

NHẬN DẠNG VÀ TRUY VẤN ĐỐI TƯỢNG BA CHIỀU VỚI RING VIEW VÀ NEURAL EMBEDDING

Bùi Ngọc Minh, Đỗ Trọng Lễ, Nguyễn Vinh Tiệp, Trần Minh Triết

Khoa Công nghệ Thông tin,

Trường Đại học Khoa học Tự Nhiên, ĐHQG-HCM

1412314@student.hcmus.edu.vn, 1412673@student.hcmus.edu.vn,

nvtiep@fit.hcmus.edu.vn, tmtriet@fit.hcmus.edu.vn

Tóm tắt

Bài toán nhận dạng và truy vấn vật thể ba chiều (3D) dành được sự quan tâm. So sánh với bài toán nhận dạng và truy vấn ảnh 2D, bài toán nhận dạng vật thể 3D có nhiều thách thức vì có nhiều cách biểu diễn dữ liệu 3D, cũng như hướng của đối tượng không được chuẩn hóa. Trong công trình này, tác giả đề xuất phương pháp nhận dạng vật thể 3D bằng cách khai thác các hình chiếu 2D của vật thể 3D từ nhiều góc nhìn khác nhau. Tác giả khai thác tính thứ tự giữa các hình chiếu 2D trong một ring, trong khi không ép buộc tính thứ tự giữa tất cả các hình chiếu. Với cấu trúc Ring View đã được định nghĩa, chúng tôi đề xuất một mạng neural network để phân lớp các vật thể 3D. Phương pháp được đề xuất không chỉ giới hạn cho các đối tượng 3D toàn vẹn, mà cũng có thể phân lớp những đối tượng không toàn vẹn, thường bắt gặp từ dữ liệu thu thập bởi robot. Tác giả thí nghiệm với tập dữ liệu từ cuộc thi SHREC 2017 và 2018, phần “RGB-D to CAD retrieval”, cũng như với tập dữ liệu ModelNet40. Phương pháp đề xuất đạt độ chính xác truy vấn mAP 85.5% trong tập dữ liệu từ SHREC 2017, 85.4% trong tập dữ liệu từ SHREC 2018 và 91.13% trong tập dữ liệu ModelNet40, tương đương với các công trình liên quan trong lĩnh vực.

Từ khóa: Máy học, 3D Vision, Nhận dạng đối tượng, Truy vấn đối tượng

3D OBJECT CLASSIFICATION AND RETRIEVAL WITH RING VIEW AND NEURAL EMBEDDING

Ngoc-Minh Bui, Trong-Le Do, Vinh-Tiep Nguyen, Minh-Triet Tran
Faculty of Information Technology, University of Science, VNU-HCM
1412314@student.hcmus.edu.vn, 1412673@student.hcmus.edu.vn,
nvtiep@fit.hcmus.edu.vn, tmtriet@fit.hcmus.edu.vn

Abstract

The problem of 3D object classification and retrieval has attracted lot of attention. Comparing to 2D images classification and retrieval problems, the 3D objects recognition problem has some different challenges and is hard to solve because there are many different representation, also the orientation of the objects are not normalized. In this work, we propose a method to recognize 3D object by exploiting the 2D projections of the 3D objects from different viewpoints. We based on the topological combination between views in a ring, which has several views, while we do not enforce the topological relationship between of all views. With the predefined structure of view-rings, we propose a neural network to classify the 3D objects. The proposed method is not limited to recognize complete 3D objects, but has the ability to classify even incomplete objects, which are commonly captured by moving robots. We experiment our method with datasets from the SHREC2017 and SHREC2018 competition, track RGB-D to CAD retrieval, also with the ModelNet40 dataset. We achieve the mAP retrieval score of 85.5% on the dataset from SHREC2017, 85.4% on the dataset from SHREC2018 and 91.13% on the ModelNet40 dataset, which are comparable with another works in the field.

Key words: Machine Learning, 3D Vision, Object Classification, Object Retrieval