Week 8: Video Editing With AI

Nate Bachmeier

TIM-7211: Research Design Methodologies

February 14, 2021

Northcentral University

# Video Editing with AI

Movie and television studios spend millions of dollars annually removing objects from videos (Shimamura et al., 2020). This painstaking process requires humans to transpose pixels between different shots and then offset any differences, such as variations of lighting or color pallets. Businesses require a mechanism to automate this challenging process. This constructive research seeks to produce an algorithm that serves as that mechanism.

# Problem Statement

Production studios enjoy shooting video in the real world because it creates an authentic feeling and saves money. However, the camera also records licensed content, pedestrians, and other distracting objects within these shots. Video Effects (VFX) teams clean up the shots by carefully transposing pixels between different frames (Trinh et al., 2019). When the scene contains fast-paced action, a compounding of effort is required. For instance, VFX teams need to account for different lighting, object orientation, and non-static positioning.

Today, VFX teams meet these requirements through painstaking manual processes that require 3-D modeling and sophisticated software projects. These programs are complex and come with a high-barrier to entry, which in turn increases costs. Businesses would prefer to have artificial intelligence solutions manage these tedious tasks. This capability would reduce costs, enable faster time-to-market, and free-up VFX resources for value-differentiating services.

The most powerful artificial intelligence applications use machines to enhance human capabilities rather than replace them (Heer, 2019; Boire, 2017). For instance, a person can write a more profound business case than a machine; however, the same machine will have fewer misspellings and grammatical errors. This dichotomy exists because humans specialize in contextualizing thought versus automation uses patterns to make predictions (Schleer et al., 2019).

# Methodological Approach

There are four approaches to studying a business use-case or phenomena (see Table 1). Software engineers are builders, and we learn best through hands-on experimentation, making constructive research the most appealing selection. Constructive design is also one of the most common research methods for information systems and technology (Silvestrini et al., 2012). These studies identify a problem, build solution artifacts, and communicate the implementation’s unique value (Hevner et al., 2004). Typically, these results (artifacts) originate from specific Proof-Of-Concept (POCs) or directed case-studies.

Table 1: Research Strategies

|  |  |  |
| --- | --- | --- |
| Approach | Description | Study Example |
| Quantitative | Studies the magnitude of a phenomena | Measure the time required for object removal tasks |
| Qualitative | Explores a concept without a numerical basis | Exploration of reasons objects require removal |
| Mixed-Method | Combines exploration and studying the magnitude of these issues | What preparation steps reduce the costs of object removal |
| Constructive | Produce artifacts to study a scenario | Create an algorithm for removing objects |

## Artifacts Production

A critical challenge with a constructive design is properly scoping the problem. If the research is too broad, then it will not complete on time. This research project seeks to produce a reusable library of video transforms powered by artificial intelligence. However, building something entirely from scratch is out-of-scope. Instead, these capabilities must extend mainstream open-source packages like Open Computer Vision (OpenCV) or similar frameworks.

The study will also produce a comprehensive survey of the current state-of-the-art approaches. For instance, Generative Adversal Networks (GAN) and Auto Encoders (AE) use machine learning to produce deep fake images. Can these sophisticated systems learn to identify commonly removed objects and “auto clean” the video? Alternatively, are there algorithms that accelerate the processing times of more traditional manual strategies?

## Addressing the Problem

The initial proof-of-concept needs to provide capabilities that identify and remove objects from video files. These features will reduce the efforts of VFX teams and allow them to focus on value-differentiating tasks. When a project requires fewer resources, it translates into lower costs for businesses and faster time-to-market. All of these benefits directly address the core problem statement.

## Measuring Success

There need to be benchmarks that measure the effectiveness of the proposed solution. For this particular use-case, the Key Performance Indicator (KPI) is the amount of human time required to clean-up a shot. The most direct way of obtaining this value is with VFX experts. They will need to process scenes under traditional methods and AI-enhanced strategies. Hence, the differences in time and effort is quantifiable.

# Ethical Issues