

Temat

Podobieństwa semantyczne w obrazowych bazach danych (CBIR)

Opiekun naukowy, dane kontaktowe opiekuna, miejsce prowadzenia badań

dr inż. Tatiana Jaworska (Tatiana.Jaworska@ibspan.waw.pl, tel. 22 3810223), IBS PAN, ul. Newelska 6, Warszawa

Opis pracy

Jedną z podstawowych funkcji obrazowych baz danych (CBIR) jest zwracanie właściwej informacji w odpowiedzi na zapytanie użytkownika. W tym przypadku właściwą informacją nie jest identyczny obraz, ale raczej obraz podobny w naszym (ludzkim) znaczeniu tego słowa. W ostatnich latach analizowane jest podobieństwo pomiędzy obrazami, ich fragmentami lub elementami obrazów, np. konturami. Modele podobieństwa kontekstowego są już rozwijane [1] jako odpowiedź na potrzeby użytkowników w zakresie wyszukiwania multimediów.

Splotowe sieci neuronowe (CNN), które obecnie się gwałtownie rozwijają, rozpoznają poszczególne obiekty na obrazie [2], takie jak np. samochody, czy ludzie, ale nie rozpoznają obiektów bardziej złożonych jak np. miasto czy tropiki.

Ponieważ semantyczne podobieństwo obrazów zawiera powiązane ze sobą koncepcje znaczeniowe, nawet jeśli wizualnie obrazy są mało podobne. Co za tym idzie, mnogość interpretacji obrazu powoduje znaczne problemy w jego wyszukiwaniu, ponieważ trzeba wziąć pod uwagę jednocześnie: złożoność obrazu, liczbę obiektów, rozkład przestrzenny obiektów, obrazowe cechy niskiego poziomu, a do tego ludzkie skojarzenia, które dany obraz wywołuje. Z tych powodów w IBS PAN zaczęto już pracować nad systemem [3] wyszukiwania oferującym użytkownikom możliwość przeszukiwania obrazowej bazy danych pod pewnymi kryteriami semantycznymi. Pozostaje jednak bardzo dużo do zrobienia w tym zakresie, ponieważ nadal nie istnieje metoda ilościowego porównywania obrazów pomimo prób zastosowania różnorodnych metryk [4].

Jeśli zadamy pytanie do CNN, to w odpowiedzi dostaniemy też prawdopodobieństwo z jakim sieć wyszukała dany obraz, ale to prawdopodobieństwo jest raczej zagregowaną miarą podobieństwa do wszystkich obrazów na jakich ta sieć była uczona. Tymczasem użytkownika znacznie bardziej interesuje podobieństwo tylko pomiędzy parą obrazów, czyli jego pytaniem i odpowiedzią systemu, które dodatkowo mają być podobne znaczeniowo (a mniej wizualnie).

Literatura

- [1] C. Beecks, M. S. Uysal i T. Seidl, "A Comparative Study of Similarity Measures for Content-Based Multimedia Retrieval," W: *Multimedia and Expo (ICME)*, Suntec City, 19-23.06, 2010.
- [2] Q. Abbas, M. E. A. Ibrahim i M. A. Jaffar, "A comprehensive review of recent advances on deep vision systems," *Artificial Intelligence Review*, vol. 52, str. 39-76, Maj on-line 2018.
- [3] T. Jaworska, "How to Compare Search Engines in CBIR?," W: *Proceedings of the SAI COMPUTING CONFERENCE 2016*, London, UK, 2016.
- [4] T. Jaworska, "An Asymmetric Approach to Signature Matching," in *Multimedia and Network Information Systems*, vol. 506, A. Zgrzywa, K. Choraś i A. Siemiński, Red., Wrocław, Springer, 2016, str. 27-37.