**Ý kiến chỉnh sửa bài báo 615:**

**“TỔNG HỢP TIẾNG VIỆT CÓ CẢM XÚC”**

**1. Nội dung hiện tại trong bản 615-2826-2-CE – Phần tên tác giả đầu bài báo :**

Lê Xuân Thành1, Trịnh Văn Loan1, Nguyễn Hồng Quang1, Đào Thị Lệ Thủy1, Đinh Đồng Lưỡng2

1 Viện Công nghệ Thông tin và Truyền thông, Trường Đại học Bách Khoa Hà Nội.

2 Khoa Công nghệ thông tin, Trường Đại học Nha Trang.

***Đề xuất chỉnh sửa của nhóm tác giả:***

Lê Xuân Thành1, Trịnh Văn Loan1, Nguyễn Hồng Quang1, Đào Thị Lệ Thủy1,2, Đinh Đồng Lưỡng3

1 Viện Công nghệ Thông tin và Truyền thông, Trường Đại học Bách khoa Hà Nội.

2 Khoa Công nghệ thông tin, Trường Cao đẳng nghề Công nghệ cao Hà Nội.

3 Khoa Công nghệ thông tin, Trường Đại học Nha Trang.

**2. Nội dung hiện tại trong bản 615-2826-2-CE – Phần tóm tắt:**

Tiếng Việt là ngôn ngữ đơn âm tiết và có thanh điệu. Để tổng hợp tiếng Việt chất lượng tốt, việc đảm bảo chất lượng của thanh điệu tổng hợp sao cho càng gần với thanh điệu tự nhiên là rất quan trọng. Bài báo này đề xuất một phương pháp tổng hợp tiếng Việt dựa trên ghép nối âm vị kép, trong đó các biến thiên F0 của các âm được tổng hợp giống như biến thiên F0 của tiếng nói tự nhiên. Hơn nữa, để tích hợp cảm xúc vào tiếng Việt tổng hợp, bài báo trình bày một phương pháp tổng hợp dựa trên mô hình Fujisaki. Ba cảm xúc khác nhau được thử nghiệm là: buồn, tức giận và vui. Các kết quả đánh giá khách quan và chủ quan chất lượng tiếng Việt tổng hợp cũng được trình bày trong nghiên cứu này.

***Đề xuất chỉnh sửa của nhóm tác giả:***

Tiếng Việt là ngôn ngữ đơn âm tiết và có thanh điệu. Để tổng hợp tiếng Việt chất lượng tốt, việc đảm bảo chất lượng của thanh điệu tổng hợp sao cho càng gần với thanh điệu tự nhiên là rất quan trọng. Bài báo này đề xuất một phương pháp tổng hợp tiếng Việt dựa trên ghép nối âm vị kép, trong đó các biến thiên F0 của các âm được tổng hợp giống như biến thiên F0 của tiếng nói tự nhiên. Hơn nữa, để tích hợp cảm xúc vào tiếng Việt tổng hợp, bài báo trình bày một phương pháp tổng hợp dựa trên mô hình Fujisaki. Ba cảm xúc khác nhau được thử nghiệm là: buồn, tức và vui. Các kết quả đánh giá khách quan và chủ quan chất lượng tiếng Việt tổng hợp cũng được trình bày trong nghiên cứu này.

**3. Nội dung hiện tại trong bản 615-2826-2-CE – Phần Tài liệu tham khảo:**

[27]L. X. Thành, o. T. L. Thủy, T. V. Loan, and N. H. Quang, “So sánh hiệu năng một số phương pháp nhận dạng cảm xúc tiếng việt nói,” in *Kỷ yếu hội nghị khoa học công nghệ quốc gia lần thứ IX, Nghiên cứu cơ bản và ứng dụng công nghệ thông tin*, 2016.

[28] o. T. L. Thủy, T. V. Loan, Q. N. Hồng, and L. X. Thành, “Ảnh hưởng của đặc trưng phổ tín hiệu tiếng nói đến nhận dạng cảm xúc tiếng Việt,” in *Kỷ yếu hội nghị khoa học công nghệ quốc gia lần thứ X, Nghiên cứu cơ bản và ứng dụng công nghệ thông tin*, 2017.

[20] Đ. T. T. Lê Xuân Thành, T. V. Loan, and N. H. Quang, “Cảm xúc trong tiếng nói và phân tích thống kê ngữ liệu cảm xúc tiếng việt,” *Chuyên san Các công trình Nghiên cứu, Phát triển và Ứng dụng Công nghệ Thông tin; Tạp chí Bưu chính Viễn thông*, pp. 86–98, 2016.

***Đề xuất chỉnh sửa của nhóm tác giả:***

[27] Lê Xuân Thành, Đào Thị Lệ Thủy, Trịnh Văn Loan và Nguyễn Hồng Quang, “So sánh hiệu năng một số phương pháp nhận dạng cảm xúc tiếng việt nói,” in *Kỷ yếu hội nghị khoa học công nghệ quốc gia lần thứ IX, Nghiên cứu cơ bản và ứng dụng công nghệ thông tin*, trang 656-662, 2016.

[28] Đào Thị Lệ Thủy, Trịnh Văn Loan, Nguyễn Hồng Quang và Lê Xuân Thành, “Ảnh hưởng của đặc trưng phổ tín hiệu tiếng nói đến nhận dạng cảm xúc tiếng Việt,” in *Kỷ yếu hội nghị khoa học công nghệ quốc gia lần thứ X, Nghiên cứu cơ bản và ứng dụng công nghệ thông tin*, trang 36-43, 2017.

[20] Lê Xuân Thành, Đào Thị Lệ Thủy, Trịnh Văn Loan và Nguyễn Hồng Quang, “Cảm xúc trong tiếng nói và phân tích thống kê ngữ liệu cảm xúc tiếng việt,” *Chuyên san Các công trình Nghiên cứu, Phát triển và Ứng dụng Công nghệ Thông tin; Tạp chí Bưu chính Viễn thông*, tập V-1, số 15 (35), trang 86–98, 2016.

**4. Nội dung hiện tại trong bản 615-2826-2-CE – Phần Giới thiệu các tác giả:**

(1) Phần giới thiệu tác giả Nguyễn Hồng Quang: Nguyễn Hồng Quang nhận bằng Kỹ sư Công nghệ Thông tin và Thạc sỹ ngành Xử lý thông tin và Truyền thông vào các năm 2000 và 2004 tại Đại học Bách Khoa Hà Nội. Năm 2008, ôngnhận bằng Tiến sỹ ngành Công nghệ thông tin tại trường Đại học Avignon, Pháp. Từ năm 2000 tới nay, ônglà ~~Giảng~~ giảng viên Viện Công nghệ Thông tin và Truyền thông, ~~trường~~ *Trường* Đại học Bách Khoa Hà Nội. Hướng nghiên cứu quan tâm *của ông là* xử lý tiếng nói, học máy và học sâu.

***Đề xuất chỉnh sửa của nhóm tác giả:***

(1) Phần giới thiệu tác giả Nguyễn Hồng Quang: Nguyễn Hồng Quang nhận bằng Kỹ sư Công nghệ Thông tin và Thạc sỹ ngành Xử lý thông tin và Truyền thông vào các năm 2000 và 2004 tại Đại học Bách Khoa Hà Nội. Năm 2008, ôngnhận bằng Tiến sỹ ngành Công nghệ thông tin tại trường Đại học Avignon, Pháp. Từ năm 2000 tới nay, ông là giảng viên Viện Công nghệ Thông tin và Truyền thông, Trường Đại học Bách khoa Hà Nội. Hướng nghiên cứu quan tâm của ông là xử lý tiếng nói, học máy và học sâu.

**5. Nội dung hiện tại trong bản 615-2826-2-CE – Phần nội dung:**

(1) Các đề xuất chỉnh sửa cảm xúc "tức" thành "tức giận"

(2) Đoạn *"…không thuần túy phân tích ảnh hưởng của các tham số như* F*0, thời hạn, năng lượng đến cảm xúc mà kết hợp giữa tiếng nói và các dữ liệu hình ảnh, hoặc là các kết quả nghiên cứu thực hiện trên các bộ ngữ liệu còn hạn chế về số lượng cũng như…"* ở trang 1 phần I. GIỚI THIỆU

(3) Đoạn *"Tiếp theo, sử dụng kết quả xây dựng bộ ngữ liệu cảm xúc tiếng Việt (BKEmo [20]) để xây dựng mô hình tổng hợp tiếng Việt có cảm xúc bằng cách điều chỉnh các tham số thời hạn ~~,~~ và cường độ với công cụ Praat [21], và điều chỉnh quy luật biến thiên F0 theo mô hình Fujisaki."* ở trang 2 phần I. GIỚI THIỆU

(4) Các đề xuất của ban biên tập về chỉnh sửa các lỗi chính tả cũng như cách trình bày câu chữ ở các phần còn lại

***Đề xuất chỉnh sửa của nhóm tác giả:***

(1) Nhóm tác giả đề nghị giữ nguyên cảm xúc "tức" để mô tả đúng cảm xúc đã thể hiện trong các nghiên cứu và công bố của nhóm tác giả.

(2) Đoạn *"…không thuần túy phân tích ảnh hưởng của các tham số như* F*0, thời hạn, năng lượng đến cảm xúc song lại kết hợp giữa tiếng nói và các dữ liệu hình ảnh, hoặc là các kết quả nghiên cứu thực hiện trên các bộ ngữ liệu còn hạn chế về số lượng cũng như…"* ở trang 1 phần I. GIỚI THIỆU

(3) Đoạn *"Tiếp theo, sử dụng kết quả xây dựng bộ ngữ liệu cảm xúc tiếng Việt (BKEmo [20]) để xây dựng mô hình tổng hợp tiếng Việt có cảm xúc bằng cách điều chỉnh các tham số thời hạn, cường độ với công cụ Praat [21] và điều chỉnh quy luật biến thiên F0 theo mô hình Fujisaki."* ở trang 2 phần I. GIỚI THIỆU

(4) Nhóm tác giả xin tiếp thu toàn bộ đề xuất chỉnh sửa của ban biên tập trong các đề xuất này

**Xin trân trọng cảm ơn!**