

# Ukladanie výsledkov extrakcie informácií do IPTC hlavičiek športových obrázkov<sup>2</sup>

Miroslav Kolínský, Jan Nemrava, Vojtěch Svátek

Katedra informačního a znalostního inženýrství, Vysoká škola ekonomická v Praze,  
nám. W. Churchilla 4, 130 67 Praha 3  
{xkolm44, nemrava, svatek}@vse.cz

**Abstrakt.** IPTC Manipulátor vyvíjaný na Vysokej škole ekonomickej v Prahe je pomocou nástroja SPRouT schopný extrahovať sémantické koncepty z textu obklopujúceho obrázky a tieto informácie ukladať priamo do IPTC hlavičiek JPG súborov. Informácie o objektoch, udalostiach alebo osobách zachytených na obrázku sa stanú jeho priamou súčasťou, v dôsledku čoho odpadá potreba externého nosiča informácií. Dáta získané pomocou tohto nástroja sa ďalej používajú ako vstup pre K-Space Annotation Tool, anotačný nástroj pre multimédiá vyvíjaný v Koblenzi v rámci projektu K-Space.

**Kľúčová slova:** anotácia, IPTC, KAT, K-Space

## 1 Úvod

Pre vyhľadávanie multimediálnych súborov v rozsiahlejších databázach je potrebné stanoviť vhodné kritériá vyhľadávania. Tieto kritériá majú v prípade multimédií väčšinou sémantický charakter. Užívateľ sa snaží vyhľadať obrázky alebo videá, ktoré zobrazujú určité osoby, objekty alebo udalosti. Aby bolo vyhľadávanie na základe sémantických kategórií možné, je potrebné k jednotlivým multimédiám pripojiť informácie o ich obsahu. Tento proces je pri malom objeme dát možné vykonávať ľudskou činnosťou. Akonáhle ale začne objem dát narastať, stáva sa táto úloha pre človeka stále obtiažnejšou, až nereálnou. Na generovanie anotácie veľkých objemov multimediálnych dát je vhodný automatický systém. V tomto článku sa budeme zaoberať nástrojom generujúcim sémantický opis obrázkov pomocou analýzy textu, ktorý tento obrázok obklopuje. Popisovaný program vzniká na Vysokej škole ekonomickej v Prahe ako podpora pre anotačný nástroj KAT – K-Space Annotation Tool. KAT je anotačný nástroj umožňujúci automatickú a semiautomatickú anotáciu textu, obrázkov a videa. Vzniká v rámci európskeho projektu K-Space<sup>3</sup>, ktorý sa zaoberá sémantikou v multimédiách. Cieľom je umožniť programu KAT využívať výsledky automatickej analýzy textu pri semiautomatickej analýze obrázkov.

## 2 Metadáta v obrázkoch

Hlavná myšlienka analýzy obklopujúceho textu spočíva vo fakte, že text nachádzajúci sa bezprostredne pred alebo za obrázkom často obsahuje informácie o osobách, objektoch alebo udalostiach, ktoré obrázok zachytáva. Náš software je schopný

---

<sup>2</sup> Výskum je podporovaný Európskou komisiou pod zmluvou FP6-027026, projekt K-Space.

<sup>3</sup> <http://www.k-space.eu/>

získavať koncepty z futbalového prostredia pomocou nástroja SProuT [1]. SProuT je schopný extrahovať z textu koncepty na základe ontológie a uložiť ich vo formáte XML. Komunikácia so SProuTom prebieha pomocou webovej služby. Po získaní konceptov je potrebné spojiť nové informácie s obrázkom.

Jednou z možností ako toho docieľiť je uloženie konceptov ako metadát priamo do obrázku. Najvhodnejšou formou uloženia dát do obrázkov sú IPTC hlavičky. IPTC<sup>4</sup> - The International Press Telecommunications Council definoval IPTC hlavičky ako množinu atribútov metadát, ktoré je možné vložiť do obrázkov podporujúcich túto štruktúru. Súčasnú podobu nadobudli v roku 1994. IPTC hlavičky obsahujú atribúty všeobecného charakteru, ako napríklad City, Country, Creator, Description, Source a pod. Výhoda použitia tohto prístupu spočíva v priamom spojení informácií s obrázkom, ktorý popisujú. Tým je zaručené, že sa informácia o obsahu vždy preniesie spolu s obrázkom a odpadá potreba externého nosiča informácie. IPTC so sebou zároveň prinášajú niekoľko problémov. Prvým z nich je všeobecná povaha atribútov, ktorá nevyhovuje konceptom z futbalového prostredia. Pre koncept typu „hráč“ alebo „skóre“ v IPTC hlavičkách neexistuje atribút s vhodnou sémantikou. Tento problém by bolo možné odstrániť vytvorením vlastných atribútov. Takýto postup síce povaha IPTC hlavičiek umožňuje, ale zároveň činí nepoužiteľným v praxi, pretože programy pracujúce s informáciami uloženými v IPTC hlavičkách spravidla pracujú len so štandardizovanými atribútmi. Atribút IPTC hlavičky nieje označený slovnou, ale pomocou dvojice čísel. Informácia o názve atribútu sa neuchováva, z čoho vyplýva nemožnosť zistenia zamýšľaného názvu. Ďalšou nevýhodou IPTC hlavičiek je obmedzený počet znakov, ktoré môžu atribúty obsahovať. S výnimkou jedného sú všetky atribúty obmedzené dĺžkou, čo spôsobuje problém s uložením informácií pri dlhších konceptoch alebo ich viacnásobnom výskyte. Jediným riešením problému je uloženie všetkých informácií o konceptoch do atribútu „Description“, pretože práve tento atribút nie je obmedzený počtom znakov. Tým sa ale znižuje sémantická interoperabilita, pretože využitie informácií je potom úplne závislé na znalosti proprietárnej vnútornej štruktúry tohto atribútu. Alternatívou k použitiu IPTC hlavičiek je technológia EXIF<sup>5</sup>. Tento formát je však určený primárne na ukladanie technických údajov obrázkov a trpí rovnakými nedostatkami ako IPTC. Obidva spomínané prístupy majú však ešte jeden spoločný problém, ktorým je nedostatočná podpora zo strany vyhľadávacích programov. Vyhľadávať obrázok pomocou údajov, ktoré sa nachádzajú v IPTC hlavičkách je s použitím bežných prostriedkov prakticky nemožné. Aj keď sa nám podarí uložiť informácie vo vhodnej forme priamo do obrázku, niesme schopný tento výsledok za účelom vyhľadávania využiť v praxi. Anotačný nástroj KAT pre svoju prácu využíva RDF<sup>6</sup> sklady metadát. RDF – Resource Description Framework je špecifikáciou konzorcia W3C originálne navrhnutou ako metadátový model, neskôr sa vyvinula na všeobecne prijímanú metódu modelovania informácií. Všetky sémantické informácie, s ktorými KAT pracuje, sú ukladané a načítavané z formátu RDF. Semiautomatická anotácia pomocou nástroja KAT spočíva v tom, že užívateľ dostane možnosť priradovať jednotlivým segmentom obrázku sémantické koncepty. Úlohou nášho

<sup>4</sup> <http://www.iptc.org/>

<sup>5</sup> <http://www.exif.org/>

<sup>6</sup> <http://www.w3.org/RDF/>

nástroja je vďaka analýze textu vygenerovať koncepty, ktoré budú následne užívateľovi ponúknuté pri ručnej anotácii. Keďže všetky informácie, s ktorými KAT pracuje sú uložené vo formáte RDF, je vhodné aby aj náš nástroj podporoval tento formát. Nevýhodou tohto spôsobu ukladania a prenosu informácií je oddelenie textu a obrazu do nezávislých súborov. Implementácia podpory formátu RDF našim nástrojom je predmetom práce v budúcnosti.

### 3 Problémy extrakcie konceptov

Proces extrakcie konceptov z textu so sebou prináša niekoľko problémov. Prvým z nich viacnásobný výskyt niektorých typov informácií ako napríklad dátumy, krajiny alebo mestá. Nieje neobvyklé, že niektoré z týchto informácií sa vzťahujú k minulosti alebo budúcnosti a nemajú žiadnu priamu súvislosť s obsahom obrázku. Pre zlepšenie získavania relevantných konceptov vzhľadom k obrázku je teda potrebné nájsť spôsob eliminácie týchto nežiadúcich informácií. Jednou z možností je použitie jazykového modulu, ktorý by bol schopný rozlišovať vety v budúcom a minulom čase, čím by sa umožnilo filtrovanie konceptov, ktoré sa v týchto vetách nachádzajú. V niektorých prípadoch ale nieje viacnásobnosť sprevádzaná výskytom vo vetách v minulosti alebo budúcnosti. Príkladom je situácia, v ktorej text obsahuje názvy krajín, ktoré na turnaji hrajú v spoločnej skupine. V tomto prípade problém spočíva v obtiažnosti určenia krajín, ktoré sa vzťahujú k sémantike zachytenej obrázkami. Momentálne najlepším riešením tohoto problému je až manuálna selekcia konceptov priamo v nástroji KAT.

Ďalšou problematickou oblasťou je existencia textov, ktoré používajú ako opisné prostriedky typy konceptov, ktoré sú rozpoznateľné nástrojom SProuT. Tieto slová sú v texte identifikované, ale nevzťahujú sa na žiadny objekt ani udalosť zachytenú na obrázku. Príkladom môže byť veta: "A boy by the entrance shows a ball signed by the French midfielder Zinedine Zidane". SProuT identifikuje hráča s menom Zinedine Zidane, ale obrázok pritom zachytáva chlapca držiaceho loptu podpísanú Zidaneom. Tento problém by sa dal riešiť pomocou syntaktického parseru schopného identifikovať vetné členy a ich štruktúry, napríklad prívlastok predmetu v pasíve. Môžeme povedať, že ak sa prívlastok prívlastok predmetu (tu: „French midfielder...“) vo vete nachádza v pasíve, pravdepodobnosť, že bude zobrazený priamo na obrázku je veľmi nízka. Na druhej strane je možná existencia viet, ktoré obsahujú vetné členy v pasíve zobrazené na obrázkoch. Automatická eliminácia týchto vetných členov by však v niektorých iných doménach automatická eliminácia takýchto vetných členov možno mohla viesť k strate podstatných informácií. Aj tento problém sa dá čiastočne riešiť manuálnou selekciou vhodných konceptov.

### 4 Príklad využitia

Príkladom použitia nástroja môže byť veta:

*Norway's goalkeeper Thomas Myhre addresses the crowd after the 2006 World Cup qualifying first leg soccer match against the Czech Republic at the Ullevaal stadium in Oslo, Norway on Saturday Nov. 12, 2005. Czech Republic defeated Norway 1-0. (AP Photo/Petr David Josek).*

Pre túto vetu identifikuje náš nástroj nasledujúce koncepty (vo formáte XML) :

```

<player>
  <surname>Myhre</surname>
  <given_name>Thomas</given_name>
  <team>NOR</team>
  <role>goalkeeper</role>
</player>
<country>
  <name>czech republic</name>
  <continent>europe</continent>
</country>
<country>
  <name>norway</name>
  <continent>europe</continent>
</country>
<date>
  <year>2006</year>
  <month>unknown</month>
  <day_of_month>unknown</day_of_month>
</date>
<date>
  <year>2005</year>
  <month>11</month>
  <day_of_month>12</day_of_month>
</date>
<event>World Cup 2006</event>
<match_type>match</match_type>
<played_in>oslo norway</played_in>
<final_score>1 : 0</final_score>

```

Informácie o týchto konceptoch sa následne ukladajú do atribútu Description IPTC hlavičiek JPG obrázku, ktorý vstupný text obklopoval.

## Reference

1. Drozdzyński W., Krieger H.-U., Piskorski J., Schäfer U., Xu F.. Shallow Processing with Unification and Typed Feature Structures - Foundations and Applications. *Künstliche Intelligenz*, 1/2004

### Annotation:

*Storing metadata into IPTC headers of sport images*

This paper describes the IPTC Manipulator developed at University Of Economics Prague, and various issues connected with annotating, extracting semantic concepts from text and storing information directly into images.