On the development of an algorithm to recognize music patterns

Music is the only artistic discipline with purely mathematical roots, and therefore the seven musical notes have been sought to please human ears with numeric relationships. Then, notes are no more than a combination of many harmonics, each of them with different amplitudes and frequencies. We can therefore imagine a song as a three-dimensional spectrogram in which for each time instant every frequency has an amplitude associated to it. In this context, an algorithm capable of identifying short segments of music through frequency-domain techniques, while removing any source of distortion and environmental noise, and comparing it to a large database of analyzed musical pieces would be extremely useful. The latter is actually what *Shazam* application does. Starting from the above-mentioned spectrogram, a sparse set of points are extracted from it by detecting local peaks of energy with a criterion that allows a uniform coverage and representation of the spectrogram. Moreover, selecting at each time instant the harmonics with the highest amplitudes allows to overcome the mentioned distorting signals. Anyway, how does an algorithm that needs to be executed in a mobile phone store such music database and compare a recording with it to give a result in a matter of seconds? This is the question that the founder of *Shazam* tried to work out. The solution that was reached consisted of defining a series of reference points (so-called anchor points) and a target zone associated to each one of them [1]. Subsequently, each anchor point is paired with each point inside its target zone, to end up creating a vector of three components with the frequency of the two points plus the time difference between them. The purpose of the project would then be to create a similar algorithm.

References

[1] A. Li-Chun Wang, *An Industrial-Strength Audio Search Algorithm*. Shazam Entertainment, Ltd. .