

Using User Models in Music Information Retrieval Systems

Wei Chai and Dr. Barry Vercoe
Machine Listening Group • MIT Media Lab

Abstract

To make multimedia data easily retrieved, we use metadata to describe the information, so that search engines or other information filter tools can effectively and efficiently locate and retrieve the multimedia content. Since many features of multimedia content are perceptual and user-dependent, user modeling is also necessary for multimedia information retrieval systems, e.g., music information retrieval systems. Furthermore, to make the user models sharable, we need standardized language to describe them. In this paper, we explore the benefits, main issues and paradigms of user modeling for music information retrieval systems. An XML-like language is also proposed to describe the user model for music information retrieval purposes.

Benefits, main issues and paradigms

Most websites providing music services only support category-based browsing and/or text-based searching. For content-based search or feature-based filtering systems, one important problem is to describe music by its parameters or features, so that search engines or information filtering agents can use them to measure the similarity of the target (user's query or preference) and the candidates. MPEG7 is an international standard, which describes the multimedia content data to allow universal indexing, retrieval, filtering, control, and other activities supported by rich metadata. However, the metadata about the multimedia content itself are still insufficient, because many features of multimedia content are quite perceptual and user-dependent. For example, emotional features are very important for multimedia retrieval, but they are hard to be described by a universal model since different users may have different emotional responses to the same multimedia content. We therefore turn to user modeling techniques and representations to describe the properties of each user, so that the retrieval will be more accurate. Besides, user modeling can be used to reduce the search space, make push service easier and improve the user interface.

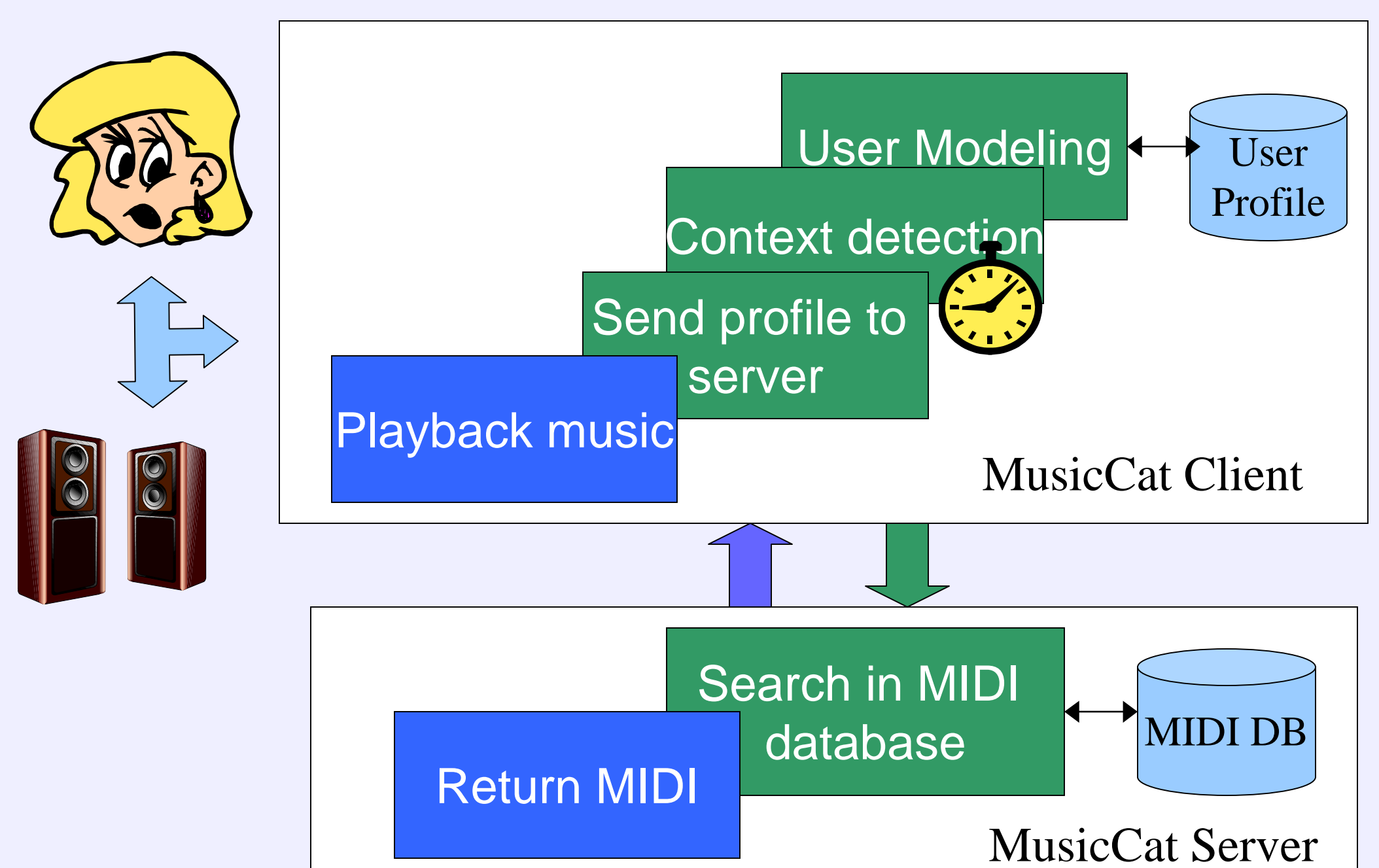
There are several important issues in user modeling for music information retrieval purpose or even more general multimedia retrieval. 1) How to model the user? User-programmed, machine-learning and knowledge-engineered methods can be used. 2) What information is necessary to describe a user for music IR purpose? It may include both the user's indirect information (e.g. age, gender, citizenship, education, music experience, etc.) and direct information (e.g. user's interests, definition of qualitative features, appreciation habits, etc.). 3) How to represent, use and share the user model? Similar to MPEG7 concepts, we can use a standard language to represent the user model, so that search engines or information filtering agents can use it to refine the result easily and efficiently, without repeating the long-time observation and learning of the user's behavior.

User modeling can be done on client-side or server-side. Issues including easy/hard to obtain the user information, hard/easy to

use collaborative filtering model, far from/close to the music data, more/less privacy or safety, more scalable/higher load on the server, etc., need to be considered when choosing either of the paradigms.

Example and conclusion

We adopt the client-side user modeling paradigm in our MusicCat system. It is an agent that allows the user to define contexts and the features of music that is preferred in those contexts correspondingly. Besides, the user can also define qualitative features of music based on quantitative features. For examples, the user just needs to tell the agent what kind of music he prefers in what kind of context, like "I need fast and exciting music when I'm happy", "I need soft music to wake me up every morning at 8:00", "I need slow classical music, when I'm thinking", "I need rhythmic music when I'm walking", etc. When the moment comes, the agent can automatically, randomly and repeatedly choose music pieces from the user's collection according to the pre-defined constraints. So far, MusicCat uses a profile-based user interface; and only MIDI files are used in the system.



In our system, we categorize the quantitative features of music into textual features, notation features, perceptual features (slow or fast, short or long, quiet or loud, non-rhythmic or rhythmic, and soft or exciting). All these features (except textual features) are computed directly from the MIDI files. Users can then define music preference corresponding to a particular context or qualitative features based on these quantitative features. The above are stored in the user's profile. Whenever a pre-defined context is activated, the client sends the user's profile to the server and the server finds the appropriate music for the user.

User information is very valuable and needs to be shared in the future to make universal information retrieval possible. Thus we propose an XML-like language, UMIRL (User Modeling for Information Retrieval Language), in which different systems may describe the user in this standard format to make the user model sharable and reusable. We believe that both user modeling techniques and descriptions are prerequisites for open and efficient personalized services.