



EditScribe: Non-Visual Image Editing with Natural Language Verification Loops

TECHNOLOGY NUMBER: 2024-606



OVERVIEW

EditScribe enables blind and low vision (BLV) users to independently edit and verify images through natural language—eliminating barriers of traditional vision-based tools.

- Users describe editing actions (e.g., blur, remove, modify objects) in everyday language and receive AI-generated verification feedback.
- Addresses a major accessibility gap, opening creative markets and digital workflows to millions who are currently underserved by visual editing tools.

BACKGROUND

Image editing underpins communication, creativity, and online engagement worldwide, but remains visually gated—leaving BLV individuals reliant on others or unable to fully participate. Modern editing software lacks non-visual, iterative control and feedback, creating persistent accessibility barriers especially as social content, digital publishing, and visual commerce expand. Trends show more BLV users engaging with digital media, and the need is growing for tools supporting independent visual creation—a market largely untapped, despite increasing advances in AI and accessibility research. Technology that removes visual barriers to editing offers opportunities to reach new creator and consumer segments, reduce inequities, and set new standards for inclusive digital workflows.

Technology ID

2024-606

Category

Software

Software & Content

Accessible Technologies/Blind

Accessibility

Inventor

Anhong Guo

Further information

Ashwathi Iyer

ashwathi@umich.edu

[View online](#)



INNOVATION

EditScribe uses large multimodal models to convert user language into image edits, and then back into structured textual feedback for evaluation. Users receive an initial summary and object-level descriptions of the image, specify desired edits through open-ended prompts, and verify results using a feedback loop of comparative summaries, AI judgement, and refreshed descriptions. This conversational editing and confirmation process allows BLV users to iteratively refine both edits and their understanding, without needing to see the image itself. The approach is novel compared to current caption-based accessibility and relies on recent AI advances to provide detailed, actionable feedback—enabling true authorship and control for BLV users, evidenced by successful user trials. EditScribe thus repositions accessibility as interactive, independent creation rather than passive consumption.

ADDITIONAL INFORMATION

REFERENCES:

["EditScribe: Non-Visual Image Editing with Natural Language Verification Loops"](#)