

Phonoscape：機械学習による記録写真から情景を再現する立体音響像

善甫 啓一，逸村 裕，松野 渉，真柴 雄一

研究背景

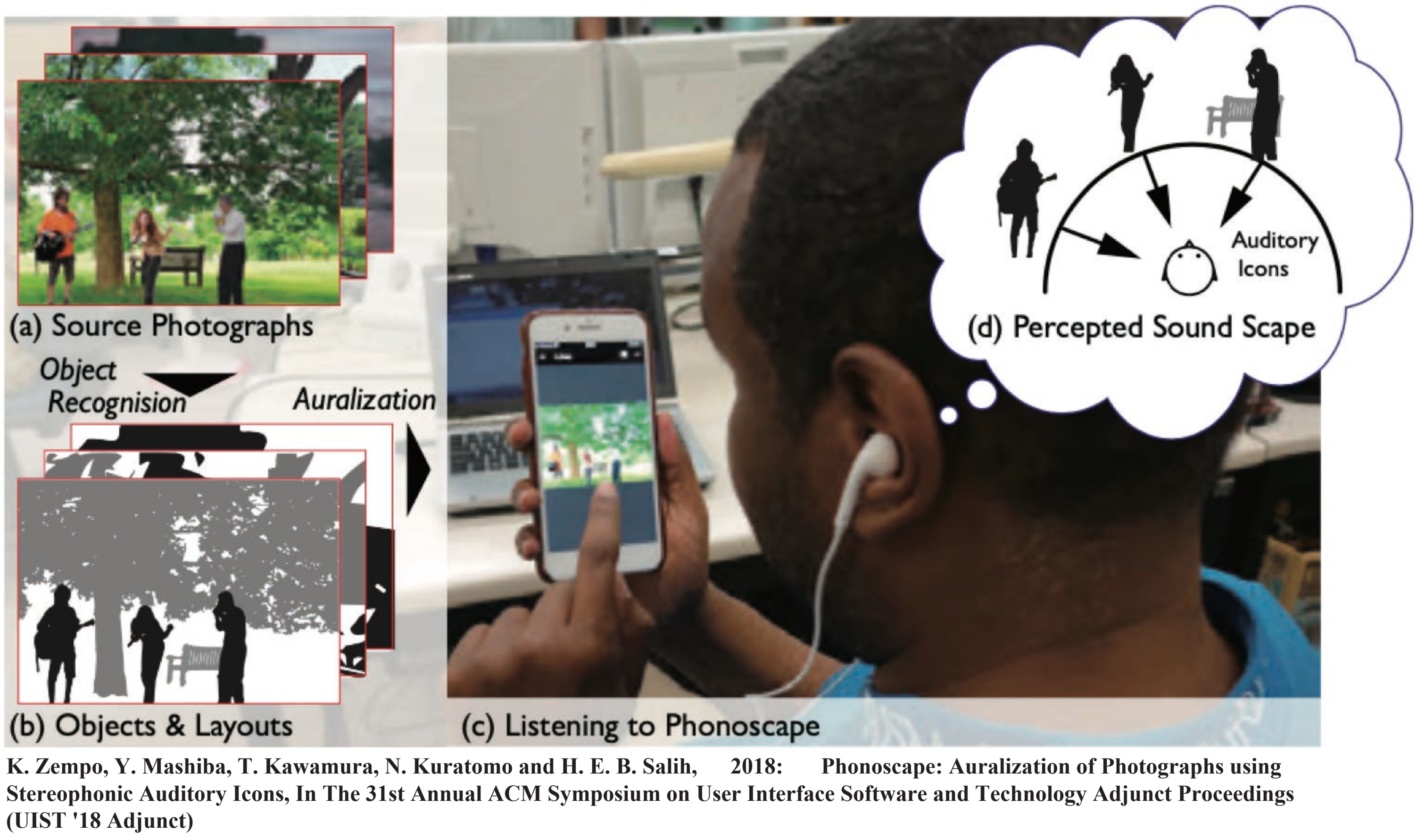
視覚障害者向けの画像可聴化手法は主にテキストベース
→ 晴眼者が画像を見て得られるような直感的な理解を
視覚障害者が得るのは困難

目的

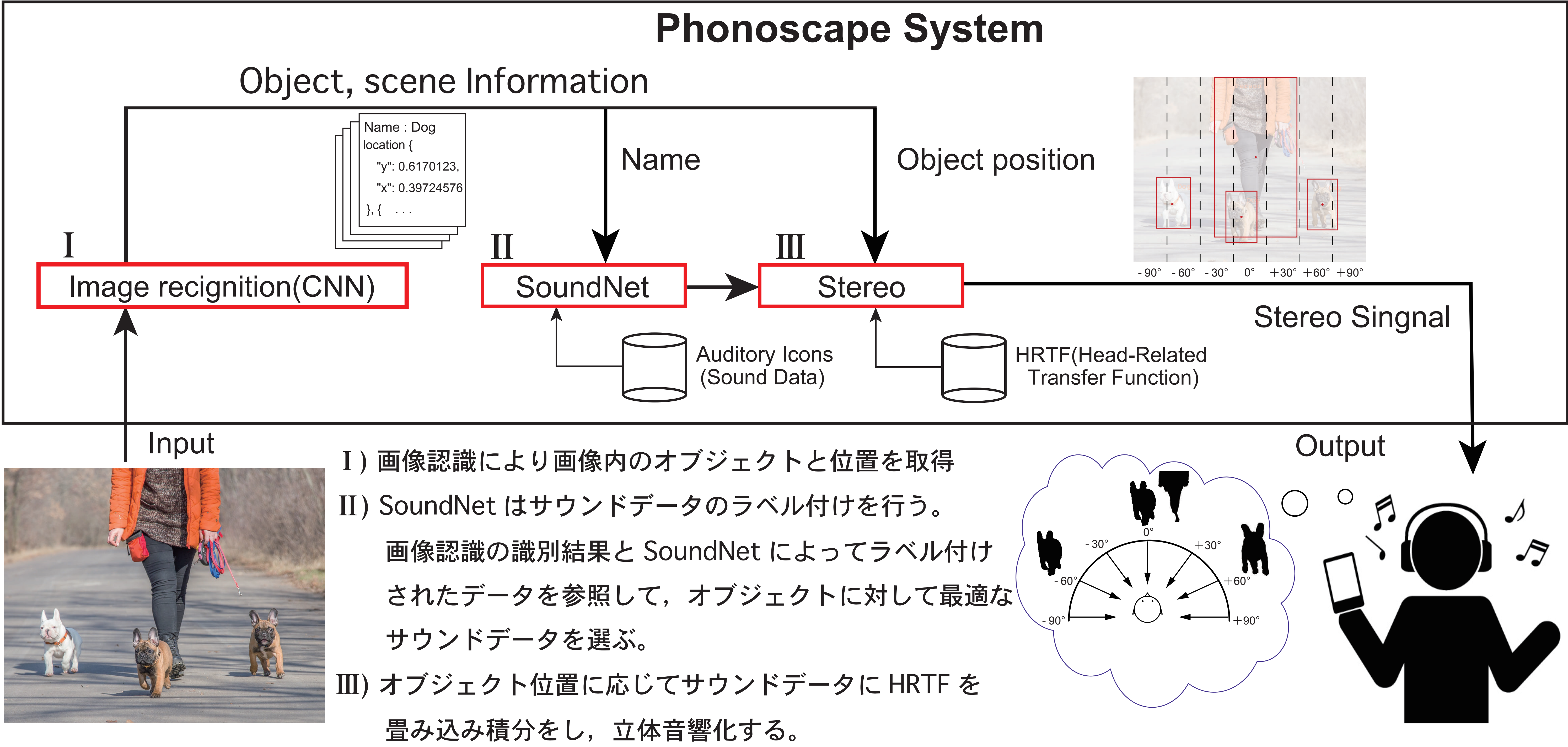
画像から得られる直感的な視覚理解を聴覚で可能とする
可聴化システムを構築する

提案手法

画像に写っているオブジェクトを Auditory icon（犬の鳴き声や、車の走行音などオブジェクトを表現するサウンド）で可聴化する。Auditory icon はオブジェクトの画像内の位置に応じて立体音響化される。ユーザーはサウンドを聞くことで画像内のオブジェクトやその位置を直感的に理解できることを目標としている。

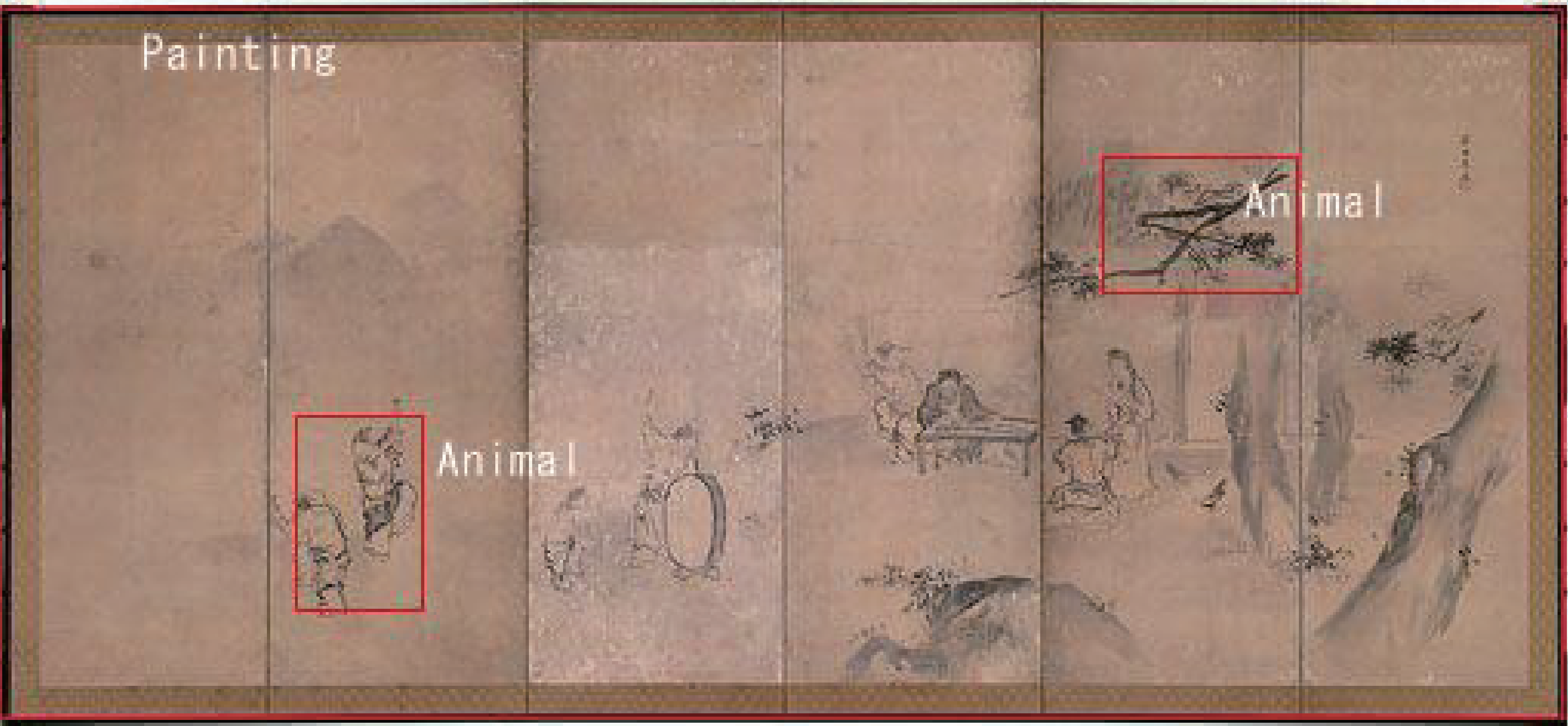


K. Zempo, Y. Mashiba, T. Kawamura, N. Kuratomo and H. E. B. Salih, 2018: Phonoscape: Auralization of Photographs using Stereophonic Auditory Icons, In The 31st Annual ACM Symposium on User Interface Software and Technology Adjunct Proceedings (UIST '18 Adjunct)



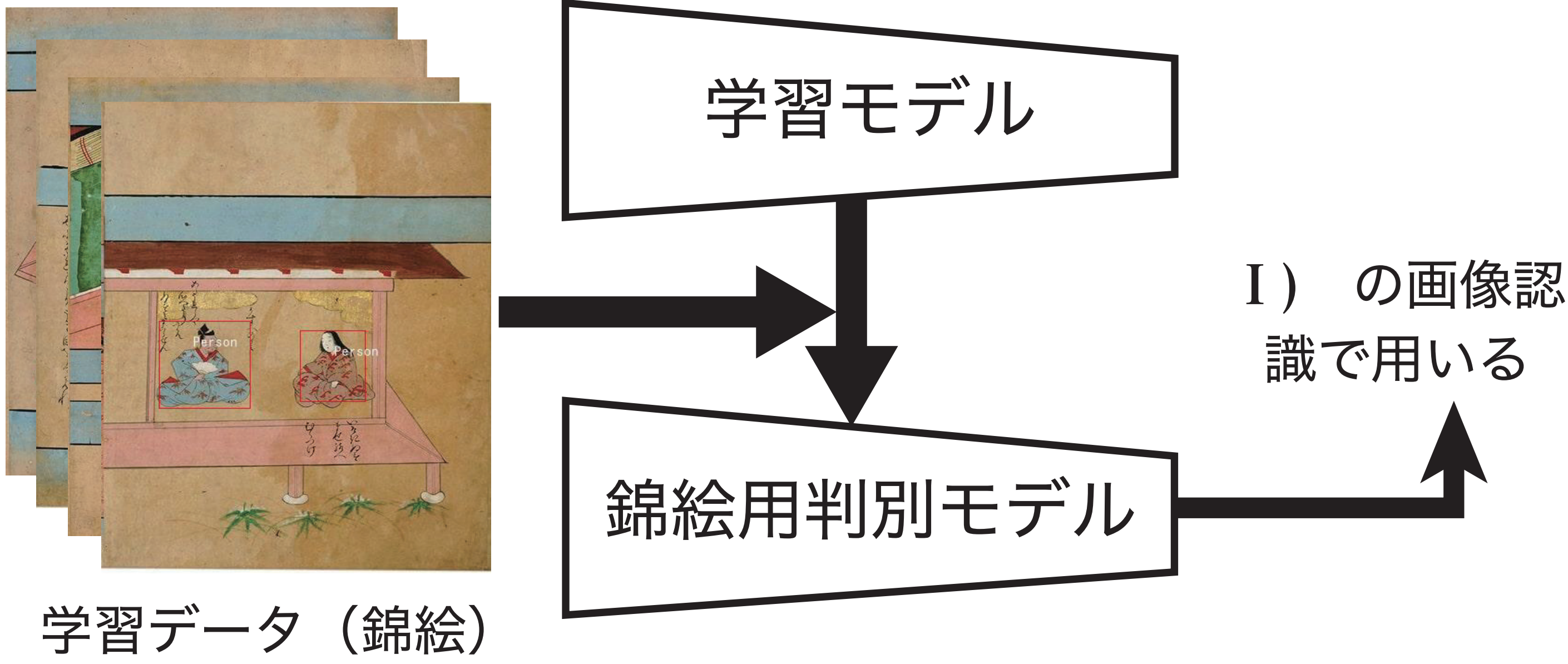
これまでの取り組み

錦絵など図書館資料の可聴化の検討



現状では画像認識の学習データ改良や資料に合わせたサウンド（錦絵をはじめとする「和風な絵画」であることを想起させる音源）のデータセットが不可欠となる。そのため附属図書館から資料の提供を受けた。

Fine - turning



従来の画像認識では自然画像を学習データとしているため、錦絵に代表される図書館資料のオブジェクトを十分に識別できない。そのためが既存の学習モデルを再利用して、新しいモデルを構築する。これにより学習データを減らすことができる。