

1 Základní informace

Jméno projektu	Systém pro hledání známých scén ve videu
Zkratka	KIS-S
Vedoucí	RNDr. Jakub Lokoč, Ph.D. (lokoc@ksi.mff.cuni.cz)
Anotace	<i>Cílem projektu je vytvořit modulární framework pro potřeby vyhledávání známých scén ve videu. Odevzdaná verze by měla podporovat automatizované základní předzpracování zadané kolekce video souborů (tj. extrakce vlastností a dat pro vyhledávání) a umožnit efektivní vyhledávání v uživatelsky přívětivém prototypu vyhledávače. Navržený framework by měl umožňovat snadnou rozšiřitelnost.</i>

1 Motivace

Vyhledávání ve videu je intenzivně rozvíjená disciplína, pro kterou se neustále objevují nové aplikace. Tento projekt by se měl zaměřit na problematiku hledání známé scény (z anglického *known-item search*, KIS) ve videu, kdy uživatel hledá jednu konkrétní scénu na základě textového popisu nebo předchozího shlédnutí. Tato úloha má řadu potenciálních praktických aplikací a proto ji zahrnují do svého programu i prestižní mezinárodní soutěže (např. *Video Browser Showdown*). K řešení KIS úloh je potřeba najít vhodné kombinace vyhledávacích modelů, které dnes často silně závisí na použitých technikách strojového učení pro předzpracování dat. Vyhledávací modely je dále potřeba vhodně parametrizovat tak, aby reflektovaly distribuce vlastností v datech. Samotný vyhledávací software pak modulárně integruje komponenty pro předzpracování videa, komponenty pro zadávání/vyhodnocování dotazů a komponenty pro vizualizaci výsledků. Podobný experimentální SW byl průběžně vyvíjen studenty pod vedením J. Lokoče, nicméně, na základě nejnovějších výsledků výzkumu navrhnou řešitelé novou modulární verzi podporující nejnovější modely pro předzpracování, hledání a vizualizaci. Vzhledem k rozsahu požadovaných činností při řešení tohoto projektu je zapotřebí sestavit vyvážený tým s členy zaměřenými jak na SW inženýrství, tak i na strojové učení. Jelikož se jedná o vědecky orientovaný SW projekt, tak by měl vzniklý SW sloužit i jako platforma pro sběr experimentálních dat (např. logování interakcí, exporty výsledků hledání a jejich parametry pro další analýzy).

2 Popis projektu

Cílem projektu je vytvořit rozšiřitelný framework pro hledání známých scén ve videu, zahrnující jednoduché předzpracování dat, vhodné vyhledávací modely a funkční uživatelsky přívětivý prototyp. S tímto prototypem by mělo být možné důstojně reprezentovat MFF UK na mezinárodních soutěžích jako je *Video Browser Showdown*. Framework bude integrovat:

- *Natrénovaný model pro automatickou detekci stříhů a reprezentativních snímků. Řešitelé zdůvodní výběr daného modelu a provedou srovnávací experimenty.*
- *Natrénovaný model pro automatickou anotaci snímků. Řešitelé zdůvodní výběr daného modelu a provedou srovnávací experimenty. Zadávání klíčových slov bude uživatelsky přívětivé (např. nápověda) a budou se ukazovat statistiky z dat (např. četnost výskytu).*
- *Model pro vyhledávání pomocí skicáře. Minimálně budou podporovány barevné skicy a*

možnost zadání pozice (automaticky detekovaného) obličeje.

- *Model pro vyhledávání pomocí obrázku (vybraného z databáze nebo externího vyhledávače).*
- *Filtry pro barevnost, průměrnou intenzitu, již zobrazené snímky, nebo omezení počtu zobrazených snímků z videa/scény.*
- *Extrakce všech vlastností pro implementované modely bude plně automatizovaná.*
- *Vizualizace bude zahrnovat rychlý grid pro zobrazení snímků (včetně možného kontextu) a metodu pro organizaci snímků ve 2D na základě jejich vzájemné podobnosti.*
- *Framework bude zvládat rozsáhlejší kolekce videa. Minimálně by měl bezproblémů zvládnout vyhledávání ve V3C datasetu, který má 1000 hodin.*
- *Framework bude podporovat logování interakcí uživatele, export výsledků hledání s dodatečnými informacemi jako relevance, deskriptory, pozice hledané scény.*

3Platforma, technologie

Budou navrženy, zdůvodněny a prezentovány v rámci podrobné specifikace projektu.

4Odhad náročnosti

Projekt má 4-členný řešitelský tým:

- *Bc. Gregor Kovalčík: návrh celkové modulární architektury, implementace a konfigurace multi-modálních vyhledávacích algoritmů, návrh rozhraní a implementace základních komponent, hlavní správce projektového repozitáře, vedení studentského týmu.*
- *RNDr. Přemysl Čech: konzultace a návrh integračních technologií, návrh architektury a rozhraní, návrh a vývoj klientské části vyhledávače a exporty dat.*
- *Bc. Tomáš Souček: výběr/návrh, optimalizace a trénování hlubokých sítí pro automatickou anotaci vybraných snímků, detekce obličejů/textů a embedování obrázků do vektorových reprezentací. Srovnávací experimenty a integrace sítí a jejich výsledků do systému.*
- *Bc. Jaroslav Moravec: předzpracování video kolekce zahrnující extrakce keyframes, jejich úprava, návrh hluboké neuronové sítě pro automatickou detekci střihů, automatický výběr reprezentativních snímků, tvorba netriviální kontrolní anotované kolekce, vizualizační nástroje pro efektivní kontrolu a srovnávací experimenty.*

*Plán prací (studenti již na projektu 4 měsíce **intenzivně** pracují)*

- *1 měsíc: analýza stávajících nástrojů, vhodného jádra vyhledávače, podrobné konzultace s vedoucím projektu J. Lokočem*
- *2-3 měsíc: návrh architektury a integrace datových zdrojů, dále pak práce na předzpracování dat - trénování detektoru střihů, modelů pro automatickou anotaci*

vybraných snímků a extrakce vlastností (pro 2TB dat)

- 3-4 měsíc: intenzivní implementace a ladění celého systému, tvorba prototypu a testování jeho efektivnosti na mezinárodní soutěži Video Browser Showdown 2019 (tým zvítězil ve dvou kategoriích a celkově získal 2. místo z 6-ti týmů - <http://www.videobrowsershowdown.org/example-browsers/infos-and-results-2019/>)
- 5 měsíc: analýza dat ze soutěže (interakční logy) a příprava vědecké publikace na konferenci ICMR, dodatečné evaluace systému a analýza slabých stránek
- 6 měsíc: rozšíření prototypu pro potřeby vyhledávání v Lifelog datech (příprava systému na mezinárodní soutěž Lifelog Search Challenge v červnu 2019), integrace poznatků z předchozí analýzy
- 7 měsíc: finalizace projektu a dokumentace

5

6 Vymezení projektu

Projekt je zaměřen na následující oblasti (zaškrtněte vyhovující):

Diskrétní modely a algoritmy	
	diskrétní matematika a algoritmy
	geometrie a matematické struktury v informatice
	optimalizace
Teoretická informatika	
	Teoretická informatika
Softwarové a datové inženýrství	
X	softwarové inženýrství
	vývoj software
	webové inženýrství
X	databázové systémy
X	analýza a zpracování rozsáhlých dat
Softwarové systémy	
	systémové programování
	spolehlivé systémy
	výkonné systémy
Matematická lingvistika	
	počítačová a formální lingvistika
	statistické metody a strojové učení v počítačové lingvistice
Umělá inteligence	
	inteligentní agenti

X	strojové učení
	robotika
Počítačová grafika a vývoj počítačových her	
	počítačová grafika
	vývoj počítačových her

7Poznámky

RNDr. Přemysl Čech je Ph.D. student a nepotřebuje zápočet. Studenti na projektu intenzivně pracují již 4 měsíce, (téměř) každý týden probíhají pravidelné konzultační schůzky. V rámci projektu již spolupracovali na přijaté publikaci na konferenci MMM a také na publikaci na konferenci ACM ICMR, která je nyní v recenzním řízení.