



Zpráva ze zahraniční služební cesty

Jméno a příjmení účastníka cesty	Dr. Jiří Polišínský
Pracoviště – dle organizační struktury	Oddělení správy NKF
Pracoviště – zařazení	Oddělení správy NKF
Důvod cesty	Návštěva SNK Martin
Místo – město	Martin
Místo – země	Slovensko
Datum (od-do)	31.10.2013
Podrobný časový harmonogram	31.10.2013 ráno odjezd z Prahy, jednání Martin, v noci příjezd zpět do Prahy
Spolucestující z NK	Mgr. Tomáš Foltýn
Finanční zajištění	Reciprocita
Cíle cesty	Prohloubení spolupráce mezi institucemi, vzájemná spolupráce při výměně bohemikálních/slovacikálních dokumentů. Představení projektu ANL retro a diskuze o zpracování digitálních dokumentů na úroveň článků včetně zkušeností NK ČR s projektem podávaným prostřednictvím VISK 7. Ukázky dat.
Plnění cílů cesty (konkrétně)	Viz cíle cesty
Program a další podrobnější informace	
Přivezené materiály	---
Datum předložení zprávy	3. 10. 2013
Podpis předkladatele zprávy	Dr. Jiří Polišínský
Podpis nadřízeného	
Vloženo na Intranet	
Přijato v mezinárodním oddělení	

25.8.

Účast na tutorialu „Performance Evaluation in Document Image Analysis“

Zpráva je pracovníkem do mezinárodního oddělení předložena nejpozději při vyúčtování cesty do 2 týdnů po jejím ukončení. Bez cestovní zprávy nebude provedeno vyúčtování. Při výjezdu více pracovníků na tutéž služební cestu s týmž programem lze odevzdat společnou cestovní zprávu.

- diskuze nad metodami pro vyhodnocování zpracování jednotlivých předem definovaných setů dat
- představení výukového programu pro sestavování vhodných metodik vyhodnocování
- diskuze nad použitými metrikami
- Ground Truth data
- praktické testování softwarových nástrojů a následné interaktivní hodnocení postupů

Účast na vstupní recepci

26.8. – 29.8. 2013

Účast na konferenci ICDAR 2013

Nejzajímavější sekce a přednášky:

Camera Based OCR

- nahrazování tradičních způsobů strojového skenování a následného zpracování OCR za využití snímání minikamerami umístěnými například na speciálních brýlích (výhodou je rychlé použití, snímání pouze těch částí dokumentu, které jsou zajímavé pro čtenáře)
- ukázky výstupů a výsledků; další možnosti použití (například v univerzitních knihovnách snímat pouze to, co studenti potřebují a následně tyto důležité pasáže transformovat přes OCR)
- ukázky transformačních algoritmů pro převody textů
- nejzajímavější přednášky v těchto sekcích:
 - Multiple Geometry Transform Estimation from Single Camera-captured Text Image (Xin Zhang and Fuchun Sun)
 - Local Subspace Classifier with Transformation-Invariance for Appearance-Based Character Recognition in Natural Images (Keisuke Higa and Seiji Hotta)

Character Recognition

- nové pohledy na písmenná schémata
- vývojové projekty na zlepšení binarizace původních textů
- uvedení možností využití pravidel vektorové logiky včetně představení základních vzorců
- nejzajímavější přednášky v těchto sekcích:
 - Moment-based Character Normalization Methods using Contour Image Combined with Original Image (Toshinori Miyoshi, Takeshi Nagasaki)
 - Character Recognition using CRF based Matching Engine (Anupama Ray)

Applications

- obsahová návaznost na sekce s tematikou Camera Based OCR
- představení aplikací pro využití „Mobile Eye Tracking“
- nástroje pro extrakci obsahu dle relevance pro uživatele
- nástroje pro automatické zaznamenávání preferencí a návyků čtenářů
- nástroje pro tzv. smart-binarizaci
 - Document Information Exploitation and its Evaluation based on Client's Relevance (K.C. Santosh and Abdel Belaid.)
 - The Reading-life Log - Technologies to Recognize Texts That We Read (Takashi Kimura, Rong Huang)

Zpráva je pracovníkem do mezinárodního oddělení předložena nejpozději při vyúčtování cesty do 2 týdnů po jejím ukončení. Bez cestovní zprávy nebude provedeno vyúčtování. Při výjezdu více pracovníků na tutéž služební cestu s tímž programem lze odevzdat společnou cestovní zprávu.

Layout Analysis

- představení nových metod v zónování textu – komplexní zónování textů, obrazových příloha či tabulek v jednom kroku
- představení nástroje pro automatickou detekci vnitřních i vnějších regionů
- Hybrid Page Segmentation with Efficient Whitespace Rectangles Extraction and Grouping (Kai Chen, Fei Yin)

Účastník se během akce zúčastnil i dalších přednášek týkajících se zejména metod rozpoznávání rukopisných textů a dále i všech poster sekcí. Během konference se seznámil s mnoha zahraničními experty, z nichž nejhodnotnější náměty předali pracovníci University Salford, Technické university Kaiserslautern, Technologického institutu v Rochestru či z japonské University Kyushu.

30.8.

Dopolední účast na neformálním semináři k nástrojům pro budování jazykových databází a jejich zpětném využití v rámci procesů OCR

- představení nástroje CODEG vyvíjeném v rámci projektu NAKI „Nástroje pro zpřístupnění tištěných textů...“
 - ukázky segmentace a strukturování dokumentu
 - převody do databází
 - diskuze nad možnostmi systému a způsobu generování dalších slovních tvarů
- prohlédnutí dalších nástrojů – například systému Aletheia
- diskuze nad produkcí Ground Truth textů a jejich následnému použití při testování jednotlivých systémů

Odpoledne prohlídka Kongresové knihovny.

Zpráva je pracovníkem do mezinárodního oddělení předložena nejpozději při vyúčtování cesty do 2 týdnů po jejím ukončení. Bez cestovní zprávy nebude provedeno vyúčtování. Při výjezdu více pracovníků na tutéž služební cestu s týmž programem lze odevzdat společnou cestovní zprávu.