため注意しなければならない。
- 制御しやすさ：画像特徴を反映させながら音を変化させることが容易である必要がある。ビープ音は変化を与えやすいが、曲については特徴量によって変化を与えることは容易ではない。

4.2.3 効果音の制御

効果音について、サイン波に基づくピープ音は、基本的に以下の式によって表現される。

$$s(n) = A \sin\left(\frac{2\pi f_0 n}{f_s}\right) \quad (1)$$

この式のうち、振幅 (A) と周波数 ($\frac{f_0}{f_s}$) に相当するパラメータを対応させる画像の特徴量に応じて変更することによって、特徴量に基づく多様な効果音を生成する。図 6 にパラメータの変化に応じたサイン波の変化を示す。

実際の効果音を生成するにあたり、音に関する以下のパラメータを操作することにより、様々な音を生成する。

- 連続パラメータ

- 音量：効果音による特徴の表現に関して音の大きさは重要な要素であるが、大きすぎたり、小さすぎたりすると、集中できなくなり、思考の妨げになる恐れがある。
- 周波数、速度：音の波形のピッチを変化させることにより、曲の速さを変更することができる。ピープ音等には有効であるが、楽曲のピッチを変化させ続けると違和感が生じる。

- 離散パラメータ

- リズム、鳴動時間：効果音を生成する上で、これらのパラメータは、診断の行動にも大きな影響を与える重要な要素である。しかしながら、病状や診断者の興味などの個人差の影響が大きいので、それらに応じた設定を可能にする必要がある。

- その他

- 音色：現段階では、MIDI などを利用して、聴き心地の良いピアノ等の楽器の音の利用も検討している。そのために、音と画像に関する感性評価を行い、聴き続けても苦にならない音のデザインを検討する必要がある。
- 音源の位置：仮想的に設定される 3 次元空間の音源位置を変化させることにより、画像内の大まかな位置を表現することができる。

4.3 効果音の割り当て

効果音生成の概要を図 7 に示す。医用画像診断について、診断者は、異常箇所を発見するために、膨大な数の断面画像を病患の有無を含めて調査する必要がある。この時に、患者の症例や状態、病患の状態や経過時間によって、画像における異常箇所の見え方は大きく異なるため、画像特徴と効果音の対応付けを一概に決めることはできない。

例えば、画像の特徴量と音量や音の高さを単純な線形比例関係にすると、特徴値が非常に大きくなったときに非常に不快な音になったり、人間の可聴域を超えたりしてしまう恐れがある。また、面積が大きい領域と小さい領域では、前後の画像間での変化の割合が同じでも、絶対的な変

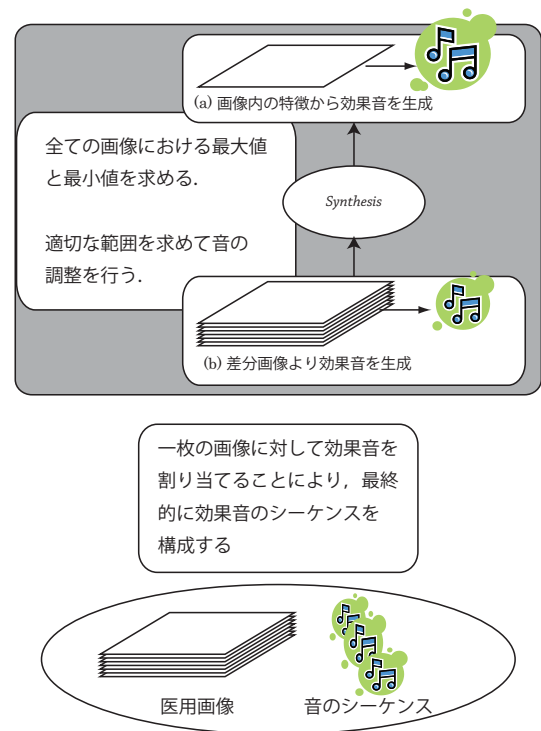


図 7 効果音の割り当て

化量が異なるため、小さい領域の微小な変化は音として表現できにくくなる。

したがって、画像特徴量と音のそれぞれに対して正規化を行う必要がある。また、あらかじめ全ての画像に対して特徴量を抽出し、その総変化量から最大値と最小値をあらかじめ求め、このような問題を回避する。最終的に、図 7 に示すように、単一画像における特徴と連続画像の特徴を併用して、全ての画像に効果音を付与し、画像に沿った効果音のシーケンスが得られる。

5. ユーザインタフェース

本手法の利用者インタフェースを図 8 に示す。インタフェース上の画像表示部には、適宜処理中の画像が表示される。診断時には、サイズを変えてできるだけ高解像度で閲覧ができるようになる。音に関しては、つまみなどを設定し、好きな音へ調整可能となるようにする。領域分割された画像とひとつ前の画像との間の差分画像が提示される。診断者は、マウスのホイールを用いて画像を一枚ずつ逐次的に閲覧する。

6. 動作例

現在のテストベッドを用いて、簡単に動作させた。実験用のパターンは、57 枚の肺の断面画像のシーケンスを利用した。

画像処理に関しては、以下のような処理を行った。

- (1) 局所画像特徴

- (2 値画像) 元画像を 2 値化し、領域の膨張と収縮を

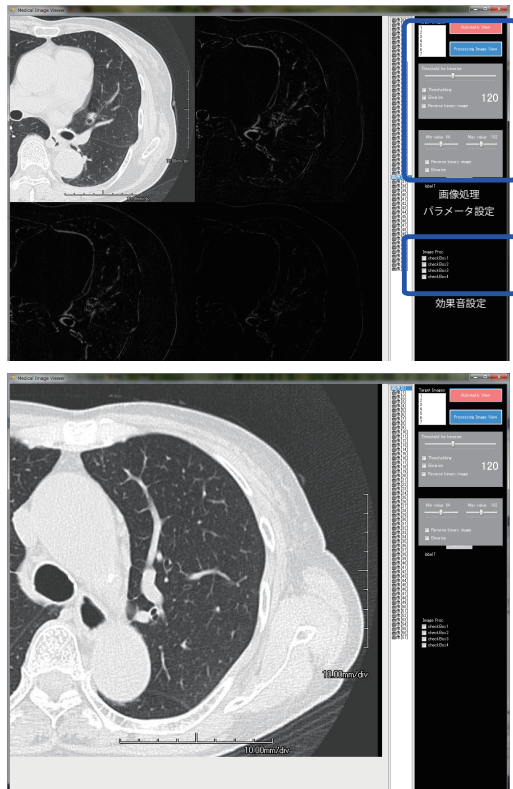


図 8 ユーザインタフェース

行うことによりマスク画像（図 9(b)）を生成し，それを利用して筋肉や骨の部分を除去する．残った画像の白領域（図 9(c)）の面積（画素数）を計測する．

(2) 連続画像特徴

- （差分画像）差分画像 n 番目の画像に関して $n-1$ 番目の画像との差分画像と，同様に $n+1$ 番目の画像との差分画像を生成し，それらの差分（2 次差分）画像を生成する．その後，2 値化し，白領域の面積を計測する（図 10）．
- （構造画像）シーケンス内の全ての画像に対して 2 値化を行い，重みづけ（ $\frac{1}{N}$ ， N は，対象画像の枚数）を行い，各画素の和を求める（図 11）．この画像は，断面画像によって構成される 3 次元の構造を表現している．この中で円形度の高い領域は，図 5 における異常個所である可能性がある．そこで，上記の局所特徴である 2 値画像の重なる面積を計測する．

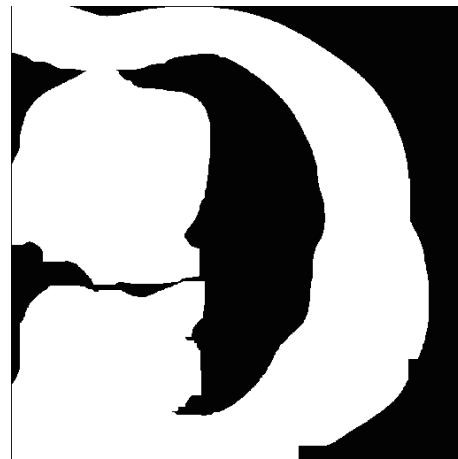
効果音については，ピープ音とジングル（ピアノのワンフレーズ）等のピッチと音量をそれぞれ，2 値画像における面積，差分画像の面積，構造画像の面積に比例するように対応させた．組み合わせは適宜変更して動作させた．

このような動作例により本手法を適用したところ，全体的にただ見るよりは，効率的に画像を閲覧することができる．

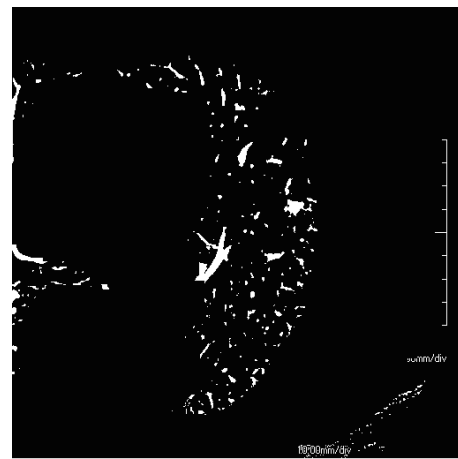
局所的な画像特徴に関しては，インタフェース上の画像をみて比較的に分かりやすい画像特徴に基づき効果音を



(a) 元画像



(b) マスク画像



(c) 2 値画像

図 9 局所画像特徴

生成しているため，利用者にとっては冗長な情報の提示となっている．画像特徴が分かりにくいのであれば，この冗長性は，特徴の検出に有効に作用するが，視認しやすい特徴の時には邪魔に感じてしまうことがある．この点については，文献 [13] と同様の結果が得られた．また，連続的な画像特徴に関しては，比較的効果的に作用した．しかしながら，効果音だけでは，何の音が出ているのか把握しづらい時があり，操作に戸惑ってしまうことがある．効果音に

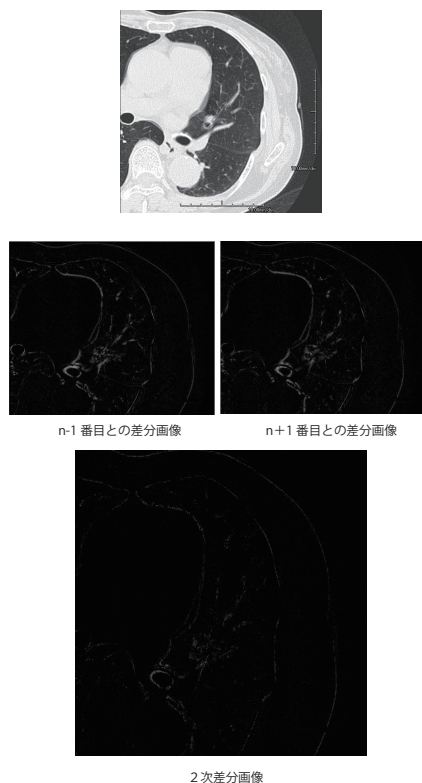


図 10 連続画像特徴 2次差分画像

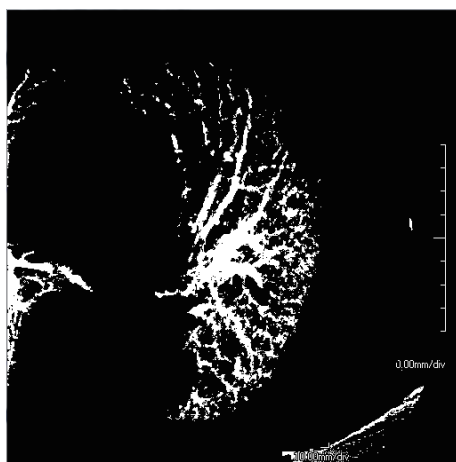


図 11 連続画像特徴 構造画像

については、ボリュームの差異よりは、ピッチの差異の方が鋭敏に感じることができた。しかし、同じ音を何度も繰り返し聞いていると差異の感じ方に揺らぎが出てくるため、効率が落ちてしまうことがある。この点については、適宜音を変えていく、楽曲と併用するなどの対策が必要である。

7. 検討

今回の評価では、簡単に一通り眺めて、異常個所が見つかるかどうかを考えると非常に単純な操作を行ったため、実際の運用形態とは必ずしも一致していない。実際には、もっと診断者が疲労した状況での利用を想定しているので、普段画像診断を行う医師が被験者となってもらい、

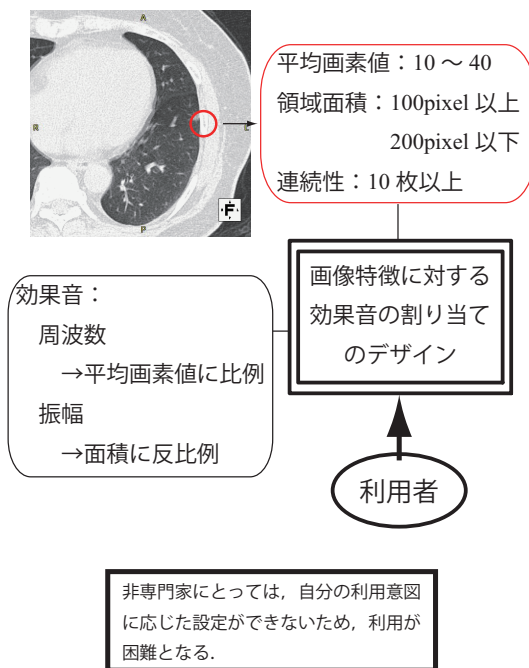


図 12 特徴と効果音の複雑な対応付けのデザイン

できるだけ膨大なデータ量を再現したうえでの運用動作を行った上での評価を行わなければならない。その上で、必要な機能の取捨選択をしながら、本手法を洗練していく必要がある。

医用画像における画像特徴については、さらに検討する必要がある。現在は、本学医学部より提供された、比較的異常個所が発見しやすい画像シーケンスを用いているが、現状のままでは、汎用性の高い手法をデザインすることは困難である。利用者である放射線医の方と緊密に意見交換を行い、どのような特徴が必要かをさらに洗い出す必要がある。

今後は、実際の運用実験を重ねるとともに、インタフェースの充実を行っていく。実際の診断においては、診断者の自分の好みや意図、興味に応じて効果音を生成する必要がある。この時に、例えば、図 12 のように見つけたい画像特徴が単純ではなく、生成される音も複雑な規則を使わないといけない場合、画像や音声に関する知識の乏しい非専門家は、設定が煩雑で困難なものとなるため、利用することができなくなってしまう。最終的には、症例に応じて適切な特徴量が対象画像より抽出されるように設定されなければならない。そのためには、診断者が利用したい特徴量を自由に選択することのできる仕組みをインタフェース上に構築すべきである。そこで、なるべく簡単な GUI 操作によって医師の注視したい特徴に合わせた音情報の生成と出力を、画像や音声の非専門家である医師が指定しカスタマイズできる環境を構築する必要がある。また、個人差による判断の揺らぎに対応するためには、各人の好みも反映できるインタフェースの設計が必要である。

8. まとめ

本研究では、医用画像の画像特徴を可聴化し、その特徴に基づく効果音を用いて表現することで、医師の画像診断作業における負担や疲労を軽減しながら支援する手法について提案した。肺断面を診断する評価ツールを構築し、その有効性に関して評価を行った。

今後は、実際に画像診断に携わる方を対象として、現段階のツールを利用して効果音生成に関する詳細な評価実験を実施する。また、より詳細な設定が可能なシステムとインタフェースの構築を行っていく。

また、例えば、監視カメラにおける効果音生成の有効利用など他の分野への適用とアプリケーションの枠組みについても検討していく。また、複数モダリティによる情報提示について、音の生成だけではなく、触覚による提示についても検討していく。

謝辞 本研究の実施にあたり、科学研究費(基盤研究(C)課題番号 24603018)の助成を受けた。

参考文献

- [1] K.Doi, *Current Status and Future Potential of Computer Aided Diagnosis in Medical Imaging*, The British Journal of Radiology. Special Issue, pp.S3–pp.S19, 2005.
- [2] 藤田広志, 医用画像のためのコンピュータ支援診断システムの開発の現状と将来, 日本写真学会誌 2003 年 66 巻 5 号, pp.484–pp.490, 2003.
- [3] X.Huang and F.Zhang, *Mult-modal Medical Image Registration Based on Gradient of Mutual Information and Hybrid Genetic Algorithm*, Proc. of The 3rd International Symposium on Intelligent Technology and Security Informatics, pp.125–pp.128, 2010.
- [4] C.Doukas, I.Maglogiannis and K.Karpouzis, *Context-Aware Medical Content Adaptation through Semantic Representation and Rules Evaluation*, Proc. of the 3rd International Workshop on Semantic Media Adaption and Personalization, pp.128–pp.134, 2008.
- [5] R.Singh, M.Vatsa, A.Norrne, *Multimodal Medical Image Fusion using Redundant Discrete Wavelet Transform*, Proc. of 7th International Conference on Advances in Pattern Recognition, pp.232–pp.235, 2009.
- [6] D.Sonntag and M.Moller, *A Multimodal Dialogue Mashup for Medical Image Semantics*, Proc. of the 15th international conference on Intelligent user interfaces, pp.381–pp.384, 2010.
- [7] A.Rodriguez, J.Labra, G.Hernandez, J.Gomez and R.Gomez, *ADONIS: Automated Diagnosis System based on Sound and Precise Logical Descriptions*, Proc. of 22nd IEEE International Symposium. 2009
- [8] 吉田翼, 木谷クリス真実, ペロンジーサージ, シュレイケビン, *EdgeSonic: 視覚障がい者のための画像特徴の可聴化*, 第 18 回インタラクティブシステムとソフトウェアに関するワークショップ WISS2010, 2010.
- [9] Peter B.L.Meijer, *An Experimental System for Auditory Image Representations*, IEEE Trans. on Biomedical Engineering, Vol.39, No.2, 1992.
- [10] Ivan Kopecek and Radek Oslejsek, *Hybrid Approach to Sonification of Color Image*. Third 2008 International Conference on Convergence and Hybrid Information Technology, pp.722–pp.727, 2008.
- [11] Jing Su, Alyssa Rosenzweig, Ashvin Goel, Eyal de Lara and Khai N. Truong, *Timbremap: Enabling the Visually-Impaired to Use Maps on Touch Enabled Devices*, International Conference on Mobile Human Computer Interface, pp.17–pp.26, 2010.
- [12] Jeremi Sudol, Orang Dialameh, Chuck Blanchard, Tim Dorcey, *LookTel-A Comprehensive Platform for Computer-Aided Visual Assistance*, IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), pp.73–pp.80, 2010.
- [13] Tsuneo Kagawa, Hideki Kudo, Shuichi Tanoue, Hiro Kiyosue, Hiromu Mori, Hiroaki Nishino, Kouichi Utsunomiya, *A Supporting Method of Medical Imaging Diagnosis with Sonification*, The 6th International Conference on Complex, Intelligent and Software Intensive Systems (CISIS-2012), pp.699–704, 2012.