Chapter 7

Conclusions and future work

For the last decade, the possibilities for Music Information Retrieval (MIR) research have been much better than fourty years ago, when Kassler gave his talk [31] about an assembler-like language called MIR. At Kassler's time, using computers for processing audio signals was so time-consuming that he did not even mention this possibility in his optimistic talk. The abundance of cheap storage space and processing power is not just a gradual, but a qualitative change since it allows a whole new range of problems to be addressed. This is probably an important reason why the last ten years have seen so much more progress than the previous thirty.

Part of that progress was the development of new ways of measuring melodic similarity. In the late nineties, most publications about query-by-humming systems described string-based methods for comparing melodic contours by using editing distances. While it is possible to efficiently find nearest neighbours when using such distance measures, for example with vantage indexing, using such distance measures for polyphonic music it is rather awkward, and these methods also do not naturally support ornamentations very well where single notes need to be matched to multiple notes.

Early geometric algorithms that are able to match arbitrary groups of notes, possibly polyphonic, to other groups of notes, suffered from a lack of suitable indexing methods. Also, distance measures that rely on counting coinciding note onsets are usually not continuous and work well only for quantized music.

This book shows that by using transportation distances, one can overcome many of the limitations of both other geometric algorithms and of string-based methods. Transportation distances can be designed such that they are continuous and support the triangle inequality. At MIREX 2006, it was shown that the result quality, response times, and scalability of our method compare favourably with those for string-based methods or other state-of-the-art geometric algorithms. However, especially for collections with a noticeable continuum of melodic similarity such as the RISM UK collection of musical incipits, the overlap of relevant items found by different algorithms was not very large, and therefore there is still quite some room for improvement.

To make methods for measuring melodic similarity more useful, it would be desirable to combine audio feature extraction methods with symbolic approaches like the distance measures that are the topic of this book. One of the big open problems of MIR, automatic polyphonic audio-to-MIDI transcription, is probably going to remain unsolved for quite a while. However, it is conceivable that with transportation

distances, even imperfect transcription results could be made searchable in a useful way thanks to the robustness against various types of errors one can achieve with transportation distances.

Transportation distances could be further improved in various ways, for example:

- To take full advantage of variants of transportation distances that obey the triangle inequality, but are only suitable for monophonic music, such as the Proportional Transportation Distance, one should work on better voice splitting algorithms and on recognizing which voices are only accompaniment and therefore not worth searching for melodies.

- To still be able to use vantage indexing for distance measures with desirable properties for measuring melodic similarities, but with the disadvantage of not guaranteeing that the triangle inequality always holds, it could be interesting to experiment with combinations of multiple vantage indices (for lowering the number of false negatives) and with vantage objects that are constructed with the aim of either making violations of the triangle inequality less likely, or at least very different from those encountered with other vantage objects.

If the recent trends in MIR research continue, we should have very nice retrieval algorithms at our disposal fourty years from now. And maybe even the pesky old Optical Music Recognition problem which Kassler thought would just require a million dollars and negligible time will eventually be solved as well.