

Abstract: Image analysis for digital surrogates of historical motion picture film, Y1

For almost a century, celluloid-based imagery was the dominant medium for recording the history of the world, creating a global library of still and moving images the full epistemological weight of which has yet to be felt. Even though still image digitization has reached a mature state, the archival digitization of motion picture film is still in a developmental phase because motion picture film is derived from a complex system and results in data intensive files. But given the rise of deep fake technology it is essential that mature systems are developed soon.

Moving Image Research Collections (MIRC) at the University of South Carolina recently entered into a partnership with the United States Marine Corps History Division to preserve, digitize and make accessible the legacy 16mm and 35mm film collection housed at Marine Corps University, Quantico. The collection is very large, containing over 18,000 cans of film (a typical can contains 7 to 8 minutes of footage). MIRC has been scanning films at 2K (2048 x 1531 pixels) and currently has over 2,000 digitized films, typically scanning 60 to 75 cans per week.

Digitizing the films is only one component of the project. This collection has a high research value for historians of many types and is of sufficient size to create a data set able to train image analysis algorithms. MIRC seeks to identify new methods for deploying these digital film assets as trusted historical resources (in contrast to the chaos of user-contributed online video). To accomplish this, MIRC is partnering with the university's Computer Vision Lab (led by Dr. Song Wang) and its Research Computing unit to develop and deploy three initial projects: Machine Learning (ML) algorithms for identifying and tracking textual information in historical imagery; ML algorithms for facial recognition in historical imagery; and a new method for certifying the chain of historical provenance from a celluloid film to a master digital surrogate copy, and then to all subsequent copies derived from that master.

MIRC seeks support to build a virtual home that not only hosts the online collections for the public, but also allows researchers and developers to collaborate with others and experiment the mechanisms for above mentioned image/video analysis projects.