

Abstrakt

Získání a reprezentace (lexikální) sémantické informace patří mezi nejdůležitější úlohy, které přispívají ke kvalitě automatického zpracování přirozeného jazyka (viz např. nedávná studie (Turney a Pantel, 2010)). V této práci je navržen nový, inovativní (ve skutečnosti první svého druhu) experimentální systém pro vytváření multimodálních distribučních sémantických modelů na základě kombinace nejlepších existujících modelů pracujících na základě zpracování textů nebo obrazu.

Multimodální sémantický model testujeme pomocí úloh odhadu sémantické podobnosti, sémantického shlukování konceptů, a pomocí nedávno vyvinutého testu BLESS. Navrhujeme také efektivní algoritmus pro nalezení vhodných parametrů pro integraci multimodálních atributů za účelem získání robustnosti modelu. Naše experimenty ukazují, že navržená technika slibuje vysokou účinnost. Napříč různými experimenty se náš nejlepší multimodální model ukázal být na prvním místě. Při srovnání s jinými modely, které pracují pouze na základě analýzy textu, je náš navržený model mezi vůbec nejlepšími.

V práci jsme zkoumali různé druhy vizuálních deskriptorů včetně metody SIFT, aby bylo možno posoudit, jak by se mohly metody počítačového vidění uplatnit v oblasti zpracování přirozeného jazyka.

V práci ukážeme důležité zjištění, že přidávání vizuálních deskriptorů do sémantického modelu je srovnatelné, a možná i lepší než přidávání dalšího množství textových atributů do nejlepších modelů založených textově. Navíc se ukazuje, že vizuální deskriptory mohou lépe zachytit sémantickou charakteristiku konkrétních konceptů, a jsou tak komplementární k atributům extrahovaným z textů.

Klíčová slova: sémantika, vizuální, kombinace, zpracování obrazu