

SO SÁNH PHƯƠNG PHÁP NHẬN DẠNG HÀNH ĐỘNG CON NGƯỜI TRONG ĐOẠN VIDEO QUAY BẰNG MỘT CAMERA DÙNG DTW VÀ HMM

COMPARISON OF HUMAN ACTION RECOGNITIONS IN MONOCULAR VIDEOS USING DTW AND HMM

Tác giả: [Hoàng Lê Uyên Thục](#), [Phạm Văn Tuấn](#), [Shian](#), [Ru Ke](#)

Tóm tắt bằng tiếng Việt:

Trong bài báo này, chúng tôi tìm hiểu và so sánh hai thuật toán nhận dạng Dynamic Time Warping (DTW) và mô hình Markov ẩn HMM. Trước tiên, từ mỗi khung video, chúng tôi ước lượng tư thế người 3D, bao gồm tọa độ 3D của các khớp đặc trưng, dùng kỹ thuật mô hình hóa cơ thể 3D; rồi chuyển các tọa độ này sang thuộc tính quan hệ hình học GRF, mô tả quan hệ hình học giữa các khớp trong một tư thế nhằm giảm số hướng và gia tăng sự khác biệt giữa các tư thế. Tiếp đến, chúng tôi áp dụng kỹ thuật k-means clustering vào các GRF để tạo ra các vector thuộc tính, nhằm giảm số hướng hơn nữa. Cuối cùng, chúng tôi lần lượt sử dụng DTW và HMM để nhận dạng hành động và so sánh hiệu quả nhận dạng của chúng. Trong hệ thống, để nhận dạng các hành động lặp lại, chúng tôi sử dụng một biến thể của HMM gốc là HMM tuần hoàn CHMM. Các kết quả thực nghiệm trên cơ sở dữ liệu IXMAS cho thấy CHMM nổi trội hơn nhiều so với DTW.

Từ khóa: Nhận dạng hành động con người; mô hình hóa cơ thể 3D; thuộc tính quan hệ hình học; Dynamic Time Warping (DTW); mô hình Markov ẩn tuần hoàn

Tóm tắt bằng tiếng Anh:

In this paper, the use of two well-known recognition algorithms which are Dynamic Time Warping (DTW) and Hidden Markov Model (HMM) are studied and compared. From each frame in monocular videos, we first estimate the 3D human pose which consists of 3D coordinates of specific human joints using an efficient 3D human modeling technique; then convert them into a set of geometrical relational features (GRF), which describe the geometric relations among body joints of a pose for dimensionality reduction and discrimination increase. Next, the k-means clustering technique is applied to those GRFs to generate feature vectors for further dimensionality reduction. Finally, we use DTW and HMM in succession for recognition of actions and then compare their recognition performance. In our system, in order to recognize the repeated actions we use a variation of original HMM which is cyclic HMM (CHMM). The experiment results on IXMAS dataset show that CHMM stands out DTW in terms of recognition rate.

Key words: Human action recognition; 3D human modeling; geometrical relational feature; Dynamic Time Warping; Cyclic Hidden Markov Model