change

is

projects a given melody onto a sequence of symbols. As an example to a quantization with 3 symbols ~~can be~~ a function that projects a melody onto the alphabet {'up','down','repeat'}. We observe that reducing a melody into a simpler melody , results in loss of result quality , though often not significant enough for success of retrieval. Where memory size plays a significant role , quantization can be a preferable process.

weighting scheme	ADR	AP	R-P
$3H, MH, 3M$	0.67	0.64	0.56
$4H, MH, 3M$	0.65	0.63	0.56
$7H, MH, 3M$	0.65	0.64	0.56
$3H, MH, 4M$	0.67	0.64	0.55
$3H, MH, 7M$	0.61	0.60	0.51
$3H, MS, 3M$	0.66	0.63	0.52

Fig. 5. Different segmentation techniques with no quantization [4]

for

Figure 2 shows use of different weight measures. As mentioned in section 2.1 weight measures are important ~~when~~ choosing the more important note among other notes in a measure. For harmonic weight the authors experimented with different alphabets size 3,4 and 7. An example to an alphabet of size 3 would consist in grouping the first , fourth and sixth degrees to a group, grouping second and seventh to another and rest to the third group. How these groupings are formed is based on Music Theory. The projection of a single note in the starting melody is based upon its harmonic meaning within the musical context. While reducing the different harmonic functionalities into one larger group, it is important to make assumptions, that result in a grouping with the least amount of information loss. With the same consideration, melodic weights have been grouped together in terms of their intervallic function. As to metric weight, there are two different schemes ~~that are proposed~~ : (1) A simple subdivision in terms of strong and weak beats and (2) *"a hierarchical organization depending on the position in the measure"*. It can be observed that, even though the differences are not significant ~~enough~~, the better results are obtained when weighting is based on more generalizing schemes.

The generalization processes that graph based algorithm ~~in Section 2.1~~ use are highly dependent upon annotations of harmonic functions of melodic data , which are rarely included in symbolic melodic data. Authors see this as a drawback since a false assumption of a harmonic function of a note in input could easily result in a false segmentation of a piece within. The tree that is formed as musical data are added to it has a tendency to grow in sublinear fashion ~~as argues by the authors.~~ We see this sublinear tendency as an advantage for memory consumption.

✓

1

?

① Was? Wer? Wie genau?

lead to

?

② Höhe oder Größe? Jedes Segment ist [illegible]

+ Ref

4.2 Urbano Melody-Shape

In the initial paper of 2011, where Urbano et al. first proposed to compare melodies based on their shape, the authors tested their implementation ~~against~~ the RISM A/II collection from MIREX 2005 and compared their results to the results of the competing algorithms from 2005 [8]. As can be seen in Fig 6, they performed best in 5 out of 11 queries and averaged the highest ADR score overall. The algorithm ~~used~~ was later improved ~~upon~~ and very similar to the time system

using

Welche Fälle ist "they"?

Query	Splines	GAM	O	US	TWV	L(P3)	L(DP)	FM
190.011.224-1.1.1	0.803	0.820	0.717	**0.824**	0.538	0.455	0.547	0.443
400.065.784-1.1.1	**0.879**	0.846	0.619	0.624	*0.861*	0.614	*0.839*	0.679
450.024.802-1.1.1	**0.722**	0.450	0.554	0.340	0.554	0.340	0.340	0.340
600.053.475-1.1.1	***0.911***	*0.883*	***0.911***	***0.911***	0.725	0.661	0.650	0.567
600.053.481-1.1.1	**0.630**	0.293	0.629	0.486	0.293	0.357	0.293	0.519
600.054.278-1.1.1	0.810	0.674	0.785	**0.864**	0.731	0.660	0.527	0.418
600.192.742-1.1.1	0.703	**0.808**	**0.808**	0.703	**0.808**	0.642	0.642	***0.808***
700.010.059-1.1.2	0.521	0.521	0.521	0.521	0.521	**0.667**	0.521	0.521
700.010.591-1.4.2	0.314	**0.665**	0.314	0.314	0.314	0.474	0.314	0.375
702.001.406-1.1.1	0.689	0.566	**0.874**	0.675	0.387	*0.722*	0.606	0.469
703.001.021-1.1.1	**0.826**	0.730	0.412	0.799	0.548	0.549	0.692	0.561
Average	**0.710**	0.660	0.650	0.642	0.571*	0.558***	0.543**	0.518***

Fig. 6. Spline based approach tested against the evaluation set of the MIREX 2005 competition and compared to the other contestants. [8]

we presented earlier, using the area between two splines and a local alignment approach to determine similarity. Based upon this system Urbano developed the ShapeH and Time systems and competed in the 2010-2015 editions of the MIREX Competition, with all submitted systems always placing in the top spots.

+ Ref

	ShapeH	ShapeL	ShapeG	ShapeTime	Time
ADR	0.609 (3)	0.483 (5)	0.542 (4)	0.671 (1)	0.657 (2)
NRGB	0.534 (3)	0.428 (5)	0.471 (4)	0.579 (1)	0.567 (2)
AP	0.532 (2)	0.273 (5)	0.418 (4)	0.541 (1)	0.487 (3)
PND	0.524 (1)	0.327 (5)	0.446 (4)	0.516 (2)	0.487 (3)
Fine	0.629 (2)	0.496 (5)	0.546 (4)	0.635 (1)	0.626 (3)
PSum	0.680 (2)	0.467 (5)	0.582 (4)	0.685 (1)	0.663 (3)
WCSum	0.629 (2)	0.391 (5)	0.532 (4)	0.636 (1)	0.609 (3)
SDSum	0.603 (2)	0.353 (5)	0.508 (4)	0.611 (1)	0.582 (3)
Greater0	0.833* (1)	0.693 (5)	0.730 (4)	0.833* (1)	0.827 (3)
Greater1	0.527 (2)	0.240 (5)	0.433 (4)	0.537 (1)	0.500 (3)
Median rank	2	5	4	1	3

Fig. 7. Results for the submissions by Urbano to the 2012 MIREX Competition [10]

Verweis?

	ShapeH		ShapeTime		Time	
ADR	0.734	(3)	0.794	(2)	0.798	(1)
NRGB	0.697	(3)	0.756	(1)	0.744	(2)
AP	0.690	(3)	0.708	(1)	0.694	(2)
PND	0.719	(1)	0.706	(2)	0.688	(3)
Fine	0.656	(1)	0.655	(2)	0.645	(3)
PSum	0.722	(1)	0.718	(2)	0.715	(3)
WCSum	0.673	(2)	0.676	(1)	0.668	(3)
SDSum	0.649	(2)	0.654	(1)	0.644	(3)
Greater0	0.867	(1)	0.847	(3)	0.857	(2)
Greater1	0.577	(2)	0.590	(1)	0.573	(3)
Median rank	2		1.5		3	

Fig. 8. Results for the submissions by Urbano to the 2013 MIREX Competition [9]

	ShapeH		ShapeTime		Time	
NRGB	0.679	(3)	0.749	(2)	0.760	(1)
AP	0.734	(3)	0.753	(2)	0.799	(1)
PND	0.736	(3)	0.744	(2)	0.761	(1)
Fine	0.538	(2)	0.546	(1)	0.512	(3)
PSum	0.558	(3)	0.565	(1)	0.563	(2)
WCSum	0.501	(3)	0.504	(2)	0.514	(1)
SDSum	0.473	(3)	0.474	(2)	0.490	(1)
Greater0	0.730	(2)	0.747	(1)	0.710	(3)
Greater1	0.387	(2)	0.383	(3)	0.417	(1)
Median rank	3		2		1	

Fig. 9. Results for the submissions by Urbano to the 2014 MIREX Competition [2]

However although the systems by Urbano seem to be performing best in the competition, the results themselves vary from year to year with the algorithms not really having changed, as Urbano himself points out in his paper to the 2014 competition [8]. While the ShapeTime system, as mentioned in ~~chapter~~ 2.2, usually achieved the best results on average, it was outperformed by the time system in 2014 (Fig.9). Scores in single categories also varied with the ShapeH system achieving an ADR score of 0.609 in the 2012 competition (Fig.7) and achieving a a score of 0.734 in the same category the following year (Fig.8).

5 Discussion

~~In general~~ there still seem to be several problems in the field of symbolic melodic music similarity. There is no definitive definition for what exactly constitutes music similarity and no agreed upon data set existing, against which methods can be evaluated. This makes comparing different methods challenging, if not impossible. Even the results from different years of the MIREX competitions cannot really be compared to each other, as the scoring appears to be inconsistent. Furthermore the used queries and corresponding Ground Truth have not been released since 2010, effectively making comparing a new method to the results from the competition impossible. There are also only limited real world applications for algorithms that only compare monophonic pieces, as most music

Selbstständigkeitserklärung

April 15, 2020

Hiermit erkläre ich, dass ich die vorliegende Arbeit selbstständig und nur unter Zuhilfenahme der angegebenen Quellen und Hilfsmittel verfasst habe.

Name	Datum	Unterschrift
Ali Bektas	15.04.2020	