CHƯƠNG TRÌNH DÒ TÌM CHUYỂN ĐỘNG CỦA MÔI CÀI ĐẶT TRÊN FPGA

LIP TRACKING PROGRAM IMPLEMENTED ON FPGA

Tác giả: *Võ thi Thu Hồng, Lê Quốc Bảo Trí, Nguyễn Ngọc Tài, Lê Trung Hiếu*

Tóm tắt bằng tiếng Việt:

Việc dò tìm chuyển động của môi là giai đoạn mang tính quyết định đối với hệ thống nhận dạng thính thị (AVSR). Đây là quá trình phân tách đường viền môi từ các chuỗi video của khuôn mặt. Đã có nhiều giải thuật hữu hiệu được đề xuất để cho kết quả phân đoạn chính xác vùng môi. Do độ phức tạp trong tính toán nên đa số các giải thuật này đều được thực hiện hoàn toàn bằng phần mềm. Trong bài viết này, chúng tôi trình bày mô hình mới sử dụng phần cứng để dò tìm chuyển động môi, cài đặt trên FPGA. Công cụ toán học ở đây kết hợp giải thuật phân cụm K-Means và giải thuật tìm đường đi ngắn nhất của Dijkstra. Kết quả thử nghiệm thích hợp với khuôn mặt có sự xuất hiện của râu, răng và xử lý tốt trong trường hợp màu môi không khác biệt nhiều so với vùng ngoài môi. Cấu trúc phần cứng được thiết kế pipeline cho phép tăng đáng kể tốc độ xử lý so với phần mềm (125 Mpps).

Từ khóa: Dò chuyển động môi; phân đoạn vùng mô i; FPGA; phân cụm K-Means; giải thuật Dijkstra; pipeline

Tóm tắt bằng tiếng Anh:

The first and crucial stage in audio visual speech recognition (AVSR) system is lip tracking. It is the process of partitioning lip contour from facial video sequences. Many efficient algorithms have been developed to achieve good and accurate segmentation. Most of these have to be performed entirely by software because of their complex computation. In this paper, we present a novel hardware- based lip tracking model, implemented on DE2 FPGA (Field Programmable Gate Array) Board of Altera. Our mathematical tool is the combination of K-Means clustering and shortest path Dijkstra algorithms. The experimental results have proved that our proposed technique is well adapted to face with the presence of beard, teeth and even to lips with weak color contrast. The pipeline design in our hardware structure considerably enhances the processing speed more than 700 FPS) in comparison with software. (FPS: Frames per second).

Key words: Lip tracking; lip segmentation; FPGA; K-Means clustering; Dijkstra algorithm; pipeline.