

ReCoDi

Refinable Colorization of Historic Imagery via Conditional Diffusion Models

Programm / Ausschreibung	KS 24/26, Bridge 2024, Bridge Ausschreibung 2024/01	Status	laufend
Projektstart	01.09.2024	Projektende	31.08.2027
Zeitraum	2024 - 2027	Projektlaufzeit	36 Monate
Keywords	Image Colorization; Historic Images; Machine Learning; Conditional Diffusion; Multi-modal Learning; Vision-Language Models		

Projektbeschreibung

Filmarchive und Museen auf der ganzen Welt verfügen über große Mengen an historischen Videos und Bildern, die nur als Grauwert- oder sogar Schwarz-Weiß-Dokumente vorliegen. Während die Digitalisierung und in einigen Fällen auch die digitale Restaurierung diese Kulturgüter vor dem Verfall bewahrt, ist die Vermarktung dieser historischen Dokumente aufgrund ihres archaischen Aussehens in der Öffentlichkeit nur schwer möglich.

Techniken zur Kolorierung machen die Darstellungen historischer Ereignisse visuell ansprechend, sodass sie für ein größeres Publikum interessant sind. Die Einfärbung historischer Bilder ist jedoch nach wie vor ein sehr arbeits- und kostenintensiver Prozess, der eine breitere Anwendung bis heute verhindert.

In den letzten fünf Jahren wurden viele Bildbearbeitungstechniken durch den Einsatz von generativen Modellen, die auf Deep Learning Ansätzen basieren, revolutioniert. Trotz dieses bemerkenswerten Fortschritts in der Forschung ist die praktische Anwendbarkeit solcher Modelle in realen Anwendungen jedoch immer noch begrenzt, insbesondere wenn man die hohen Anforderungen an Qualität und Präzision der professionellen Filmindustrie berücksichtigt. In einem früheren Projekt (FFG BRIDGE RE:Color) wurden die Wurzeln für die heute erste kommerziell verfügbare Softwarelösung für die hochwertige Videokolorierung gelegt. Das Projekt ermöglichte es dem Industriepartner HSArt, zum führenden Anbieter in diesem aufstrebenden Marktsegment zu werden.

Das beantragte ReCoDi Projekt befasst sich mit der Lösung des bisher aufwändigsten Schritts im Kolorierungsprozess historischer Filme, nämlich der Einfärbung einzelner Bilder unter Beibehaltung der historischen Authentizität. Dieses Projekt zielt darauf ab, modernste Deep-Learning-basierte Bildgenerierungsansätze, neuartige Konditionierungsstrategien für diese generativen Modelle und eine effiziente Benutzerkontrolle in einem einheitlichen Optimierungsframework zu kombinieren. Das resultierende Technologiekonzept wird sowohl interaktive als auch automatisierte Kolorierungstechniken kombinieren, um eine qualitativ hochwertige und authentische Einfärbung historischer Kulturgüter zu ermöglichen.

Die Hauptinnovation von ReCoDi liegt in der Kombination von vorhandenem Vorwissen (basierend auf historisch authentischen Farbschemata, textuellen Szenenbeschreibungen, manuell kolorierten Bildern, sowie Referenzbildern aus ähnlichen landschaftlichen Umgebungen), einer semantischen Beschreibung der Bildinformation, dem Repräsentationsvermögen aktueller Foundation Models, und innovativen Benutzerinteraktionstechniken. Diese multimodalen Beschreibungen der Bildinformation und Erwartungen der KolorierungsexpertInnen werden in einem konditionierten

Optimierungsframework vereinigt, um so eine effiziente und vollständig benutzergesteuerte Kolorierung zu erzielen. Mit ReCoDi kann eine hohe Effizienzsteigerung erzielt werden, wodurch in Zukunft eine Vielzahl an hochqualitativen und authentisch kolorierten historischen Bilddokumenten einem großen Publikum zugänglich gemacht werden können.

Abstract

Film archives and museums all over the world contain large amounts of historic videos and imagery, only available as either monochrome or even black-and-white documents. While digitization and in some cases digital restoration preserves these cultural assets from decay, marketing of these historic documents is difficult due to limited public acceptance caused by their archaic look. Colorization techniques make the depictions of historic events visually appealing such that they are interesting to a larger audience. However, colorizing historic imagery remains a highly labor- and cost-intensive process, which prevents its more widespread use to this day.

Within the past 5 years, deep learning-based generative models have revolutionized many image editing techniques. Despite this remarkable academic research progress, however, the practical usability of such models in real applications is still limited, especially if we consider the high quality and precision requirements imposed by the professional film industry. In a previous project (BRIDGE RE:Color) project, the roots for today's first commercially available software solution specialized on video color propagation have been implemented. The project enabled the industrial partner HSART to become the leading player in this emerging market segment.

The proposed ReCoDi project addresses the most costly step in the colorization process of historic films, namely the colorization of single keyframes respecting authentic historic looks. Such keyframes are the basis for any further color propagation tools. This project aims to combine state-of-the-art deep learning-based image generation approaches, novel conditioning strategies for these generative models, and efficient user control into a unified optimization framework. This will result in a significantly improved colorization workflow that combines both interactive and automated colorization techniques, allowing high quality colorization of historical cultural assets.

The main innovation of ReCoDi lies in exploiting the combination of existing prior knowledge (based on proven historically accurate color schemes, manually colorized exemplary imagery, as well as reference color images from similar environments), semantic image understanding, the representational power of emerging vision-language models, and innovative user interaction techniques, all bundled into a single workflow that allows for efficient and fully user-controlled colorization. The expected increase in efficiency makes the success of such a colorization solution on the market highly probable and will enable the large-scale reuse of visual depictions of historic events, thus making them appealing and easily accessible to a wide audience.

Projektkoordinator

- Technische Universität Graz Institute of Visual Computing (IVC)

Projektpartner

- HS-ART Digital Service GmbH