

Tính hệ số tương quan giữa các mẫu liên kề trong tín hiệu âm thanh

Hệ số tương quan đo lường mức độ phụ thuộc tuyến tính giữa hai biến. Trong tín hiệu âm thanh, các mẫu liên kề thường có sự tương quan cao trong tín hiệu gốc nhưng sẽ giảm đáng kể sau khi được mã hóa.

1. Khái niệm về hệ số tương quan

Hệ số tương quan r giữa hai biến ngẫu nhiên X và Y được tính theo công thức. Hệ số tương quan có giá trị trong khoảng từ 1 đến -1. Các biến không có sự tương quan nếu hệ số tương quan giảm xuống bằng 0. Chúng ta có thể chọn ngẫu nhiên 2048 cặp mẫu liên kề trong cả tín hiệu âm thanh gốc và âm thanh đã mã hóa liên kề để đánh giá hệ số tương quan bằng cách sử dụng các công thức:

$$r_{xy} = \frac{cov(x,y)}{\sqrt{D(x)}\sqrt{D(y)}} \quad (1)$$

$$E(x) = \frac{1}{L} \sum_{j=1}^L x_j, E(y) = \frac{1}{L} \sum_{j=1}^L y_j \quad (2)$$

$$D(x) = \frac{1}{L} \sum_{j=1}^L (x_j - E(x))^2, \\ D(y) = \frac{1}{L} \sum_{j=1}^L (y_j - E(y))^2 \quad (3)$$

$$cov(x,y) = \frac{1}{L} \sum_{j=1}^L (x_j - E(x))(y_j - E(y)) \quad (4)$$

3. Bảng hệ số tương quan

Ta có thể đánh giá hệ số tương quan giữa các mẫu liên kề trong âm thanh gốc và âm thanh mã hóa bằng thuật toán DNA như sau:

Loại tín hiệu	Hệ số tương quan (Original Audio)	Hệ số tương quan (Encrypted Audio)
Tín hiệu 1	0.879	0.007

Từ bảng trên cho thấy giá trị tương quan giữa các mẫu liên kề trong tín hiệu âm thanh khi đã mã hóa sử dụng thuật toán DNA là