

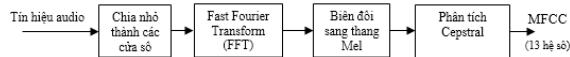
e. Đặc trưng 5: Low-Energy

Khác với các đặc trưng trên, đặc trưng Low-Energy được xác định trên toàn bộ tín hiệu miền thời gian. Nó là tỉ lệ phần trăm của các cửa sổ phân tích có RMS (Root-Mean-Square) năng lượng thấp hơn RMS trung bình năng lượng của các tín hiệu trong các cửa sổ phân tích. Trong đó, RMS năng lượng của tín hiệu ở cửa sổ t được xác định bởi công thức (10):

$$RMS_t = \sqrt{\frac{\sum_{i=1}^N (M_t[i]^2)}{N}} \quad (10)$$

f. Đặc trưng 6: Các hệ số MFCC (Mel-Frequency Cepstral Coefficients)

MFCC là một trong các tập đặc trưng được dùng phổ biến trong các hệ thống nhận dạng giọng nói, truy tìm thông tin nhạc,... Nó cung cấp cách biểu diễn nén tín hiệu audio dưới dạng phổ sao cho hầu hết năng lượng của tín hiệu được tập trung vào các hệ số đầu tiên. Hình 4 mô tả các bước thực hiện rút trích đặc trưng MFCC từ tín hiệu audio. Chi tiết về phương pháp rút trích đặc trưng MFCC (Logan and Beth, 2000) mô tả trong Hình 5.



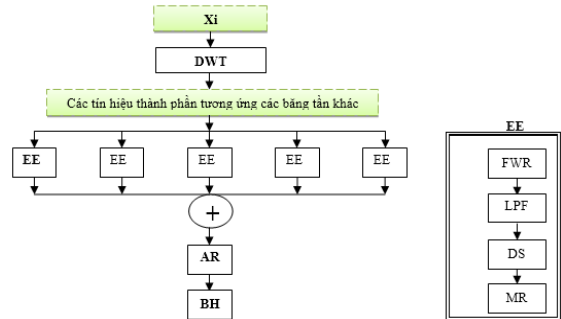
Hình 5: Sơ đồ rút trích đặc trưng MFCC

Kết quả thu được là một tập đặc trưng MFCC gồm 13 hệ số. Tuy nhiên, nhiều nghiên cứu (Li *et al.*, 2003) cho thấy 5 hệ số MFCC đầu tiên cung cấp khá đầy đủ thông tin cho việc phân loại nhạc theo thể loại. Vì vậy, để giảm số chiều cho vector đặc trưng, chúng tôi chọn 5 hệ số MFCC đầu tiên cho hệ thống phân loại nhạc theo thể loại của chúng tôi.

3.2 Đặc trưng về nhịp điệu nhạc

Vector đặc trưng về nhịp điệu cung cấp rất nhiều thông tin có ích về đặc điểm của các thể loại nhạc. Hầu hết các hệ thống dò tìm nhịp điệu nhạc cung cấp các thuật toán xác định nhịp điệu của bản nhạc và cường độ của chúng. Bên cạnh đó, chúng còn cho biết mối liên hệ giữa các nhịp của bản nhạc. Trong bài báo này, chúng tôi sử dụng phương pháp xác định tập đặc trưng về nhịp điệu nhạc được đề xuất bởi George Tzanetakis (Tzanetakis *et al.*, 2001) trong việc phân loại nhạc theo thể loại. Phương pháp này dựa trên việc dò tìm các chu kỳ (đơn vị: bpm - số nhịp/phút) có biên độ lớn nhất của tín hiệu. Tín hiệu audio X được chia nhỏ thành các tín hiệu thành phần X_i bởi cửa sổ phân tích có kích thước 65536 mẫu với tần số lấy mẫu (sampling rate) là 22050 Hz tương ứng xấp xỉ 3s. Sau đó, thuật toán xác định nhịp điệu nhạc được áp dụng đối với mỗi X_i như biểu diễn trong Hình 5.

Quá trình xác định nhịp điệu nhạc trên tín hiệu audio được áp dụng lặp đi lặp lại trên các tín hiệu thành phần X_i và tích lũy vào trong biểu đồ nhịp điệu BH. Tập các đỉnh cao nhất của hàm tự tương quan tạo nên biểu đồ nhịp điệu nhạc được sử dụng làm cơ sở cho việc xác định các đặc trưng về nhịp điệu. Trong đó, các đỉnh cao nhất trong BH tương ứng với các chu kỳ khác nhau của tín hiệu audio là các nhịp chính của bản nhạc.

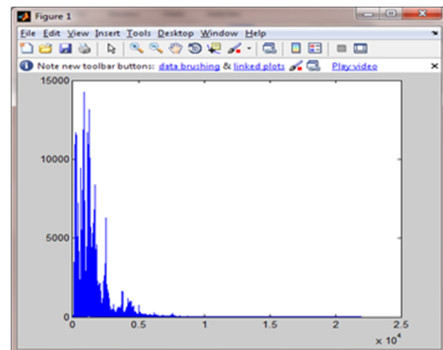


Hình 6: Sơ đồ khối xác định Histogram nhịp điệu nhạc

Xác định các đặc trưng về nhịp điệu:

Vector đặc trưng về nhịp điệu là một vector 6 chiều gồm các đặc trưng:

- A1, A2: Đặc trưng này là độ đo sự khác nhau về nhịp so với các nhịp còn lại của tín hiệu. Nó được xác định bởi tỉ số giữa biên độ của lần lượt 2 đỉnh Đ1 và Đ2 với tổng biên độ của tất cả các đỉnh trong BH.
- RA: là tỉ số giữa biên độ của đỉnh Đ2 với biên độ của đỉnh Đ1. Đặc trưng này biểu diễn mối quan hệ giữa nhịp chính và nhịp phụ đầu tiên.
- P1, P2: Chu kỳ của đỉnh Đ1 và Đ2 được tính bằng số nhịp trong 1 phút (đơn vị tính: bpm).
- SUM: Tổng biên độ của các đỉnh trong BH. Đặc trưng này cho biết độ mạnh của nhịp nhạc.



Hình 7: Quang phổ về âm thanh của 1 bản Chèo