



# Video Editing Tools with Artificial Intelligence



# কৃত্রিম বুদ্ধিমত্তা সহ ভিডিও সম্পাদনা সরঞ্জাম

হাচিক ইয়াzadzhiyan

টেকনিক্যাল ইউনিভার্সিটি অফ সোফিয়া,  
ব্রাঞ্চ প্লোভডিভ  
কম্পিউটার সিস্টেম এবং প্রযুক্তি বিভাগ

বিমূর্ত . ভিডিও সম্পাদনা একটি খুব ক্লাসিক কাজ হতে পারে, তাই এতে অবাক হওয়ার কিছু নেই যে কৃত্রিম বুদ্ধিমত্তা ক্রমবর্ধমানভাবে কর্মপ্রবাহকে প্রবাহিত করতে বা ক্লাসিক কাজগুলি স্বয়ংক্রিয়ভাবে ব্যবহার করা হচ্ছে। যাইহোক, গবেষণা সাহিত্যে কী বুদ্ধিমান ভিডিও সম্পাদনা সরঞ্জাম বিদ্যমান এবং ভিডিও সম্পাদকদের স্বয়ংক্রিয়তা প্রয়োজন তার একটি সংক্ষিপ্ত বিবরণ পাওয়া খুব কঠিন। তাই, আমরা গবেষণায় স্মার্ট ভিডিও এডিটিং টুলের ক্ষেত্র চিহ্নিত করেছি এবং পেশাদার ভিডিও সম্পাদকদের মতামত জরিপ করছি। সত্যিকারের বুদ্ধিমান ভিডিও প্রক্রিয়াকরণ সরঞ্জামগুলির প্রতি সম্ভাব্যতা এবং বর্তমান প্রযুক্তিগত সীমাবদ্ধতাগুলি কী কী তা সনাক্ত করার অভিপ্রায়ে আমরা কৃত্রিম বুদ্ধিমত্তা গবেষণায় শিল্পের বর্তমান অবস্থার সংক্ষিপ্তসারও করেছি। ফলাফলগুলি বুদ্ধিমান ভিডিও সম্পাদনা সরঞ্জামগুলির ক্ষেত্রের বোঝার ক্ষেত্রে অবদান রাখে, অটোমেশন, গবেষণার জন্য অর্পণ প্রয়োজনীয়তাগুলিকে হাইলাইট করে এবং বুদ্ধিমান ভিডিও সম্পাদনা সরঞ্জামগুলিতে আরও গবেষণার জন্য সাধারণ পরামর্শ প্রদান করে।

কীওয়ার্ড: এআই, ভিডিও এডিটিং, অ্যাডোব প্রিমিয়ার, *DaVinci* সমাধান

## I. ভূমিকা

ভিডিও ইন্টারনেটে সবচেয়ে জনপ্রিয় বিষয়বস্তু। ভিজুয়াল নেটওয়ার্ক সূচক অনুসারে, 2023 সালে ইন্টারনেট ট্র্যাফিকের 82.5% ভিডিও সামগ্রী ছিল। মোবাইল ফোন, ভিডিও শেয়ারিং এবং সোশ্যাল মিডিয়া প্ল্যাটফর্মগুলি ভিডিও তোলা এবং পোস্ট করার আগের চেয়ে সহজ এবং দ্রুত করে তোলে। যাইহোক, এই ভিডিওগুলি সম্পাদনা করতে এখনও অনেক সময় লাগে। ভিডিও সম্পাদনা করার জন্য একটি যন্ত্রণা রয়ে গেছে কারণ এটি অডিও এবং ইমেজ উভয়ের সাথে একটি দ্বি-ট্র্যাক মাধ্যম ছাড়াও পৃথক ফ্রেমে কাজ করা প্রয়োজন। ভিডিও এডিটিং সহজ করার জন্য বিভিন্ন প্রচেষ্টা রয়েছে। একটি পদ্ধতি হল কৃত্রিম বুদ্ধিমত্তা (AI) ব্যবহার করে সম্পাদনা প্রক্রিয়া স্বয়ংক্রিয় করা।

এখানে আমরা ভিডিও এডিটিং অটোমেশনের শিল্পের রাজ্যে আগ্রহী, বিশেষত বর্তমান এআই প্রযুক্তির সাথে যা কাঙ্ক্ষিত এবং কী অর্জন করা যায় তার মধ্যে অমিলের উপর ফোকাস করছি।

ওকুন এট আল। [Okun et al., 2015] একটি একক সম্পাদিত চলচ্চিত্র তৈরি করতে এক বা একাধিক উত্সের অংশগুলিকে কাটা এবং যুক্ত করার কাজ হিসাবে ভিডিও সম্পাদনাকে সংজ্ঞায়িত করে। ভিডিও সম্পাদনা সরঞ্জামগুলিকে মোটামুটিভাবে (কম্পিউটার) প্রোগ্রাম হিসাবে সংজ্ঞায়িত করা যেতে পারে যা লোকেরা ভিডিও সম্পাদনার কাজ সম্পাদন করতে ব্যবহার করতে পারে, যেমন। ভিডিও বিভাগ একত্রিত করা। ভিডিও এডিটিং হল এমন একটি ক্ষেত্র যেখানে AI ব্যবহার করা হচ্ছে মানুষের ভিডিও এডিটরদের কাজগুলিকে স্বয়ংক্রিয় বা বৃদ্ধি করতে।

স্মার্ট ভিডিও এডিটিং টুল ভিডিও এডিটিং সহজ করার জন্য ডিজিটাল ভিডিও এডিটিং এর শুরু থেকেই চেষ্টা করে আসছে। বুদ্ধিমান ভিডিও সম্পাদনা সরঞ্জামগুলির একটি মূল বিষয় হল ফ্রেমের পরিবর্তে ফ্রেম এবং ডায়ালগের মতো উচ্চ-স্তরের বিমূর্ততা থেকে ভিডিও ম্যানিপুলেশনের অনুমতি দেওয়ার সমস্যা। এই ধরনের একটি টুলের একটি প্রাথমিক উদাহরণ হল 2002 থেকে সিলভার [Casares et al., 2002], যা ভিডিও ক্লিপগুলির বুদ্ধিমান নির্বাচনের পাশাপাশি ভিডিও মেটাডেটা ব্যবহার করে ভিডিও সম্পাদনার বিমূর্ত দৃশ্য প্রদান করে। একটি স্মার্ট ভিডিও এডিটিং টুলের একটি সাম্প্রতিক উদাহরণ হল Roughcut [Leake et al., 2017]। রাফকাট ব্যবহারকারী-প্রবেশকৃত দৃশ্যের সংলাপ, কাঁচা ফুটেজ এবং এডিটিং ইন্ডিয়াম ব্যবহার করে সংলাপ-চালিত দৃশ্যের জন্য গণনামূলক সম্পাদনা সক্ষম করে। একটি ওপেন সোর্স টুল আছে অটোএডিট [Passarelli, 2019] এবং গবেষণা [Berthouzou, 2012] যা ভিডিওগুলির সাথে টেক্সট ট্রান্সক্রিপ্ট লিঙ্ক করে ভিডিও সাক্ষাত্কারের পাঠ্য সম্পাদনা করার অনুমতি দেয়।

সম্পূর্ণরূপে AI-নিয়ন্ত্রিত ভিডিও উৎপাদন সম্প্রতি গবেষণার আগ্রহকে আকর্ষণ করেছে [Xue et al.; হুয়া এট আল., 2004]। বর্তমানে, এআই-নিয়ন্ত্রিত ভিডিও উৎপাদন স্বয়ংক্রিয় ভিডিও সারাংশ বা মিশ্রণ তৈরির উপর দৃষ্টি নিবদ্ধ করে। এই সম্পূর্ণ স্বয়ংক্রিয় ভিডিও সম্পাদনা পদ্ধতিগুলি, যেমন ভিডিও সারাংশ এবং যৌগিক অ্যানালিসিস তৈরি করতে ব্যবহৃত হয়, বুদ্ধিমান ভিডিও সম্পাদনা সরঞ্জাম হিসাবে বিবেচিত হয় না, কারণ এগুলি সাধারণ অ্যালগরিদম যা একটি খুব সংকীর্ণ এবং নির্দিষ্ট প্রশ্ন সম্পাদন করে যার জন্য বুদ্ধিমত্তা বা ব্যবহারকারীর মিথস্ক্রিয়া প্রয়োজন হয় না।

চিত্র প্রক্রিয়াকরণ, দৃষ্টি এবং প্রাকৃতিক ভাষা প্রক্রিয়াকরণে কৃত্রিম বুদ্ধিমত্তার অগ্রগতি হয়েছে

ভিডিও সম্পাদনা। কিন্তু ভিডিও এডিটিং-এর কড়াকড়ি থেকে সরে আসার স্বপ্ন কি সত্যি হয়েছে? উত্তর, অবশ্যই, আমরা কার স্বপ্নের কথা বলছি তার উপর নির্ভর করে। কিভাবে স্বয়ংক্রিয় ভিডিও সম্পাদনার অগ্রগতি মূল্যায়ন করা যেতে পারে? স্বয়ংক্রিয় ভিডিও সম্পাদনার অগ্রগতি মূল্যায়নের এই প্রশ্নটি দুটি কোণ থেকে যোগাযোগ করা যেতে পারে। প্রথমটি হল সাহিত্য পর্যালোচনা করা, যখন দ্বিতীয়টি হল মানব ভিডিও সম্পাদকদের প্রত্যাশা পরীক্ষা করা এবং কৃত্রিম বুদ্ধিমত্তায় শিল্পের অবস্থার সাথে তাদের তুলনা করা। এই নিবন্ধে, আমরা উভয় করি।

স্মার্ট ভিডিও সম্পাদনা সরঞ্জামগুলি যে প্রধান চ্যালেঞ্জটি সমাধান করার চেষ্টা করেছিল তা হল ব্যবহারকারীদের জন্য ভিডিও সম্পাদনা প্রক্রিয়াকে স্ট্রিমলাইন করা। এটি সাধারণত ফ্রেমে ভিডিও ফ্রেমের মধ্য দিয়ে যাওয়ার ক্লাস্টিকর কাজগুলি দূর করে করা হয়। প্রস্তাবিত সমাধান এবং সরঞ্জামগুলি বিভিন্ন উপায়ে পৃথক হয় যা সমস্যাটির দিকে দৃষ্টিভঙ্গি, উদ্দিষ্ট লক্ষ্য, অন্তর্নিহিত প্রযুক্তি, বিমূর্তকরণের স্তর (গুলি), প্রস্তাবিত মিথস্ক্রিয়াগুলি এবং উক্তিগুলির মিথস্ক্রিয়াগুলির পদ্ধতিগুলি থেকে উদ্ভূত হয়।

আমরা দুটি দৃষ্টিকোণ থেকে অত্যাধুনিক ভিডিও এআই অ্যাপ্লিকেশন পর্যালোচনা করেছি: i) সাধারণ ভিডিও এআই প্রযুক্তি; ii) নির্দিষ্ট এআই ভিডিও এডিটিং প্রযুক্তি। ভিডিওর জন্য কমন এআই প্রযুক্তিতে বিভিন্ন ধরনের ভিডিও কাজ রয়েছে যেমন অবজেক্ট ট্র্যাকিং, অবজেক্ট ডিটেকশন, স্পিচ রিকগনিশন, ভিডিও রিজনিং, অ্যাকশন ডিটেকশন, ভিডিওতে সেন্ট্রমেন্ট ডিটেকশন। ভিডিও সম্পাদনার জন্য নির্দিষ্ট এআই প্রযুক্তি অনেক সংকীর্ণ যেমন ভিডিও স্ক্রিপ্ট, ফ্রেম এবং দৃশ্য প্রক্রিয়াকরণ এবং ভিডিও সম্পাদনা খনির নিয়ম।

আমরা 13 জন ভিডিও সম্পাদকের সমীক্ষা করেছি যাদের ভিডিও সম্পাদনার অভিজ্ঞতা 1 বছর থেকে 22 বছরের মধ্যে। এই সমীক্ষায় তাদের ভিডিও এডিটিং অভিজ্ঞতা, এআই ভিডিও এডিটর সম্পর্কে চিন্তাভাবনা এবং ভিডিও এডিটিং অটোমেশনের প্রয়োজন রয়েছে। ভিডিও সম্পাদনা সরঞ্জামগুলিতে অটোমেশন সম্পর্কিত প্রত্যাশা, প্রয়োজনীয়তা এবং সমস্যাগুলির একটি ওভারভিউ কম্পাইল করার জন্য সমীক্ষার প্রতিক্রিয়াগুলি একটি বিষয়ভিত্তিক বিশ্লেষণ করতে ব্যবহৃত হয়। আমরা সমীক্ষার দ্বারা প্রকাশিত মতামত এবং প্রত্যাশাগুলিকে সামগ্রী তৈরি/কারচুপি, স্বয়ংক্রিয় ভিডিও সম্পাদনা এবং অন্যান্য AI সরঞ্জামগুলির জন্য মেশিন লার্নিংয়ের বর্তমান স্তরের সাথে তুলনা করি। আমরা আলোচনা করি যে কীভাবে AI-তে সাম্প্রতিক ভিডিও নির্মাতাদের জন্য একটি আদর্শ AI ভিডিও সম্পাদনা সরঞ্জাম তৈরি করতে পারে। এই কাগজটি নিম্নরূপ গঠন করা হয়েছে: বিভাগ 2 ভিডিওতে বুদ্ধিমান ভিডিও সম্পাদনা সরঞ্জাম এবং এআই কৌশলগুলির একটি ওভারভিউ প্রদান করে। সাহিত্যে বুদ্ধিমান ভিডিও সম্পাদনার সরঞ্জামগুলি বিভাগ 3-এ তুলনা এবং সংক্ষিপ্ত করা হয়েছে। পদ্ধতি এবং সংক্ষিপ্ত ফলাফল সহ (মানব) ভিডিও সম্পাদকদের সমীক্ষাটি 4 তে রয়েছে। বিভাগ 5-এ, বুদ্ধিমানদের জন্য সমাধানগুলির উপর পূর্ববর্তী কাজগুলি

ভিডিও সম্পাদনা সরঞ্জাম এবং সমীক্ষা থেকে ব্যবহারকারীর প্রত্যাশার বিপরীতে এবং কিছু এআই কৌশল ব্যবহারকারীর প্রত্যাশা পূরণের সম্ভাব্য সমাধান হিসাবে প্রস্তাব করা হয়েছে। পরিশেষে, কাগজের সারাংশ, উপসংহার এবং ভবিষ্যত কাজ সেকশন 6 এ উপস্থাপন করা হয়েছে।

## II. ব্যবহার সহজ

ডিজিটাল ভিডিও প্রবর্তনের পর থেকে ভিডিও সম্পাদনা সহজতর করার জন্য আরও ভাল সরঞ্জাম তৈরি করা সর্বদা একটি গবেষণা বিষয়বস্তু।

প্রথমত, আমরা বুদ্ধিমান এবং স্বয়ংক্রিয় ভিডিও প্রকাশনার সরঞ্জাম তৈরি করার জন্য বিভিন্ন পদ্ধতির পর্যালোচনা করেছি। এরপরে, আমরা ভিডিও এডিটিংয়ে প্রত্যক্ষ বা পরোক্ষভাবে প্রয়োগ করা হয়েছে এমন AI পদ্ধতির দিকে ফিরে যাই।

ভিডিও সম্পাদনা সহজ করার চেষ্টা করা প্রথম প্রকল্পগুলির মধ্যে একটি হল সিলভার [ম্যার্স এট আল।, 2001; লং এট আল।, 2004]। সিলভার প্রকল্পের প্রথম সংস্করণে, বিভিন্ন ধরনের ভিউ রয়েছে, যথা ট্রান্সক্রিপ্ট, টাইমলাইন, প্রিভিউ এবং স্ক্রিপ্ট ভিউ।

এই এডিটিং টুলটি স্মার্ট সিলেকশন, কাট, ডিলিট, কপি, পেস্ট এবং ফটো এবং সিন ডিটেকশন ব্যবহার করে রি-অ্যাটাচ সহ স্মার্ট এডিটিং এক্সপ্লোর করে। মেটাডেটা তৈরি করতে টেক্সট ট্রান্সক্রিপশন, শর্ট বাউন্ডারি ডিটেকশন এবং অপটিক্যাল ক্যারেক্টার রিকগনিশন (OCR) আকারে ভিডিওর মেটাডেটা লেয়ার ব্যবহার করে স্মার্ট এডিটিং তৈরি করা হয়েছে। সিলভার টুলের দ্বিতীয় পুনরাবৃত্তিতে [লং এট আল।, 2004], টুলটি লেন্স প্রয়োগ করেছে (ক্রিপ, ফটো, ফ্রেম) এবং শব্দার্থিক স্কেলিং সহজ কাজগুলি করার জন্য যেমন দুটি ভিডিও সেগমেন্টে যোগ দেওয়ার সময় কাটতে সঠিক ফ্রেমগুলিকে কল্পনা করা। বেশিরভাগ স্মার্ট ভিডিও টুল শুধুমাত্র একটি নির্দিষ্ট ধরনের ভিডিও তৈরি করার জন্য ডিজাইন করা হয়েছে। উদাহরণ স্বরূপ, QuickCut [Truong et al., 2016] তৈরি করা হয়েছিল বর্ণনা করা ভিডিও রচনা করার জন্য, এবং ভিডিও ডাইজেস্ট [Pavel et al., 2014] ভিডিও লেকচারের সারসংক্ষেপের জন্য তৈরি করা হয়েছিল। এর পরে, আমরা সম্পূর্ণ স্বয়ংক্রিয় ভিডিও সম্পাদনার পদ্ধতিগুলি দেখি। স্বয়ংক্রিয় ভিডিও এডিটিং হল কোনো মানব সম্পাদকের কোনো ইনপুট ছাড়াই ভিডিও সেগমেন্টের গণনামূলক প্রক্রিয়াকরণ এবং সংমিশ্রণ। স্বয়ংক্রিয় ভিডিও সম্পাদনা রেকর্ড করা ভিডিওগুলিতে বা, অনেক বড় স্কেলে, ভিডিও সংরক্ষণাগারগুলিতে সঞ্চালিত হতে পারে। স্বয়ংক্রিয় ভিডিও সম্পাদনার প্রাথমিক কাজ ভিডিও সিকোয়েন্স তৈরির জন্য নিয়ম-ভিত্তিক কৌশলগুলির উপর দৃষ্টি নিবদ্ধ করে [বাটলার এবং পার্কস, 1997] বা ভিডিও ডকুমেন্টারিগুলির ডোমেনে ব্যবহারকারীর অনুরোধগুলি নির্বাচন এবং স্বয়ংক্রিয়ভাবে সম্পাদনা করার জন্য একটি শব্দার্থ-ভিত্তিক পদ্ধতি [বোকোনি, 2004]। ম্যাশআপ, একটি একক ইভেন্টের জন্য একাধিক ভিডিওর সংমিশ্রণ, অন্য ধরনের স্বয়ংক্রিয় ভিডিও সম্পাদনা। ভার্সুয়াল ডিরেক্টর নামে কাজ [শ্রেষ্ঠ এট আল।, 2010] কনসার্ট রেকর্ডিংয়ের জন্য মিশ্রণ তৈরি করার জন্য একটি পদ্ধতি তৈরি করেছে

ভিডিও এডিটর এবং ফিল্ম ব্যাকরণ সাহিত্যের সাক্ষাত্কার থেকে প্রাপ্ত নিয়মের উপর ভিত্তি করে যা একটি ভাল কনসার্ট ভিডিও তৈরি করে তা সর্বাধিক করে তোলে। লাইভ ইভেন্ট সম্প্রচার করতে স্বয়ংক্রিয় ভিডিও সম্পাদনাও ব্যবহার করা যেতে পারে। [Radut et al., 2020]-এর কাজ লাইভ ইভেন্টগুলির জন্য শুধুমাত্র প্রোটোটাইপ AI ভিডিও সম্পাদক নয়, AI-সম্পাদিত লাইভ ইভেন্টগুলির গুণমান পরিমাপের জন্য মূল্যায়ন এবং মূল্যায়ন পদ্ধতির আলোচনাও করেছে। এআই স্বয়ংক্রিয়ভাবে সম্পাদিত ভিডিও তৈরি করতেও ব্যবহার করা যেতে পারে। মেশিন দ্বারা তৈরি যখন AI BBC থেকে আর্কাইভের সম্মুখীন হয়, তখন এটি BBC আর্কাইভ থেকে 150টি ছোট জটিলতা তৈরি করে [R&D, 2018]। [তাসকির এট আল।, 2006] স্পিচ রিকগনিশন সিস্টেম থেকে স্পিচ ট্রান্সক্রিপ্ট ব্যবহার করে ভিডিও প্রোগ্রাম স্কিম করার জন্য একটি সমষ্টি পদ্ধতি উপস্থাপন করেছে। [Truong et al., 2016] ভিডিও বিমূর্তকরণ বা সংক্ষিপ্তকরণ পদ্ধতির একটি সারাংশ প্রদান করে, যেমন একটি ভিডিওর সম্বন্ধে তথ্য সরবরাহ করার জন্য একটি সিরিজ কীফ্রেম তৈরি করা বা চিত্রগুলি সরানো। [Wu et al., 2020] দ্বারা কর্পোরেট মিটিংয়ের স্বয়ংক্রিয় ভিডিও সম্পাদনার কাজ মানব-সম্পাদিত ভিডিও থেকে শেখা সম্পাদনার সিদ্ধান্তগুলি ব্যবহার করে এবং অডিও এবং ভিডিও উভয় ক্ষেত্রে দুটি মনোযোগ মডেল ব্যবহার করে। EDL- সম্পাদনা সিদ্ধান্তের তালিকা হল একটি পাঠ্য-ভিত্তিক ভাষা যা ক্লিপ এবং টাইমকোড ডেটার একটি অর্ডারকৃত তালিকা সহ রচনামূলক সিদ্ধান্তগুলিকে এনকোড করে। এটি অনেক স্বয়ংক্রিয় ভিডিও সম্পাদনা সিস্টেমের আউটপুট হিসাবে ব্যবহৃত হয় [তাসকির এট আল।, 2006; উ এট আল।, 2020; Passarelli, 2019], কোন ভিডিও সম্পাদকরা তখন Adobe Premier Pro বা Davinci Resolve-এর মতো সফটওয়্যারে তাদের সম্পাদনা চালিয়ে যেতে EDL-এর সাথে এনকোড করা স্বয়ংক্রিয় সম্পাদনা সমাধানগুলি ব্যবহার করতে পারে।

ভিডিওগুলি থেকে তথ্য বের করতে AI কৌশলগুলি কীভাবে ব্যবহার করা যেতে পারে তা গবেষণার একটি খুব বিচিত্র ক্ষেত্র। আমরা মুখ শনাক্তকরণ, বস্তু সনাক্তকরণ, অবজেক্ট ট্র্যাকিং, দৃশ্য সনাক্তকরণ, অনুভূতি বিশ্লেষণ, ভিডিও মুক্তি এবং ভিডিও ক্যাপশনে বিশেষভাবে আগ্রহী। ফেস রিকগনিশন বলতে বোঝায় যে কোনও ছবিতে মানুষের মুখ উপস্থিত আছে কিনা এবং সম্ভবত যার, যখন বস্তু সনাক্তকরণ হল কোনও ছবিতে কোনও নির্দিষ্ট বস্তুকে শনাক্ত করার সমস্যা। অবজেক্ট ট্র্যাকিং হল একটি ভিডিওর ফ্রেমের মাধ্যমে একটি নির্দিষ্ট বস্তুকে চিহ্নিত করা এবং তার গতিবিধি ট্র্যাক করার সমস্যা। দৃশ্য সনাক্তকরণ বা ভিডিও বিভাজন হল সেগমেন্টগুলির সনাক্তকরণ যা একটি ভিডিওতে শব্দার্থগতভাবে বা দৃশ্যত সম্পর্কিত। সেন্টিমেন্ট বিশ্লেষণ হল মেজাজের সাথে মেলানোর সমস্যা যা কিছু বিষয়বস্তুর দ্বারা প্রকাশ করা হবে: এটি খুশি, দুঃখ, বিদ্ৰূপাল্পক, ইত্যাদি কিনা। ভিডিও ক্যাপশনিং [Wu et al., 2016] হল একটি AI কৌশল যা প্রাকৃতিক বর্ণনা তৈরি করে যা ক্যাপচার করে ভিডিওর গতিশীলতা।

ভিডিও এআই কৌশল রয়েছে যা ভিডিও সম্পাদনার জন্য আরও নির্দিষ্ট। [মাতসুও এট আল।] পুনরুৎপাদনযোগ্য সম্পাদনা নিদর্শন তৈরি করার জন্য ভিডিও থেকে সম্পাদনার ধরণগুলি (আলগা, মার্কারি-আর্টসার্ট ফ্রেম এবং নিয়মগুলির সমন্বয়ে গঠিত) সনাক্ত করার জন্য একটি ডেটা মাইনিং কৌশল উপস্থাপন করেছেন। [Butler and Parkes, 1997] এর আগের কাজ ভিডিও সম্পাদনা স্বয়ংক্রিয় করার জন্য একটি নিয়ম এবং প্রশ্ন-ভিত্তিক পদ্ধতি উপস্থাপন করেছিল। সম্পাদনা প্রক্রিয়ার মডেলিং এবং শব্দার্থবিদ্যা ব্যবহার করে স্বয়ংক্রিয় ভিডিও সম্পাদনা [Nack and Parkes, 1997] এ উপস্থাপন করা হয়েছে।

### III. স্মার্ট ভিডিও এডিটিং টুল

এই অধ্যায়ে, আমরা সাহিত্য পর্যালোচনা করি এবং সংক্ষিপ্তভাবে বর্ণনা করি যে কীভাবে পূর্ববর্তী বুদ্ধিমান ভিডিও সম্পাদনা সরঞ্জামগুলি ভিডিও সম্পাদনাকে একটি সহজ কাজ করার সমস্যা তৈরি করে এবং সমাধান করে। বুদ্ধিমান ভিডিও সম্পাদনা সরঞ্জামগুলির জন্য সাহিত্য অনুসন্ধানগুলি কম্পিউটার বিজ্ঞান সাহিত্য ডাটাবেসে কীওয়ার্ডগুলি (বুদ্ধিমান বা বুদ্ধিমান বা স্বয়ংক্রিয় বা এআই) এবং (ভিডিও সম্পাদক বা ভিডিও সম্পাদনা) ব্যবহার করে সঞ্চালিত হয়েছিল, যা হল DBLP, ACM ডিজিটাল লাইব্রেরি এবং Google স্কলার। তারপরে শিরোনাম এবং বিমূর্তগুলি অন্তর্ভুক্তির মানদণ্ডের উপর ভিত্তি করে পড়া এবং ফিল্টার করা হয়েছিল, যা নির্ধারণ করেছিল যে অন্তর্ভুক্ত কাগজগুলি ব্যবহারকারীদের জন্য একটি ভিডিও সম্পাদনা সরঞ্জাম তৈরি করার জন্য একটি বুদ্ধিমান পদ্ধতির বর্ণনা করা উচিত। অন্তর্ভুক্ত নথিতে ইউজার ইন্টারফেসের বর্ণনা এবং/অথবা বাস্তবায়নও থাকতে হবে। আরও সম্পর্কিত সাহিত্য খুঁজে পেতে অন্তর্ভুক্ত কাগজপত্রের রেফারেন্সগুলিও স্ক্যান করা হয়েছিল। ফলস্বরূপ পেপারগুলিকে সংক্ষিপ্ত করা হয়েছিল এবং তিনটি বিষয়ের মধ্যে গোষ্ঠীভুক্ত করা হয়েছিল, যথা ভিডিও সম্পাদনা কাজ, একটি স্বয়ংক্রিয় সম্পাদকের সাথে মিথস্ক্রিয়া (মানব-কম্পিউটার মিথস্ক্রিয়া), এবং এআই প্রযুক্তি। এই বিভাগে অন্তর্ভুক্ত কাগজপত্রের তালিকা সারণি 1 এ রয়েছে

#### A. ভিডিও সম্পাদনার কাজ - স্মার্ট ভিডিও সম্পাদকদের

এই উপবিভাগটি ভিডিও সম্পাদনা কর্মপ্রবাহে সম্বোধন করা বিভিন্ন কাজের পরিপ্রেক্ষিতে বুদ্ধিমান ভিডিও সম্পাদনা সরঞ্জামগুলির ক্ষেত্র প্রবর্তন করে এবং প্রতিটি কাজের জন্য পদ্ধতির সংক্ষিপ্ত বিবরণ দেয়। ভিডিও সেগমেন্টেশন হল সবচেয়ে সাধারণ কাজ যা স্মার্ট ভিডিও এডিটররা সমাধান করার চেষ্টা করে। এই বিভাগে পর্যালোচনা করা সমস্ত পূর্ববর্তী কাজ ভিডিও বিভাজন পদ্ধতির কিছু ফর্ম ব্যবহার করে, তবে বিভাজন সম্পাদন করতে বিভিন্ন পদ্ধতি ব্যবহার করা হয়। এখন থেকে, একটি ভিডিওকে একটি একক উৎস ফাইল থেকে ভিডিওর একটি অবিচ্ছিন্ন অংশ হিসাবে সংজ্ঞায়িত করা হয়। শট সনাক্তকরণের ব্যবহার চিহ্নিত প্রথম বিভাজন পদ্ধতি।

একটি ফ্রেম হল ওকুন এবং অন্যান্য ছবিগুলির একটি অবিচ্ছিন্ন ধারাবাহিক ক্রম। [2015]। শট সনাক্তকরণ বিভাজন চিত্র বিশ্লেষণ পদ্ধতির সাথে সংশ্লিষ্ট হয় [ক্যাসারেস এট আল।, 2002; লং এট আল।, 2004] এবং শট সনাক্তকরণ বৈশিষ্ট্যগুলি ব্যবহার করে সংশ্লিষ্ট হয় যা ক্যামেরার গতি, উজ্জ্বলতা এবং সময়কাল [শিপম্যান, 2008]। [Wu et al., 2015] ব্যবহারকারী-উত্পাদিত ভিডিওগুলিকে ভাগ করতে ফ্রেম এবং সাবফ্রেম ব্যবহার করে (একটি সাবফ্রেমকে ভিডিওর একটি মৌলিক ইউনিট হিসাবে সংজ্ঞায়িত করা হয় যাতে সামঞ্জস্যপূর্ণ ক্যামেরা চলাচল এবং স্বয়ংসম্পূর্ণ শব্দার্থবিদ্যা রয়েছে)। [Casares et al., 2002] এছাড়াও L-কার্টের ক্ষেত্রে বিভিন্ন স্থানে ভিডিও এবং অডিওর বিভাজন বিবেচনা করে।

দ্বিতীয় বিভাজন পদ্ধতিটি ভিডিও অডিওর বিষয়বস্তু বা অর্থের উপর নির্ভর করে এমন কাট তৈরি করতে একটি পার্থক্য প্রতিলিপির সাথে সিঙ্ক্রোনাইজেশন ব্যবহার করে। [লিক এট আল।, 2017] ভিডিওটি সেগমেন্ট করার জন্য দৃশ্যের ডায়ালগ স্ক্রিপ্ট থেকে পূর্ব-লিখিত লাইন ব্যবহার করেছে, যখন [পাভেল এট আল।, 2014] টেক্সট ট্রান্সক্রিপশনটি বেয়েসিয়ান টপিক সেগমেন্টেশনের মাধ্যমে চালিয়েছে [আইজেনস্টাইন এবং বারজিলে, 2008] সনাক্তকরণের জন্য BSec তথ্যমূলক ভিডিওর বিভাগ এবং উপধারা।

অন্যান্য বিভাজন পদ্ধতির মধ্যে রয়েছে দৃষ্টিভঙ্গি ব্যবহার করা [Kimura et al., 2005], ব্যবহারকারী-চিহ্নিত বিন্দু থেকে শুরু করে, সেই পয়েন্টগুলির সাথে ফ্রেমের মিলের পরিমাপ [Chi et al., 2013], এবং কথা বলার হেড ইন্টারভিউ হ্রাসের জন্য ফিটনেস মূল্যায়ন। ভিডিওগুলি [বার্তুসো, 2012]। [Cattelan et al., 2008] দ্বারা ব্যবহারকারী-উত্পাদিত ভিডিও সারাংশ ব্যবহারকারীর দেখার ক্রিয়া এবং মন্তব্য ব্যবহার করে বিভাজন তৈরি করেছে। ভিডিও সেগমেন্টেশন পরবর্তী কাজ নিয়ে আলোচনা করার জন্যও অপরিহার্য, যা হল ভিডিও সেগমেন্ট রচনা করা।

আমাদের AI ভিডিও এডিটিং টুলের তালিকায় ভিডিও সেগমেন্ট কম্প্যাজ করা হল দ্বিতীয় সবচেয়ে সাধারণ কাজ। ভিডিও সেগমেন্টের সংশ্লিষ্টকে স্ট্রিমলাইন করার সবচেয়ে সাধারণ পদ্ধতি হল ইন-সিন ডায়ালগ [লিক এট আল।, 2017] বা টেক্সট ট্রান্সক্রিপশনের ব্যবহার [বার্তুসো, 2012; Wu et al., 2015] রচনার শুরু বিন্দু হিসাবে। ভিডিও কথোপকথন অবশ্যই লিখতে হবে এবং ইনপুট হিসাবে সরবরাহ করতে হবে, তবে স্পিচ রিকগনিশন প্রযুক্তি ব্যবহার করে একটি পার্থক্য প্রতিলিপি তৈরি করা যেতে পারে।

উদাহরণস্বরূপ, [Truong et al., 2016] ম্যানুয়ালি তৈরি করা টেক্সট ট্রান্সক্রিপ্টের পরিবর্তে বর্ণিত অডিও বা ভয়েস-ওভার থেকে রূপান্তরিত একটি পার্থক্য প্রতিলিপি ব্যবহার করেছে।

Roughcut[Leake et al., 2017]-এ, অটোমেশন ব্যবহার করে প্রতিটি নিজ নিজ সংলাপ এবং স্পিকারের জন্য ভিডিও সেগমেন্ট তৈরি করা হয়েছে। সংলাপের ক্রম প্রদত্ত স্ক্রিপ্ট অনুসরণ করে। যাইহোক, সৃজনশীল রচনাটি ব্যবহারকারী দ্বারা পরিবর্তন করা যেতে পারে যিনি ভিডিও সম্পাদনা বাগধারাগুলির সংমিশ্রণ চয়ন করেন। একইভাবে, সম্পাদক দ্বারা তৈরি গল্পের স্ক্রিমা অংশগুলি রচনা করতে ব্যবহৃত হয়েছিল [Truong et al., 2016]। ভিডিও রেকর্ডিং [Berthouzoz, 2012] থেকে সংশ্লিষ্ট পার্থক্য নির্বাচন করে একটি ইন্টারভিউ থেকে নির্দিষ্ট বাক্যাংশ বা পুনরাবৃত্তি শব্দের মতো অবাঞ্ছিত অংশগুলি কেটে সেগমেন্টেশন করা হয়।

[Chi et al., 2013] ভিডিও ক্লিপগুলিতে ব্যবহারকারী দ্বারা সরবরাহ করা ট্যাগগুলিকে বিভাগগুলি সংগঠিত এবং রচনা করার উপায় হিসাবে ব্যবহার করে এবং সম্পাদকরা একটি ডেমো ভিডিওর ধাপগুলির সাথে সামঞ্জস্যপূর্ণ ট্যাগগুলিকে সাজিয়ে সামগ্রিক সেগমেন্টের রচনা পরিবর্তন করতে পারেন। ব্যবহারকারী-উত্পাদিত ভিডিও সারাংশ তৈরির উদ্দেশ্যে রচনাটি দৃষ্টিভঙ্গি [কিমুরা এট আল।, 2005] এবং [ক্যাটেলান এট আল।, 2008]-এ দেখার কার্যকলাপ এবং ব্যবহারকারীর মন্তব্যগুলির মাধ্যমে মডেল করা দর্শকের অভিপ্রেয় ব্যবহার করে করা হয়।

টাইমলাইন এবং ভিডিওগুলিকে ভিজুয়লাইজ করা বিমূর্ততার বিভিন্ন স্তরে টাইমলাইন এবং ভিডিও দেখার আকারে আসে এবং বিকল্প উপায়ে ভিডিও টাইমলাইনের বিভিন্ন উপস্থাপনা যেমন পার্থক্য। ভিডিও রেন্ডারিং বিমূর্ততা ফ্রেম, ফ্রেম এবং ক্লিপ আকারে হতে পারে।

একটি ফ্রেম হল একটি ভিডিওর একটি স্থির চিত্র, যখন একটি ফ্রেম হল একটি অবিচ্ছিন্ন চিত্রের ক্রম [Okun et al., 2015]। ভিডিও অ্যাবস্ট্রাকশনের বিভিন্ন রূপের একটি ওভারভিউ [ট্রুং এবং ভেক্টেশ-এ পাওয়া যায়। একটি প্রতিনিধি ফ্রেম ব্যবহার করে ক্লিপগুলির ভিজুয়লাইজেশনকে টাইমলাইনে ভিডিও বিষয়বস্তুর দ্রুত বিচারের অনুমতি দেওয়ার একটি পদ্ধতি হিসাবে আলোচনা করা হয়েছে [লং এট আল।, 2004]।

[ক্যাসারেস এট আল।, 2002] বিভিন্ন স্তরের বিমূর্ততায় টাইমলাইন ভিজুয়লাইজেশন অফার করে, যেগুলি হল স্ক্রিপ্ট, সম্পাদনাযোগ্য প্রতিলিপি এবং টাইমলাইন ভিউ। টাইমলাইন ভিজুয়লাইজেশনের দ্বিতীয় পদ্ধতি হল টেক্সচুয়াল ট্রান্সক্রিপ্টের পরিপ্রেক্ষিতে টাইমলাইনের প্রতিনিধিত্ব করা [ক্যাসারেস এট আল।, 2002; ট্রুং এট আল।, 2016; বার্তুসো, 2012; পাভেল এট আল।, 2014]। পূর্ববর্তী অনুচ্ছেদগুলিতে উল্লিখিত হিসাবে, ভিডিও ফ্রেমের প্রকৃত রচনায় পরিবর্তন করতে টাইমলাইনের পার্থক্য উপস্থাপনা কখনও কখনও শব্দ স্তরে ম্যানিপুলেট করা যেতে পারে।

বুদ্ধিমান ভিডিও ম্যানিপুলেশন শুধুমাত্র দুই বুদ্ধিমান ভিডিও এডিটর আলোচনা করা হয়। প্রথম কাজ [Casares et al., 2002] স্মার্ট সিলেকশন, ক্লিপিং, ক্লিপিং, পেস্ট এবং রি-পিপিং ব্যবহার করা হয়েছে। এই সমস্ত ক্রিয়াগুলি চিত্র বিশ্লেষণের সাথে ছবির সীমানা ব্যবহার করে সঞ্চালিত হয়। দ্বিতীয় কাজটি সহজভাবে ভিডিও ট্রান্সক্রিপ্ট ব্যবহার করে বুদ্ধিমান নির্বাচন বা ভিডিও সেগমেন্টের বুদ্ধিমান কাটিং প্রদান করে [Berthouzoz, 2012]। এই কাজের অনেক উদাহরণের অভাব এবং নীচে উল্লিখিত কাজগুলি এই কারণে হতে পারে যে সমস্ত স্মার্ট ভিডিও এডিটিং সরঞ্জামগুলি ধারণার প্রমাণ, এইভাবে এই অত্যন্ত গুরুত্বপূর্ণ কিন্তু অ-প্রয়োজনীয় বৈশিষ্ট্যগুলির অভাব রয়েছে।

রূপান্তর তৈরি করুন। ট্রানজিশন তৈরির সহজ পদ্ধতি দুটি পৃথক কাজের মধ্যে কভার করা হয়েছে। [Truong et al., 2016] ট্রানজিট কাজগুলিকে গতিশীল প্রোগ্রামিং হিসাবে প্রণয়ন করে একটি নান্দনিকভাবে আনন্দদায়ক রূপান্তর স্বয়ংক্রিয় করার একটি পদ্ধতি তৈরি করেছে যেখানে জাম্পের মতো খারাপ ট্রানজিশন পয়েন্টগুলিকে শাস্তি দেওয়া হয়। [Berthouzoz, 2012] ফ্রেমের অনুক্রমিক ক্লাস্টারিং ব্যবহার করে এবং লেনদেন পয়েন্ট হিসাবে ফ্রেমের মধ্যে সংক্ষিপ্ততম পথ খুঁজে বের করে লুকানো রূপান্তর তৈরি করতে একটি ভিন্ন পদ্ধতি ব্যবহার করে।

ভিডিও রেকর্ড করুন। [Truong et al., 2016] ক্যাপচার প্রক্রিয়া চলাকালীন ভিডিও নিবন্ধন করে অডিও টীকা সহ ভিডিও ক্লিপ নিবন্ধন করার জন্য একটি নতুন পদ্ধতি উপস্থাপন করেছে। তাদের কাজে, রেকর্ডিং পর্যালোচনার সময় ট্যাগ দিয়ে লগিং ছাড়াও অডিও দ্বারা লগিং করা যেতে পারে।

### B. অটোমেশনের সাথে মিথস্ক্রিয়া

এই বিভাগে, মিথস্ক্রিয়া পদ্ধতির পাশাপাশি ভিডিও বিমূর্তকরণের স্তর [ড্রুং এবং ভেক্টেশ] সংক্ষিপ্ত করা হবে। আমরা পরীক্ষা করে দেখেছি বেশিরভাগ স্মার্ট ভিডিও এডিটিং টুলে ব্যবহৃত ইন্টারঅ্যাকশনের প্রাথমিক মোড হল একটি কীবোর্ড এবং মাউস সহ একটি গ্রাফিক্যাল ইউজার ইন্টারফেস (GUI)। একটি ব্যতিক্রম হল [Kimura et al., 2005] থেকে একটি কেস, যা দৃষ্টি-ভিত্তিক মিথস্ক্রিয়া তদন্ত করে। যাইহোক, ব্যবহারকারীদের নিয়ন্ত্রণের বিমূর্ততা এবং গ্রানুলারিটির মাত্রা বিভিন্ন টুল জুড়ে পরিবর্তিত হয়। কোনো বিমূর্ততা ছাড়াই ভিডিও এডিটিং টুলে, এডিটিং অবশ্যই স্বতন্ত্র ফ্রেম লেভেলে করতে হবে, যা খুবই শ্রম-নিবিড়। যাইহোক, কিছু স্মার্ট ভিডিও এডিটিং টুল বিমূর্ততার একাধিক স্তরে ভিডিও ম্যানিপুলেশন অফার করে। একাধিক বিমূর্ততা প্রদানকারী সরঞ্জামগুলির দুটি উদাহরণ হল সিলভার [ক্যাসারেস এট আল।, 2002] এবং কুইককাট [ড্রুং এট আল।, 2016]। সিলভার যা

তিনটি বিমূর্ততা অফার করে যা হল: ক্লিপ, শট এবং ফ্রেম। Quickcut কথ্য শব্দ এবং ফ্রেমের পরিপ্রেক্ষিতে বিমূর্ততা প্রদান করে।

কিছু ভিডিও এডিটিং টুল বিমূর্ততার খুব উচ্চ স্তরে কাজ করে। DemoCut [Chi et al., 2013]-এ, উদাহরণস্বরূপ, ভিডিও সম্পাদনা করা হয় সেই ধাপগুলির জন্য বিমূর্ত পদক্ষেপ এবং মার্কার দ্বারা। একইভাবে, RoughCut [Leake et al., 2017]-এ, ব্যবহারকারী একটি সংলাপ স্ক্রিপ্ট সংলাপ লাইন ব্যবহার করে এবং বাগধারার আকারে সিদ্ধান্তগুলি সম্পাদনা করে টাইমলাইন ম্যানিপুলেট করতে পারে। উচ্চ স্তরে ম্যানিপুলেশন, তবে, ফ্রেম স্তরে সূক্ষ্ম সমন্বয় করার ক্ষমতার মূল্যে আসে। যাইহোক, তিনটি বুদ্ধিমান ভিডিও এডিটিং টুলে [লিক এট আল।, 2017; ড্রুং এট আল।, 2016; Passarelli, 2019] ভিডিও সম্পাদনার কাজটি EDL (এডিট বর্ণনা ভাষা) হিসাবে রপ্তানি করা যেতে পারে, যা ফ্রেম-স্তরের সংশোধন করতে এবং ভিডিওতে সম্পাদনা প্রক্রিয়া সম্পূর্ণ করতে অন্যান্য বাণিজ্যিক ভিডিও সম্পাদনা সফটওয়্যারের সাথে ব্যবহার করা যেতে পারে।

### C. AI প্রযুক্তি ব্যবহার করা হয়

ভিডিও বিভাজন। বুদ্ধিমান ভিডিও সম্পাদনা সরঞ্জামগুলির পূর্ববর্তী কাজগুলি ফ্রেমের সীমানা সনাক্ত করতে এবং প্রতিটি ফ্রেমের জন্য প্রতিনিধি ফ্রেমগুলি খুঁজে পেতে ইমেজ বিশ্লেষণের উপর নির্ভর করত [Casares et al., 2002] অথবা চিত্র বিশ্লেষণের সংমিশ্রণ ব্যবহার করে, [Shipman, 2008]-এ পটভূমি জ্ঞান এবং মডেল ম্যাচিং তৈরি করুন।

ফ্রেম সনাক্তকরণ নিয়মগুলি লক্ষ্যযুক্ত ভিডিও প্রকারের জন্য হাতে তৈরি করা হয়। [Casares et al., 2002]-এ, হাতে তৈরি ট্রান্সক্রিপ্টগুলি বক্তৃতা শনাক্তকরণ ব্যবহার করে ভিডিওগুলির সাথে সারিবদ্ধ করা হয়েছিল। থিম্যাটিকভাবে সুসংগত ইউনিটে ভিডিও লেকচারের বিভাজন [পাভেল এট আল।, 2014]-এ ভিডিওটির পাঠ্য প্রতিলিপিতে বিষয় বিভাজন দ্বারা সঞ্চালিত হয়। অডিও টীকা সহ বিভাজনের আরেকটি রূপ [Truong et al., 2016] এ অনুসন্ধান করা হয়েছে। এটি শব্দার্থগতভাবে প্রাসঙ্গিক বিভাগে অডিও টীকা দ্বারা গতি-ভিত্তিক বিভাজন এবং পরিমার্জন ব্যবহার করে কাজ করে। ভিডিওতে ক্রমাগত গতি শনাক্ত করে মোশন-ভিত্তিক সেগমেন্টেশন করা হয়, যখন শব্দার্থিক সেগমেন্টেশন ভিডিওর ক্রিয়া বা বিষয়ের সাথে মিলে যায়। কম্পিউটেশনাল কৌশলগুলি ব্যবহার করে ইমেজিংয়ের নীতিগুলির একটি ভূমিকা [উ এট আল।, 2015] এবং [লিক এট আল।, 2017] এ পাওয়া যেতে পারে। ভিডিও ক্লিপিং পয়েন্ট, নির্বাচিত ভিডিও এবং নির্বাচিত অডিও খণ্ডগুলি সনাক্ত করার জন্য ডোমেন-নির্দিষ্ট নীতিগুলি সাক্ষাতকারের মাধ্যমে নির্বাচন করা হয় এবং অস্ট্রাইজেশন সমস্যা হিসাবে উপস্থাপন করা হয় [Wu et al., 2015]। ইন [লিক এট আল।, 2017], 12

মৌলিক ফিল্ম এডিটিং ইন্ডিয়ামগুলি (জাম্প এডানো, আবেগ বর্ধন, ইত্যাদি) বেশিষ্ট্য প্যারামিটারের পরিপ্রেক্ষিতে উপস্থাপন করা হয় যা সম্পাদনা সমাধান তৈরি করতে একটি লুকানো মার্কাভ মডেলে ইনপুট হিসাবে প্রবেশ করে।

একটি লুকানো মার্কাভ মডেল (এইচএমএম) হল মডেলিং সিকোয়েন্সের একটি পরিসংখ্যানগত পদ্ধতি যেখানে অভ্যন্তরীণ অবস্থার সিরিজ লুকানো থাকে। HMM-এ ব্যবহৃত বৈশিষ্ট্যগুলির মধ্যে রয়েছে স্পিচ-টু-টেক্সট, ফেস রিকগনিশন এবং ক্লিপ থেকে স্ট্রাকচারাল তথ্য ব্যবহার করে তৈরি করা লেবেল।

#### IV. একজন স্মার্ট ভিডিও এডিটিং সহকারীর কাছ থেকে সম্পাদকরা কী চান

এই বিভাগে, আমরা একটি আদর্শ AI ভিডিও সম্পাদক কী গঠন করে সে সম্পর্কে (মানব) ভিডিও সম্পাদকদের মতামত অন্বেষণ করার জন্য আমরা যে সমীক্ষাটি করেছি তার প্রতিবেদন করি।

অধ্যয়ন পদ্ধতি সাধারণ তথ্য.

আমাদের সমীক্ষায় অংশগ্রহণকারীদের মধ্যে ভিডিও সম্পাদনার গড় অভিজ্ঞতা ছিল 9.75 বছর, যার মধ্যে সবচেয়ে কম সময় 1 বছর থেকে দীর্ঘতম 22 বছর। ভিডিও অংশগ্রহণকারীদের সাথে কাজ করার ধরনের পরিপ্রেক্ষিতে, প্রতিটি অংশগ্রহণকারী প্রায় 3 ধরনের ভিডিও তালিকাভুক্ত করেছে। বিজ্ঞাপন, তথ্যচিত্র, উপস্থাপনা, খেলাধুলা, সোশ্যাল মিডিয়া এবং সংবাদ ভিডিওগুলি সবচেয়ে সাধারণ ধরনের ভিডিও।

সফটওয়্যার প্রোগ্রামের পরিপ্রেক্ষিতে, অংশগ্রহণকারীরা গড়ে 5টি ভিডিও সম্পাদনা প্রোগ্রাম ব্যবহার করেছেন। সর্বাধিক উল্লিখিত সম্পাদনা প্রোগ্রামগুলি হল অ্যাডোবি প্রিমিয়ার প্রো এবং ডাভিঞ্চি রেজলুশন। এছাড়াও, কম পরিচিত প্রোগ্রাম যেমন Avid, Flimora 9 এবং VizStory এর পাশাপাশি Rev.com এবং Descript.com এর মতো ভিডিও পরিষেবাগুলিও উল্লেখ করা হয়েছে। যখন জিজ্ঞাসা করা হয়েছিল যে তারা ভিডিও সম্পাদনার প্রসঙ্গে কোন AI প্রযুক্তির কথা শুনেছে, 13 জনের মধ্যে 8 জন উত্তরদাতা ব্র্যান্ডেড এবং বাণিজ্যিক অফার যেমন Adobe Premier Pro CC এবং Magisto-এ AI প্রযুক্তির সাথে উত্তর দিয়েছেন। অবশিষ্ট 5 জন উত্তরদাতাদের জন্য, উত্তরগুলি AI কৌশলগুলিকে কভার করেছে যা হল স্বয়ংক্রিয় সংশোধন, শব্দ ট্রাস, পটভূমি অপসারণ, ভিডিও স্থিতিশীলতা, অবজেক্ট সনাক্তকরণ/ইমেজ টাঁকা, বিভাজন, জুমিং, গভীর মিথ্যা, স্বয়ংক্রিয় ভিডিও সংকেত, মুখের স্বীকৃতি এবং পার্শ্বের সাথে বক্তৃতা।

নিখুঁত এআই সম্পাদক। "আপনি নিখুঁত AI ভিডিও সম্পাদনা সরঞ্জামটি কী হতে চান?" প্রশ্নের উত্তর উত্তর একে অপরের থেকে ব্যাপকভাবে পরিবর্তিত হয়। যাইহোক, আমরা প্রতিক্রিয়াগুলিতে পাঁচটি থিম চিহ্নিত করেছি। এগুলি হল ভিডিও এডিটিং কাজের টুল হিসেবে এআই, প্রজেক্টের টুল হিসেবে এআই

ব্যবস্থাপনার কাজ, নান্দনিক গুণমান উন্নত করার জন্য স্বয়ংক্রিয় সরঞ্জাম, মানুষের এআই সমস্যা এবং বিষয়বস্তু আবিষ্কারের জন্য এআই।

ভিডিও এডিটিং কাজের জন্য একটি টুল হিসেবে AI হল সবচেয়ে বড় শ্রেণী যেটা বেশিরভাগ প্রতিক্রিয়ার মধ্যে পড়ে। এতে কীওয়ার্ড রয়েছে যা শট রিকগনিশন, ভিডিও কম্পোজিটিং, ডায়ালগ লাইনের উপর ভিত্তি করে খারাপ ভিডিও ফিল্টার করা, গান এবং সাবটাইটেল সিঙ্ক করা, অনুবাদ এবং ভাষা বোঝা। প্রজেক্ট ম্যানেজমেন্ট এআই বিষয়ের মধ্যে ভিডিও মেটাডেটা তৈরি, ডেটা ম্যানেজমেন্ট এবং গ্রহণের শর্ত রয়েছে। পরবর্তী বিষয় হল নান্দনিক গুণমান উন্নত করার জন্য AI, যা স্বয়ংক্রিয় রঙের গ্রেডিং এবং স্বয়ংক্রিয় অডিও সমতা কভার করে। এর পাশাপাশি, এআই সম্পর্কিত মানব-এআই উদ্বেগ রয়েছে, যেমন নিয়ন্ত্রণ এবং অটোমেশনের ভারসাম্য, ব্যবহারকারী-কেন্দ্রিক এআই এবং ব্যক্তিগতকরণ। অবশেষে, বিষয়বস্তু আবিষ্কার এআই বিষয়ের শর্তাবলীর মধ্যে টাইমলাইনে বিদ্যমান বিষয়বস্তুর উপর ভিত্তি করে স্টক ভিডিও এবং স্টক মিউজিক ভিডিও সাজেস্ট করা অন্তর্ভুক্ত।

এআই ভিডিও এডিটরের সাথে কীভাবে ইন্টারঅ্যাক্ট করতে হয় তা জিজ্ঞাসা করা হলে, বেশিরভাগ প্রতিক্রিয়া উল্লেখ করেছে যে তারা ভয়েসের মাধ্যমে ইন্টারঅ্যাক্ট করতে চায়, তারপরে যারা কীবোর্ড এবং মাউস GUI এর মাধ্যমে ইন্টারঅ্যাক্ট করতে চায়। এই দুটি প্রধান মিথস্ক্রিয়া মোড ছাড়াও, অন্যান্য বিভিন্ন মোড যেমন স্পর্শ ইন্টারফেস, অঙ্গভঙ্গি, মস্তিষ্কের কম্পিউটার ইন্টারফেস বেশ কয়েকবার উল্লেখ করা হয়েছে। কিছু উত্তর যোগাযোগের জন্য প্রয়োজনীয় হিসাবে প্রকল্পের অবস্থার উপর ভিত্তি করে প্রাসঙ্গিক কমান্ড উল্লেখ করেছে, সেইসাথে একটি এআই অফ বোতাম যা অটোমেশনকে সহজেই বন্ধ করার অনুমতি দেয়।

নিখুঁত এআই সম্পাদকের বিষয়ে শেষ প্রশ্নটি বিমূর্তকরণ (মানব) সম্পাদকদের সাথে কাজ করতে চান এমন স্তর সম্পর্কে জিজ্ঞাসা করে। বেশিরভাগ উত্তরদাতারা বলেছেন যে তারা একটি নিখুঁত AI ভিডিও সম্পাদনা সরঞ্জামের তাদের দৃষ্টিভঙ্গিতে কীফ্রেম স্তরে ভিডিওটি পরিচালনা করতে চান। বিমূর্তকরণের দ্বিতীয় এবং তৃতীয় সর্বাধিক জনপ্রিয় স্তর হল ক্লিপ এবং ফ্রেম। অন্যান্য ধরনের বিমূর্ততা যা একবার উল্লেখ করা হয়েছে তা হল ক্রম, ইতিহাস এবং ফ্রেম। দুইজন উত্তরদাতা উল্লেখ করেছেন যে তারা একটি নমনীয় ধরনের বিমূর্ততা চান যেখানে তারা মিথস্ক্রিয়ার অন্তর্নিহিত নীতিগুলির জন্য বিমূর্তকরণের স্তরকে সামঞ্জস্য করতে পারে। এআই এবং ওয়ার্কফ্লো। অধ্যয়নের দ্বিতীয় অংশটি নিম্নলিখিত বিষয়গুলি অন্বেষণ করে: ভিডিও সম্পাদনা কর্মপ্রবাহের কাজগুলি যা অংশগ্রহণকারীরা স্বয়ংক্রিয় করতে চায়, এবং এই কাজের জন্য স্বায়ত্তশাসনের স্তর এবং মিথস্ক্রিয়া মোড সম্পর্কিত সমস্যাগুলি। ওয়ার্কফ্লো বিভাগে প্রশ্নগুলির উত্তরগুলি চারটি বিষয়ভিত্তিক কাজ নিয়ে গঠিত: ভিডিও সম্পাদনা কার্য, নান্দনিক উন্নতি, ভিডিও প্রাক-সম্পাদনা কার্য, এবং সাজেশন টাস্ক।

ভিডিও সম্পাদনার কাজগুলি বর্ণনা করতে ব্যবহৃত সর্বাধিক জনপ্রিয় কীওয়ার্ডগুলি ছিল বিভাজন এবং সাবটাইটেলিং, যা তিনটি প্রতিক্রিয়াতে উল্লেখ করা হয়েছিল। ভিডিও এডিটিং কাজের জন্য দ্বিতীয় জনপ্রিয় কীওয়ার্ড হল ভিডিও সেগমেন্টেশন এবং খারাপ ফ্রেম ফিল্টারিং (যা দুবার উল্লেখ করা হয়েছে)। এছাড়াও, নিম্নলিখিত ভিডিও সম্পাদনা কাজগুলি একবার উল্লেখ করা হয়েছে: বিষয়বস্তু বিশ্লেষণ, ভিডিও সংরক্ষণাগার, মুখ সনাক্তকরণ, রূপ বসানো, রূপান্তর ফ্রেম নির্বাচন, অডিও-টু-ভিডিও সিঙ্ক্রোনাইজেশন, এবং দুই-চ্যানেল অডিও এবং ভিডিও নির্বাচন।

উত্তরগুলির পরবর্তী ঘন ঘন উল্লেখিত বিষয়ের কাজ যা আমরা স্বয়ংক্রিয় করতে চাই তা হল নান্দনিক মানের উন্নতি। এগুলোর জন্য সবচেয়ে জনপ্রিয় পদ হল কালার কারেকশন এবং অডিও ইকুয়ালাইজেশন। উপরন্তু, চাক্ষুষ উন্নতি, পটভূমি অপসারণ, এবং তোতলা অপসারণের মতো কাজগুলি প্রতিটি একবার উল্লেখ করা হয়েছিল। প্রাক-সম্পাদনার কাজগুলির জন্য, ভিডিও রেকর্ডিং দুবার উল্লেখ করা হয়েছে এবং স্বয়ংক্রিয় টাইমকোড তৈরির কথা একবার উল্লেখ করা হয়েছে। পরামর্শমূলক কাজের পরিপ্রেক্ষিতে, উত্তরগুলির মধ্যে সম্পাদনা শৈলী, সাধারণ সহায়তা এবং সঙ্গীত পরামর্শ সহ একটি ক্লিপ পরামর্শ অন্তর্ভুক্ত রয়েছে।

চূড়ান্ত প্রশ্ন হল সম্পাদকরা তাদের ব্যবহার করা সরঞ্জামগুলির তুলনায় সম্পাদককে এআই হতে কতটা ভাল আশা করে। এই প্রশ্নে, চারজন উত্তরদাতা বলেছেন যে তারা এটিকে খুব অনুরূপ বা পরিচিত হতে চান। দুটি প্রতিক্রিয়া বলেছে যে তারা একটি মৌলিক স্তরের মিল থাকতে চায়। দুটি উত্তরে পাওয়া "ব্যবহারে সহজ" কীওয়ার্ডটি আরেকটি শব্দ যা "পরিচিত" এর অনুরূপ অর্থ হতে পারে। একজন উত্তরদাতা AI ভিডিও সম্পাদনা সরঞ্জামগুলিকে বিদ্যমান সরঞ্জামগুলিতে সংহত করার জন্য একটি প্লাগ-ইন পদ্ধতির উল্লেখ করেছেন। শুধুমাত্র দুই জন উত্তরদাতা বলেছেন যে তারা আশা করছেন ভবিষ্যতে এআই ভিডিও এডিটর খুব আলাদা হবে বা একেবারেই হবে না।

## V. এআই ভিডিও এডিটিং টুলের জন্য চ্যালেঞ্জ

ভিডিও সম্পাদনার কাজ সাহিত্যে চিহ্নিত এবং আমাদের গবেষণার ফলাফলে উল্লেখ করা ভিডিও সম্পাদনার কাজগুলির মধ্যে যথেষ্ট ওভারল্যাপ রয়েছে। এখানে আমরা ভিডিও সম্পাদনা কর্মপ্রবাহের অনাবিষ্কৃত অংশগুলিতে ফোকাস করি। প্রথমটি হল বিভিন্ন ট্র্যাক থেকে অডিও এবং ভিডিও সিঙ্ক করা। এই কাজটি ইতিমধ্যেই স্বয়ংক্রিয় ভিডিও মিশ্রণ তৈরির একটি ভিন্ন কিন্তু সম্পর্কিত প্রসঙ্গে তদন্ত করা হয়েছে [Wu et al., 2015; শ্রেষ্ঠা এট আল।, 2010]। ভিডিওতে খারাপ ফ্রেম বা খারাপ অংশগুলিকে ফিল্টার করার বিষয়ে অধ্যয়ন করা হয়নি, তবে ভিডিও সম্পাদকরা যখন বলেছিল যে তারা খারাপ ফ্রেম ছিল তখন তারা কী বোঝাতে চেয়েছিল তা বোঝার জন্য এটি সচেতন গবেষণার মাধ্যমে করা উচিত। অবশেষে, আমরা যেমন ভাষা সমস্যা আছে

ভিডিও সম্পাদনা করার সময় স্বয়ংক্রিয় অনুবাদ, সাবটাইটেল এবং ভাষা বোঝা। সাবটাইটেলিং এবং স্বয়ংক্রিয় অনুবাদ সিস্টেম মেশিন অটোমেশন ব্যবহার করে স্বয়ংক্রিয় হতে পারে। যাইহোক, ভিডিও সম্পাদনার প্রেক্ষাপটে ভাষা বোঝার জন্য প্রাকৃতিক ভাষা প্রক্রিয়াকরণে অগ্রগতি এবং ভিডিও সম্পাদনার পরিপ্রেক্ষিতে ভিডিও সম্পাদনার শর্তাবলী এবং ভাষা ব্যবহারের বোঝার প্রয়োজন।

ভিডিও লগিং এবং মেটাডেটা তৈরি ভিডিও লগিং হল একটি ভিডিও রেকর্ডিং দেখা এবং টাইমকোড ব্যবহার করে এর বিষয়বস্তু ট্যাগ করা। এটিকে বুদ্ধিমান ভিডিও সম্পাদনা সাহিত্যে এবং আমাদের গবেষণায় ব্যবহারকারীরা স্বয়ংক্রিয় করতে চান এমন একটি কাজ হিসাবে চিহ্নিত করা হয়েছে। সাহিত্যে বর্তমান ভিডিও রেকর্ডিং কৌশলগুলি নির্দিষ্ট অ্যাপ্লিকেশনের মধ্যে সীমাবদ্ধ; যথা ডেমো ভিডিও এবং সংলাপ ভিত্তিক ভিডিও। ভিডিওকে টেক্সটে রূপান্তর করতে এবং টেক্সট প্রসেসিং কৌশল প্রয়োগ করতে স্পিচ রিকগনিশনের ভিডিও রেকর্ডিং এবং মেটাডেটা তৈরিতে অনেক সম্ভাব্য ব্যবহারের ক্ষেত্রে রয়েছে। AI এর আরেকটি আকর্ষণীয় ক্ষেত্র যা দেখার জন্য তা হল ভিজুয়াল এবং ভাষাগত উভয় ইনপুট থেকে ভিডিও যুক্তি এবং বোঝার সংমিশ্রণ [Wu et al., 2016]।

ভিডিও সম্পাদনার জন্য ভয়েস-ভিত্তিক মিথস্ক্রিয়া মিথস্ক্রিয়া করার সবচেয়ে সাধারণ মোড যা লোকেরা বলে যে তারা একটি AI ভিডিও সম্পাদনা সরঞ্জামের সাথে যোগাযোগ করতে ব্যবহার করতে চায়। ভিডিও এডিটিং টুলে ভয়েস ইউজার ইন্টারফেসের সম্ভাব্যতা অন্বেষণ করা হয়নি। [চ্যাং এট আল।, 2019] নির্দেশমূলক ভিডিও নেভিগেট করার জন্য ভয়েস-ভিত্তিক ইন্টারঅ্যাকশনের ডিজাইন স্পেস অন্বেষণ উপস্থাপন করেছে। যেহেতু ভিডিও এডিটিংয়ে ভয়েস ইন্টারঅ্যাকশন সম্পূর্ণরূপে অজানা অঞ্চল, তাই শুরুর পয়েন্টটি ডিজাইন গবেষণা হওয়া উচিত। আরেকটি সম্ভাবনা হল ইন্টারঅ্যাকশনের পরিবর্তে প্রসঙ্গ-মুক্ত একক ভয়েস কমান্ড অধ্যয়ন করা। যে কাজগুলি ব্যবহারকারী সম্পূর্ণরূপে স্বয়ংক্রিয় করতে চান, যেমন নান্দনিক গুণমান বৃদ্ধি, ফাইল পরিচালনা বা প্রাক-সম্পাদনা কাজগুলি, ভয়েস কমান্ড ডিজাইনের জন্য আদর্শ হওয়া উচিত।

এই প্রেক্ষাপটে ব্যক্তিগতকরণ হল একটি স্মার্ট ভিডিও এডিটিং টুলের ক্ষমতা যা প্রক্রিয়াকৃত ভিডিও, উত্পাদিত ভিডিও এবং সঙ্কেতের ব্যবহারের ধরণগুলি থেকে শিখে ব্যবহারকারীর সাথে খাপ খাইয়ে নিতে পারে। এই বিষয়টি সাহিত্য থেকে অনুপস্থিত কিন্তু ভিডিও সম্পাদনা এবং ব্যক্তিগতকরণের প্রেক্ষাপটে বোঝার আকারে আমাদের গবেষণায় পাওয়া গেছে। ভিডিও এডিটিং নিয়মের নিয়ম-ভিত্তিক শিক্ষা [মাতসুও এট আল।] এ আলোচনা করা হয়েছে।

এটি লক্ষ করা গুরুত্বপূর্ণ যে বিভাগ 3 এ আলোচিত স্মার্ট ভিডিও সম্পাদনা সরঞ্জামগুলি একটি নির্দিষ্ট ধরনের ভিডিও সম্পাদনা সমাধানের জন্য ডিজাইন করা হয়েছে: প্রকৃতপক্ষে, তাদের মধ্যে আটটি শুধুমাত্র এক ধরনের ভিডিওকে লক্ষ্য করে। যাইহোক, ভিডিও এডিটরদের আমাদের সমীক্ষায় গড়ে প্রতিটি দ্বারা প্রক্রিয়া করা তিন ধরনের ভিডিও তালিকাভুক্ত করা হয়েছে। যেহেতু বেশিরভাগ স্মার্ট ভিডিও এডিটিং টুল এক ধরনের ভিডিওর জন্য তৈরি করা হয়েছে, তাই ফলাফলে উল্লিখিত কৌশলগুলির সাধারণ প্রয়োগযোগ্যতা নিয়ে উদ্বেগ রয়েছে।

স্মার্ট ভিডিও এডিটিং টুল লিটারেচারে বর্ণিত ভিডিও এডিটিং টাঙ্কগুলি মৌলিক ভিডিও এডিটিং টাঙ্কের সময় মিথস্ক্রিয়াগুলির উপর ফোকাস করা হয়। যাইহোক, গবেষণা অংশগ্রহণকারীদের অতিরিক্ত কাজ যেমন ফাইল/মিডিয়া সংগঠন, নান্দনিক মানের উন্নতি, প্রাক-সম্পাদনা কাজ এবং বিষয়বস্তু পরামর্শের জন্য অটোমেশন প্রয়োজন। একটি ভিডিও সম্পাদনা বিষয়বস্তুর পরামর্শ একটি বিদ্যমান ভিডিও বা গল্পে যোগ করার জন্য ভাল ভিডিও বা সঙ্গীত বিভাগগুলির পরামর্শ দেয়।

সমীক্ষার ফলাফলে, ভয়েস ইন্টারঅ্যাকশন ছিল অংশগ্রহণকারীদের দ্বারা পছন্দসই মিথস্ক্রিয়ার সবচেয়ে সাধারণ মোড। যাইহোক, একটি বুদ্ধিমান ভিডিও সম্পাদকের শুধুমাত্র একটি কাজ ভয়েস [চি এট আল।, 2013] অন্তর্ভুক্ত করে, যেখানে ভিডিও ট্যাগ করতে ভয়েস টীকা ব্যবহার করা হয়। ভয়েস ইন্টারঅ্যাকশনের এই জনপ্রিয়তা মোবাইল ফোন এবং স্মার্ট হোম স্পিকারগুলিতে AI-ভিত্তিক ভয়েস সহায়তার জনপ্রিয়তার পাশাপাশি বিজ্ঞান কল্পকাহিনীতে একটি ভয়েস হিসাবে AI-এর উপস্থাপনার জন্য দায়ী করা যেতে পারে।

স্মার্ট ভিডিও এডিটিং টুলে ব্যবহৃত AI কৌশল হল হিউরিষ্টিক-ভিত্তিক সিস্টেম। অন্যান্য পদ্ধতি যেমন নিউরাল নেটওয়ার্ক এবং মেশিন লার্নিং-ভিত্তিক পন্থা কম অধ্যয়ন করা হয়। [Dove et al., 2017]-এর সমীক্ষা এই উপসংহারে পৌঁছেছে যে ব্যবহারকারীর অভিজ্ঞতা তৈরি করার জন্য মেশিন লার্নিং একটি কঠিন ডিজাইনের উপাদান, কারণ এটি মেশিন লার্নিং-ভিত্তিক প্রোটোটাইপ তৈরি করা কঠিন এবং এটি সম্পাদন করার জন্য মেশিন লার্নিং সহযোগীদের প্রশিক্ষণের প্রয়োজন।

## VI. উপসংহার

এই কাগজে, আমরা বুদ্ধিমান ভিডিও সম্পাদনা সরঞ্জামগুলিকে সংজ্ঞায়িত করেছি এবং (বুদ্ধিমান) ভিডিও সম্পাদনা, ব্যবহারকারীর মিথস্ক্রিয়া এবং এআই প্রযুক্তির উপর বিদ্যমান সাহিত্যের একটি ওভারভিউ উপস্থাপন করেছি।

আমরা ভিডিও এডিটরদের তাদের ভিডিও এডিটিং ওয়ার্কফ্লোতে অটোমেশনের প্রয়োজনীয়তা সম্পর্কেও জরিপ করেছি। গবেষণার এই ক্ষেত্রে গবেষণা করছেন জ্ঞান প্রয়োজন

ভিডিও সম্পাদনা, মানুষের কম্পিউটার মিথস্ক্রিয়া, এবং এআই বা মেশিন লার্নিং। ইন্টেলিজেন্ট ভিডিও এডিটিং টুলের জন্য এই তিনটি ভিন্ন ভিন্ন দৃষ্টিভঙ্গির প্রয়োজন একটি কারণ হল প্রতিটি ক্ষেত্রের তুলনায় পৃথকভাবে তাদের উপর সাহিত্য খুবই সীমিত।

এই গবেষণায়, ভিডিও সম্পাদনা কাজের ক্ষেত্রে সাহিত্য এবং গবেষণার ফলাফলের মধ্যে প্রচুর পরিমাণে ক্রসওভার রয়েছে। যাইহোক, এমন কিছু ক্ষেত্র রয়েছে যেমন ভিডিও রেকর্ডিং বা ভিডিও সম্পাদনা প্রকল্পগুলি সংগঠিত করা, নান্দনিক গুণমান সংশোধন, এবং বিষয়বস্তু পরামর্শগুলি যা গবেষণার প্রয়োজনগুলি পূরণ করতে আরও অন্বেষণ করা দরকার। আমাদের উপসংহার হল যে মেশিন লার্নিং সম্প্রদায়ের বৃহত্তর সম্পৃক্ততার সাথে, আদর্শ এআই সম্পাদক অর্জন করা যেতে পারে। ভবিষ্যতের কাজে, আমরা এই লক্ষ্যে অবদান রাখতে চাই।

## তথ্যসূত্র

- [1] এফ বাথুজোজ। ইন্টারভিউ ভিডিওতে কাট এবং ট্রানজিশন রাখার জন্য টুল। পৃষ্ঠা 8, 2012। এস. বোকোনি। শব্দার্থ-সচেতন স্বয়ংক্রিয় ভিডিও সম্পাদনা। মাল্টিমিডিয়া - মাল্টিমিডিয়া '04, পৃষ্ঠা 971, 2004-এর 12 তম বার্ষিক ACM আন্তর্জাতিক সম্মেলনের কার্যক্রমে।
- [2] এস. বাটলার এবং এ. পার্কস। স্বয়ংক্রিয় বুদ্ধিমান ভিডিও সম্পাদনার জন্য ফিল্ম সিকোয়েন্স প্রজন্মের কৌশল। ফলিত কৃত্রিম বুদ্ধিমত্তা, 11(4):367-388, জুন 1997।
- [3] জে. ক্যাসারেস, এসি লং, বিএ ম্যাসার্স, আর. ভাটনগর, এসএম স্ট্রিভেনস, এল. ড্যাবিশ, ডি. ইয়োকাম, এবং এ. করবেটা। মেটাডেটা ব্যবহার করে ভিডিও সম্পাদনা সহজ করা। পৃষ্ঠা 10, 2002।
- [4] RG Cattelan, C. Teixeira, R. Goularte, এবং MDGC Pimentel। সর্বব্যাপী ইন্টারেক্টিভ ভিডিও সম্পাদনার প্রতি দৃষ্টি হিসাবে দেখুন এবং মন্তব্য করুন। মাল্টিমিডিয়া কম্পিউটিং, কমিউনিকেশনস এবং অ্যাপ্লিকেশনের উপর ACM লেনদেন, 4 (4):1-24, অক্টোবর 2008।
- [5] এম. চ্যাং, এ. টুং, ও. ওয়াং, এম. আগ্রাওয়ালা, এবং জে. কিম। কিভাবে ভিডিওর জন্য ভয়েস ভিত্তিক নোভিগেশন ডিজাইন করবেন। কম্পিউটিং সিস্টেমের মানবিক উপাদানের উপর 2019 CHI সম্মেলনের কার্যপ্রণালীতে, পৃষ্ঠা 1-11, মে 2019।
- [6] P.-Y. চি, জে. লিউ, জে. লিন্ডার, এম. ডনচেভা, ডব্লিউ লি, এবং বি হাটম্যান। ডেমোকট: শারীরিক প্রদর্শনের জন্য সংক্ষিপ্ত নির্দেশমূলক ভিডিও তৈরি করা। পৃষ্ঠা 10, 2013, ভি. সিসকো। সিসকো ভিজ্যুয়াল নেটওয়ার্কিং সূচক: পূর্বাভাস এবং প্রবণতা, 2017-2022। শ্বেতপত্র, 1:1, 2018।
- [7] জি. ডভ, কে. হালসকত্ত, জে. ফোরলিজি এবং জে. জিমারম্যান। ইউএক্স ডিজাইন ইনোভেশন: ডিজাইন মেটেরিয়াল হিসেবে মেশিন লার্নিং এর সাথে কাজ করার জন্য চ্যালেঞ্জ। পৃষ্ঠা 278-288, 2017। জে. আইজেনস্টাইন এবং আর. বারজিলে। Bayesian unsupervised বিষয় সেগমেন্টেশন. ন্যাচারাল ল্যাঙ্গুয়েজ প্রোসেসিং-এ অভিজ্ঞতামূলক পদ্ধতির সম্মেলনের কার্যক্রমে - EMNLP '08, পৃষ্ঠা 334, 2008।
- [8] এক্স-এস. হ্যা, এল. ল. এবং এইচ. জে. বাং। অস্ট্রাইজেশান-ভিত্তিক স্বয়ংক্রিয় হোম ভিডিও এডিটিং সিস্টেম। ভিডিও প্রযুক্তির জন্য সার্কিট এবং সিস্টেমের উপর IEEE লেনদেন, 14(5):572-583, মে 2004।
- [9] টি. কিমুরা, কে. সুমিয়া, এবং এইচ. তানাকা। ব্যবহারকারীর দৃষ্টিশক্তি ব্যবহার করে একটি ভিডিও সম্পাদনা সমর্থন সিস্টেম। ইন

প্যাক্রিম 2005 IEEE Pacific Rim Conference on Communications, Computers and Signal Processing, 2005., পৃষ্ঠা 149–152, 2005।

- [১০] মি. লিক, এ. ডেভিস, এ. টুং এবং এম. আগরওয়াল। সংলাপ-চালিত দৃশ্যের জন্য গণনামূলক ভিডিও সম্পাদনা। গ্রাফিক্সে ACM লেনদেন, 36(4):1–14, জুলাই 2017।
- [১১] এসি লং, বিএ মায়ার্স, জে. ক্যাসারেস, এসএম স্টিভেনস এবং এ. করবেট। লেক্স এবং শব্দার্থিক জুমিং ব্যবহার করে ভিডিও সম্পাদনা। পৃষ্ঠা 10, 2004।
- [12] ওয়াই. মাতসুও, এম. আমানো এবং কে. উয়েহার। ভিডিও স্ক্রিমগুলিতে মাইনিং ভিডিও সম্পাদনা করার নিয়ম। পৃষ্ঠা 4. বিএ মায়ার্স, জেপি ক্যাসারেস, এস. স্টিভেনস, এল ড্যাব্রিশ, ডি. ইয়োকাম, এবং এ. করবেট। ডিজিটাল ভিডিও লাইব্রেরির জন্য একটি মাল্টি-ভিডিও ইন্টেলিজেন্ট এডিটর। ডিজিটাল লাইব্রেরিতে প্রথম ACM/IEEECS যৌথ সম্মেলনের কার্যক্রমে - JCDL '01, পৃষ্ঠা 106–115, 2001।
- [১৩] এফ. নাক এবং A. Parkes, স্বয়ংক্রিয় ভিডিও সম্পাদনায ভিডিও শব্দার্থবিদ্যা এবং থিম প্রতিনিধিত্বের প্রয়োগ। HJ Zhang, P. Aigrain, এবং D. Petkovic, সম্পাদক, মাল্টিমিডিয়া সিস্টেমে ভিডিও ডেটার প্রতিনিধিত্ব এবং পুনরুদ্ধার, পৃষ্ঠা 57-83। 1997।
- [১৪] জেএ ওকুন, এস. জাওয়ারম্যান, কে. রাফারটি, এবং এস. স্কয়ারস, সম্পাদক। ভিজ্যুয়াল এক্শনের ভিইএস হ্যান্ডবুক: ইন্ডাস্ট্রি স্ট্যান্ডার্ড ভিএফএক্স অনুশীলন এবং পদ্ধতি। 2015।
- [১৫] পু. প্যাসারেলি। স্বয়ংক্রিয় সম্পাদনা দ্রুত পার্থ্য ভিত্তিক ভিডিও সম্পাদনা, 2019।
- [16] ক. পাতেল, সি. রিড, বি. হার্টম্যান এবং এম. আগরওয়াল। ভিডিও ডাইজেস্ট: তথ্যমূলক বক্তৃতা ভিডিওগুলির জন্য একটি ব্রাউজযোগ্য, স্ক্রিমযোগ্য বিন্যাস। ইউজার ইন্টারফেস সফটওয়্যার এবং প্রযুক্তির উপর 27 তম বার্ষিক ACM সিম্পোজিয়ামের কার্যপ্রণালীতে, পৃষ্ঠা 573-582, অক্টোবর 2014।
- [১৭] এম. Radut, M. Evans, K. To, T. Nooney, এবং G. Phillipson. কত ভালো ভালো যথেষ্ট? লাইভ ইভেন্টের এআই-সম্পাদিত ভিডিও কভারেজের বিষয়গত গুণমানের মূল্যায়নের চ্যালেঞ্জ। ইন্টেলিজেন্ট সিনেমাটোগ্রাফি এবং এডিটিং-এর উপর কর্মশালা, পৃষ্ঠা 8 পৃষ্ঠা, 2020।
- [18] আর্টওয়ার্ক সাইজ: 8 পৃষ্ঠা। ISBN: 9783038681274 প্রকাশক: দ্য ইউরোগ্রাফিক্স অ্যাসোসিয়েশন সংস্করণ নম্বর: 017-024।
- [১৯] বি. R&D. এআই এবং আর্কাইভ - মেশিন দ্বারা তৈরি, 2018। এফ. শিপম্যান। হাইপারভিডিও লেখা, দেখা এবং জেনারেট করা: হাইপার-হিচককের একটি ওভারভিউ। 5(2):19, 2008।
- [20] পি. শ্রেষ্ঠা, পিএইচ ডি উইথ, এইচ. ওয়েডা, এম. বারবিয়েরি এবং ইএইচ আর্টস। একাধিক ক্যামেরা কনসার্ট রেকর্ডিং থেকে স্বয়ংক্রিয় ম্যাশআপ জেনারেশন। মাল্টিমিডিয়ায় আন্তর্জাতিক সম্মেলনের কার্যক্রমে - এমএম '10, পৃষ্ঠা 541, 2010।
- [২১] সি. তাসকির, জেড. পিজলো, এ. আমির, ডি. পোনসেলিয়ন, এবং ই. ডেলপা। বক্তৃতা প্রতিলিপি ব্যবহার করে স্বয়ংক্রিয় ভিডিও প্রোগ্রামের সারাংশ। মাল্টিমিডিয়ায় IEEE লেনদেন, 8(4):775–791, আগস্ট 2006।
- [২২] ক. টুং, এফ. বাথজোজ, ডব্লিউ লি, এবং এম. আগরওয়াল। QuickCut: বর্ণনাকৃত ভিডিও সম্পাদনা করার জন্য একটি ইন্টারেক্টিভ টুল। ইউজার ইন্টারফেস সফটওয়্যার এবং প্রযুক্তির উপর 29 তম বার্ষিক সিম্পোজিয়ামের কার্যক্রমে, পৃষ্ঠা 497-507, অক্টোবর 2016।
- [২৩] টি. টুং এবং এস ভেক্টেশ। ভিডিও বিমর্ততা: একটি পদ্ধতিগত পর্যালোচনা এবং শ্রেণীবিভাগ। 3(1):37। H.-Y. উ. টি. সান্তারা, এম. লিস, আর. ভাগাস, এবং এ. ঝালা। স্বয়ংক্রিয় ভিডিও সম্পাদনার জন্য যৌথ মনোযোগ। ইন্টারেক্টিভ মিডিয়া অভিজ্ঞতার উপর ACM আন্তর্জাতিক সম্মেলনে, পৃষ্ঠা 55-64, জুন 2020।
- [২৪] ওয়াই. Wu, T. Mei, Y.-Q. জু, এন. ইউ, এবং এস. লি. MoVieUp: স্বয়ংক্রিয় মোবাইল ভিডিও ম্যাশআপ। ভিডিও প্রযুক্তির জন্য সার্কিট এবং সিস্টেমের উপর IEEE লেনদেন, 25(12):1941–1954, ডিসেম্বর 2015।
- [২৫] জেড। উ. টি. ইয়াও, ওয়াই ফু এবং ওয়াই-জি। জিয়াং। ভিডিও শ্রেণীবিভাগ এবং ক্যাপশনিং এর জন্য গভীর শিক্ষা। arXiv প্রিন্ট arXiv:1609.06782, 2016. C. Xue, L. Li, F. Yang, P. Wang, T. Wang, এবং Y. Zhang. স্বয়ংক্রিয় হোম ভিডিও সম্পাদনা: একটি মাল্টি-কোর সমাধান। পৃষ্ঠা 2।