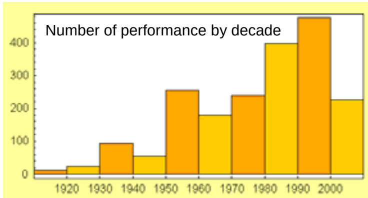
Similarity Measurements in Chopin Mazurka Performances

Craig Stuart Sapp
C4DM Seminar, 11 July 2007
Queen Mary, University of London

Mazurka Project

- **2,210 recordings of 49 mazurkas**
  = 45 performances/mazurka on average
  ***least*: 31 performances of 41/3**
  ***most*: 64 performances of 63/3**
- **105 performers on 160 CDs, 98 hours of music**
- **Earliest is 1907 Pachmann performance of 50/2**



Performers of mazurka 63/3:

Afanassiev (2001)
Anderszewski (2003)
Ashkenazy (1981)
Biret (1990)
Blet (2003)
Block (1995)
Blumental (1952)
Boshniakovich (1969)
Brailowsky (1960)
Bunin (1987)
Chiu (1999)
Cohen (1997)
Cortot (1951)
Czerny-Stefańska (1949)
Ezaki (2006)
Falvay (1989)
Ferenczy (1958)
Flière (1977)
François (1956)
Friedman (1923)
Friedman (1930)
Gierżod (1998)
Gornostaeva (1994)
Harasiewicz (1955)
Hatto (1988)
Horowitz (1949)
Indjic (1988)
Kapell (1951)
Kissin (1993)
Kushner (1989)
Luisada (1991)
Lushtak (2004)
Magaloff (1978)
Magin (1975)
Michałowski (1933)
Milkina (1970)
Mohovich (1999)
Moravec (1969)
Neighaus (1950)
Osinska (1989)
Pachmann (1927)
Paderewski (1930)
Perlemuter (1992)
Pobłocka (1999)
Rabcewiczowa (1932)
Rachmaninoff (1923)
Rangell (2001)
Rosen (1989)
Rosenthal (1931)
Rubinstein (1939)
Rubinstein (1952)
Rubinstein (1966)
Schilhawsky (1960)
Shebanova (2002)
Smith (1975)
Ts'ong (1984)
Uninsky (1932)
Uninsky (1971)
Wasowski (1980)
Zak (1937)

http://mazurka.org.uk/info/discography

Expressive Audio Features (Piano)

Note Attack

Note Loudness

Note Release

- Not much else a pianist can control
- String instruments have more control variables
- Voice has even more…

Expressive Performance Features

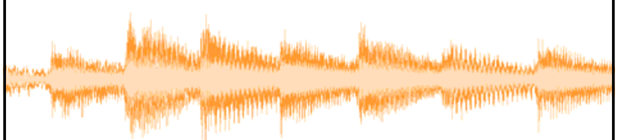
Note Attack → beat durations, offbeat timings, individual notes
- beat durations → meter, phrasing, tempo
- offbeat timings → tempo
- individual notes → ornamentation, articulation, hand synchrony

Note Loudness → dynamics, articulation (accents)
- dynamics → meter, phrasing

Note Release → articulation (staccato, legato), pedaling

Data Extraction (1)

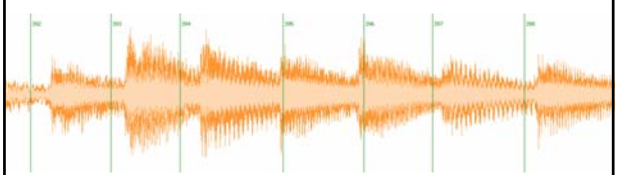
- Extracting beat times from audio files



- Using Sonic Visualiser for data entry processing
  http://www.sonicvisualiser.org

Data Extraction (2)

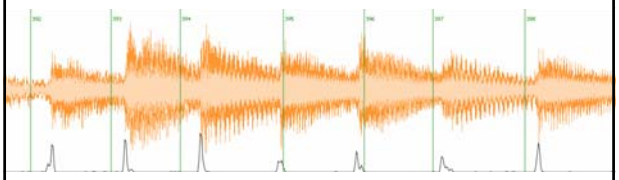
- Step 1: Listen to music and tap to beats (; key)



- Notice taps do not fall on the audio attacks:
  - 23.22 ms hardware granularity built into program
  - Human: ~30 ms SD for constant tempo; ~80 ms SD for mazurkas

Data Extraction (3)

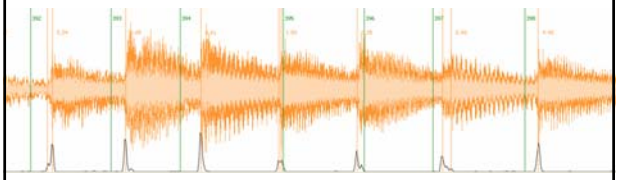
• Step 2: Add onset detection function to the display



• $_{Mz}$SpectralReflux plugin for Sonic Visualiser:
http://sv.mazurka.org.uk/download (Linux & Windows)

Data Extraction (4)

• Step 3: Estimate onset times from function peaks

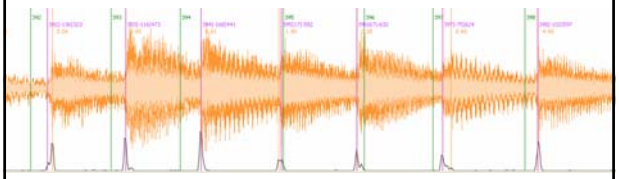


• Send taps (green) and onsets (orange) to external program:
http://mazurka.org.uk/cgi-bin/snaptap

(no interlayer processing plugins for Sonic Visualiser yet…)

Data Extraction (5)

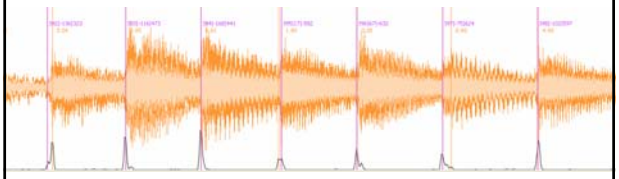
• Step 4: Load snaptap results into SV (purple):



• $_{Mz}$SpectralReflux currently sentitive to noise (old recordings)
so snaptap only works on clean recordings.

Data Extraction (6)

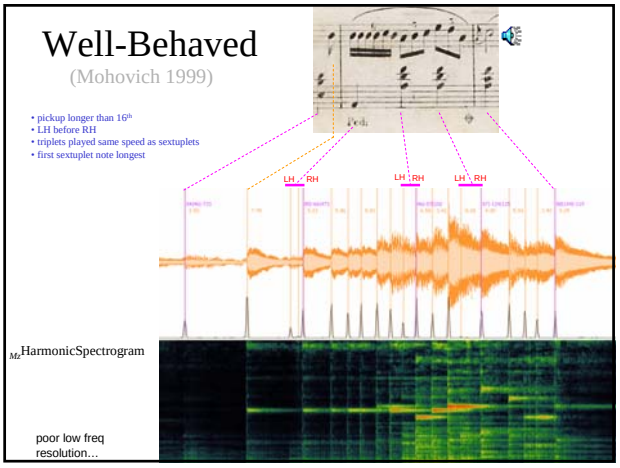
• Step 5: Correct errors



• Hardest part of data entry.
• Down to ~30 min/mazurka for clean recordings.
• 278 performances (of ~6 mazurkas) at this stage.

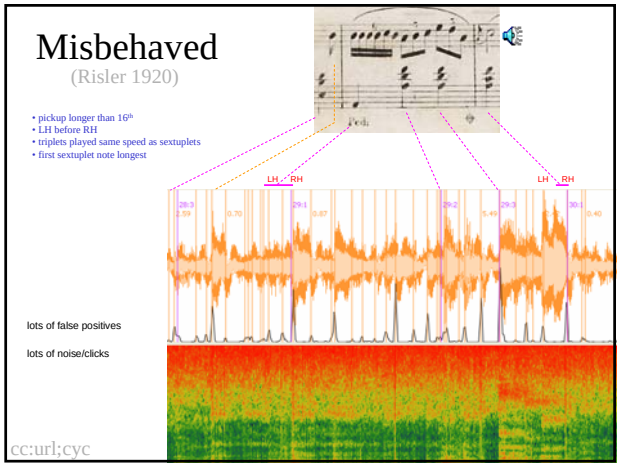
Well-Behaved

(Mohovich 1999)



Misbehaved

(Risler 1920)



cc:url;cyc

Extracted Feature Data

• Further feature extraction by Andrew Earis (note timings/dynamics)

• Beat times durations or tempos

0.150	0.604	99
0.754	0.530	113
1.284	0.500	120
1.784	0.512	117
2.297	0.567	106
2.864	0.567	106
3.432	0.641	94
4.073	0.998	60
5.072	1.870	32
6.942	1.555	39
8.498	1.821	33
10.320	1.177	51
11.497	0.811	74
12.309	0.835	72
13.145	0.658	91
13.804	0.610	98
14.415	0.619	97
15.034	0.622	96
15.657	0.637	94
16.294		
(seconds)	*(seconds)*	*(BPM)*

http://mazurka.org.uk/info/excel/beat

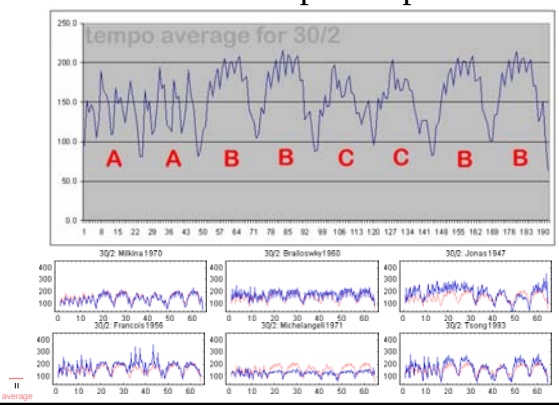
• Audio amplitude at beat locations

61.6
65.0
68.9
69.9
66.3
66.9
66.1
63.4
61.4
62.5
62.5
60.7
70.4
65.3
71.2
65.5
66.9
77.9
72.7
70.7
(~dB)

http://mazurka.org.uk/info/excel/dyn/gbdyn

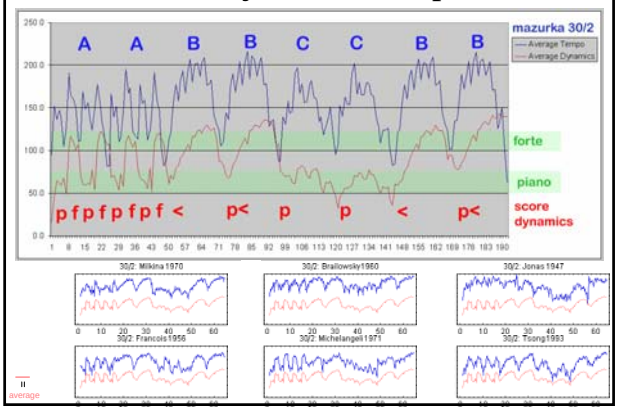
http://mazurka.org.uk/ana/tempograph

Beat-Tempo Graphs



http://mazurka.org.uk/ana/dyngraph

Beat-Dynamics Graphs



Pearson Correlation

$$r = \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_i (x_i - \bar{x})^2 \sum_i (y_i - \bar{y})^2}}$$

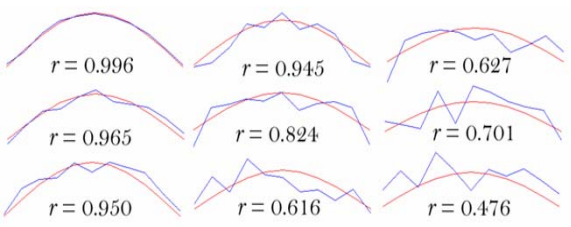
Example:

x = (1,3,4,2,7,4,2,3,1) y = (4,4,3,5,5,6,4,3,2)

$\bar{x} = 3$ (average of x) $\bar{y} = 4$ (average of y)

$r = 0.436436$

Shape Matching



Correlation & Fourier Analysis

correlation: multiply & sum

$$X(k) = \sum_n x(n)\, e^{-2\pi j n k}$$

spectrum = $X(k)$, signal = $x(n)$, sinusoids = $e^{-2\pi j n k}$

$X(k)$ = spectrum, indexed by k (frequency)

$x(n)$ = signal, indexed by n (time)

$e^{-2\pi j n k}$ = set of k complex sinusoids indexed by n

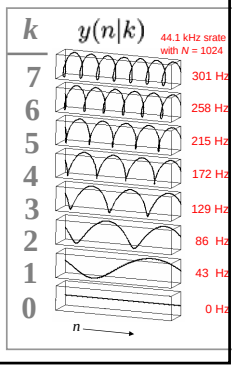
Correlation & Fourier Analysis (2)

Let $y(n|k) = e^{-2\pi jnk}$

Then the DFT can then be written as:

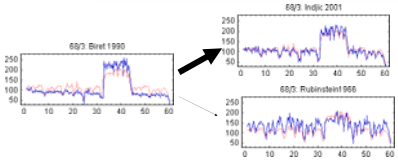
$$X(k) = \sum_n x(n)\ y(n|k)$$

$$X \begin{cases} \cdots \\ X_3 = \sum x(n)\ y_3(n) \\ X_2 = \sum x(n)\ y_2(n) \\ X_1 = \sum x(n)\ y_1(n) \\ X_0 = \sum x(n)\ y_0(n) \end{cases}$$



Performance Tempo Correlations

	Bi	Br	Ch	Fl	In	Lu	R8	R6	Sm	Un
Biret	1.	0.92	0.81	0.83	0.95	0.85	0.62	0.5	0.55	0.86
Brailowsky	0.92	1.	0.81	0.86	0.91	0.84	0.66	0.55	0.65	0.85
Chiu	0.81	0.81	1.	0.86	0.86	0.81	0.76	0.74	0.67	0.89
Friere	0.83	0.86	0.86	1.	0.88	0.84	0.73	0.7	0.74	0.89
Indjic	0.95	0.91	0.86	0.88	1.	0.88	0.66	0.59	0.63	0.9
Luisada	0.85	0.84	0.81	0.84	0.88	1.	0.67	0.61	0.56	0.89
Rubinstein 1938	0.62	0.66	0.76	0.73	0.66	0.67	1.	0.77	0.62	0.75
Rubinstein 1966	0.5	0.55	0.74	0.7	0.59	0.61	0.77	1.	0.59	0.69
Smith	0.55	0.65	0.67	0.74	0.63	0.56	0.62	0.59	1.	0.64
Uninsky	0.86	0.85	0.89	0.89	0.9	0.89	0.75	0.69	0.64	1.

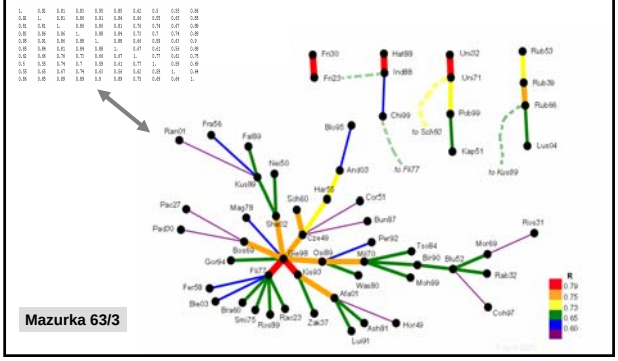


Highest correlation to Biret

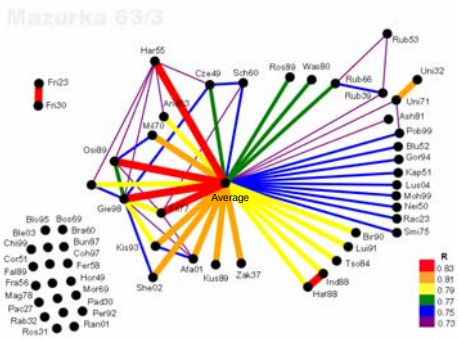
Lowest correlation to Biret

Correlation Maps – Nearest Neighbor

- Draw one line connecting each performance to its closest correlation match
- Correlating to the entire performance length.

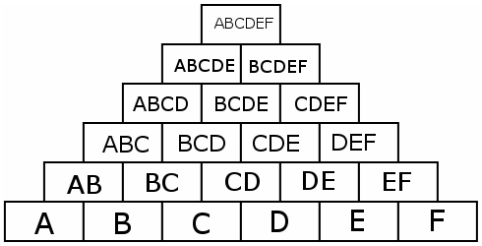


Absolute Correlation Map



Scape Plotting Domain

- 1-D data sequences chopped up to form a 2-D plot
- Example of a composition with 6 beats at tempos A, B, C, D, E, and F:



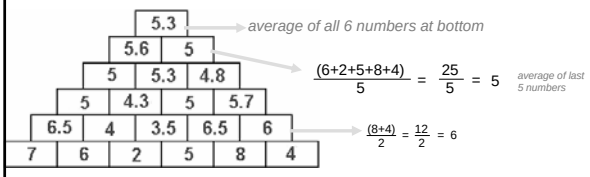
← analysis window size →

original sequence:

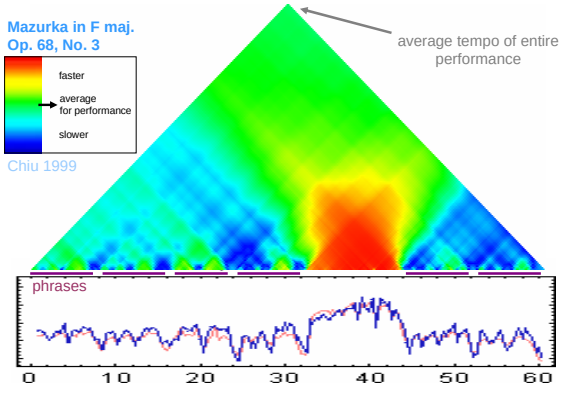
start ← time axis → **end**

Scape Plotting Example

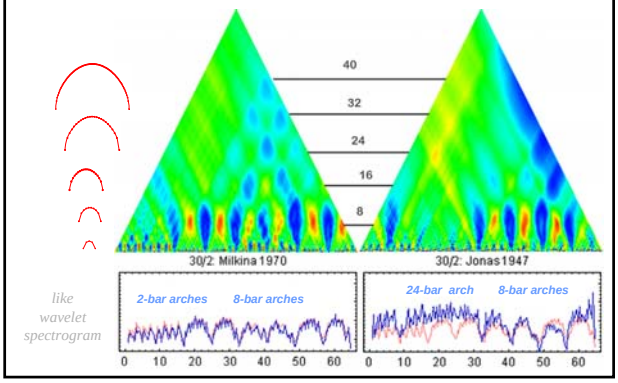
- Averaging in each cell with base sequence (7,8,2,5,8,4):



Average-Tempo Timescapes



Arch Correlation Scapes

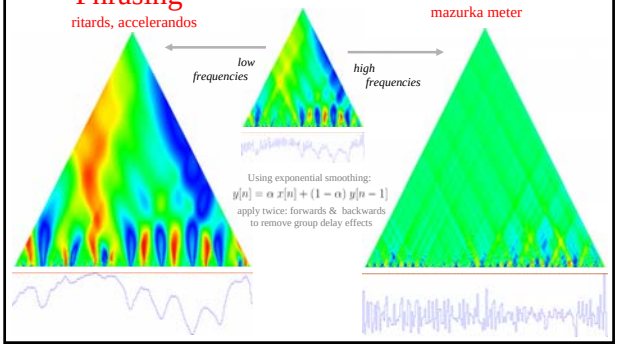


Composite Features

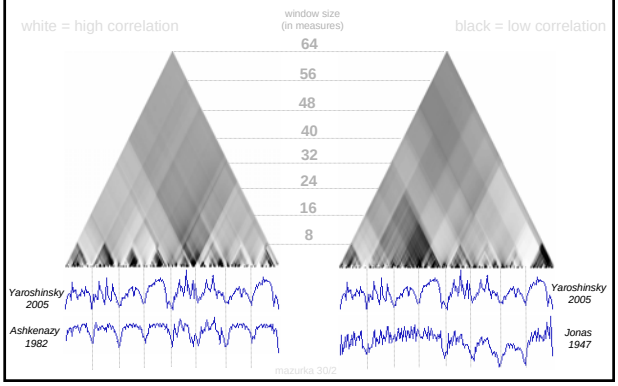
multiple performance features embedded in data

Phrasing
ritards, accelerandos

Accentuation
mazurka meter

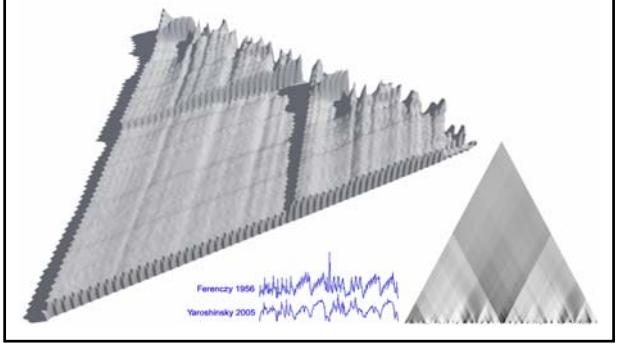


Binary Correlation Scapes

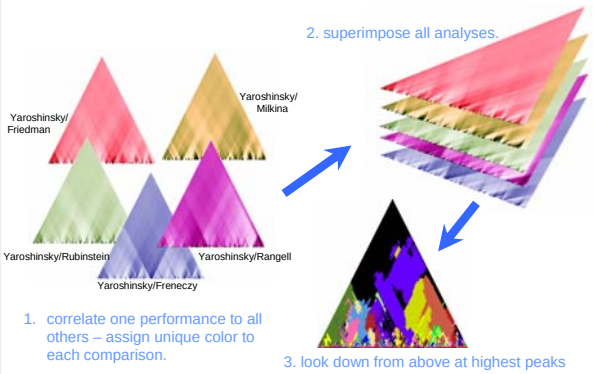


Scapes Are 3D Plots

2D plotting domain + 1D plotting range

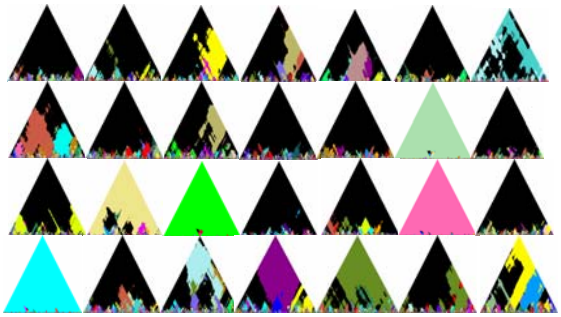


Scapes of Multiple Performances



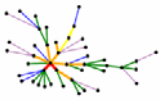
Perormance Correlation Scapes

• **Who is most similar to a particular performer at any given region in the music?**

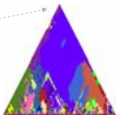


mazurka.org.uk/ana/pcor mazurka.org.uk/ana/pcor-gbdyn

Maps and Scapes



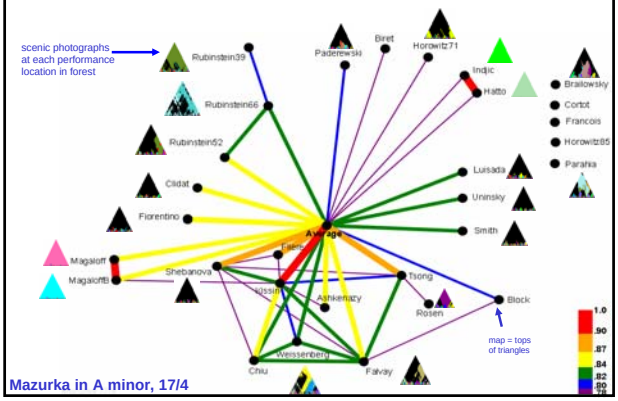
map points are tops of scapes



Correlation maps give gross detail, like a real map:

Correlation scapes give local details, like a photograph:

Map and Photos of the Forest



Mazurka in A minor, 17/4

Boring Timescape Pictures

Occasionally we get over-exposed photographs back from the store, and we usually have to throw them in the waste bin.

The same performance by Magaloff on two different CD re-releases:

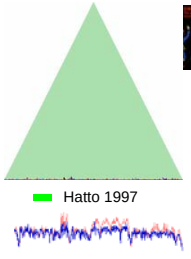
Philips 456 898-2 mazurka 17/4 in A minor Philips 426 817/29-2

• Structures at bottoms due to errors in beat extraction, measuring limits in beat extraction, and correlation graininess.

Boring Timescape Pictures?

Two difference performances from two different performers on two different record labels from two different countries.

Calliope 3321 mazurka 17/4 in A minor Concert Artist 20012



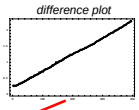
Indjic 1988 Hatto 1997

see: http://www.charm.rhul.ac.uk/content/contact/hatto_article.html

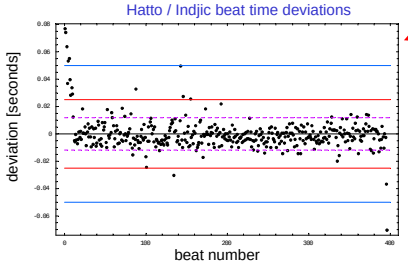
Beat-Event Timing Differences

Hatto beat location times: 0.853, 1.475, 2.049, 2.647, 3.278, etc.

Indjic beat location times: 0.588, 1.208, 1.788, 2.408, 3.018, etc.

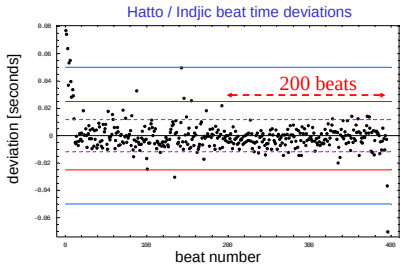


Hatto / Indjic beat time deviations



remove 0.7% timeshift

Timing Difference Probability



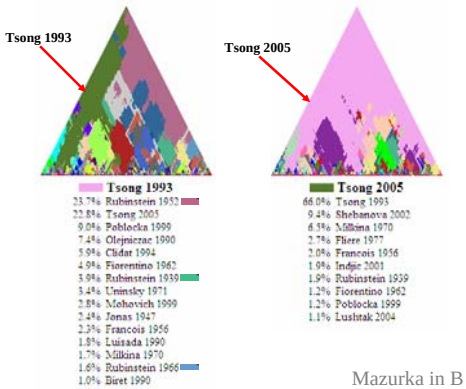
$1 : 2^{200}$

$1 : 10^{59}$

1 in $\sqrt{\text{googol}}$

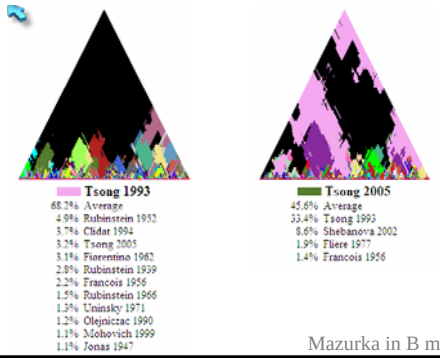
probability that same performer can produce a second performance so closely is equivalent to one atom out of an entire star.

Same Performer Over Time



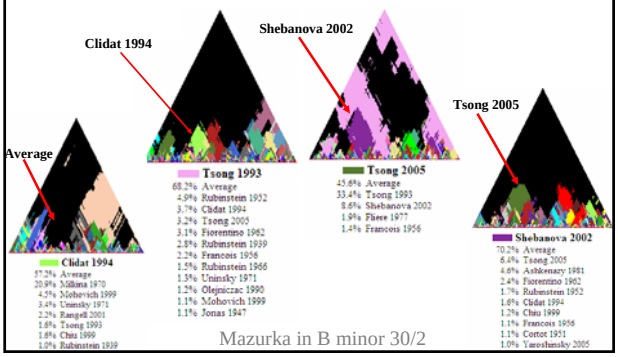
Mazurka in B minor 30/2

Including the Average



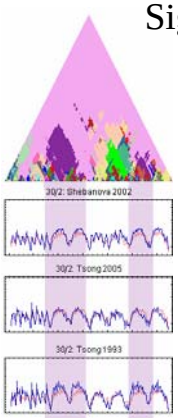
Mazurka in B minor 30/2

Mutual Best Matches



Mazurka in B minor 30/2

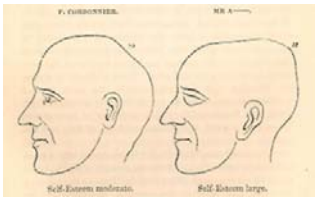
Significance of Similarity



- Ts'ong 2005 performance matches best to Shebanova 2002 in 3 phrases when comparing 36 performances of mazurka 30/2.
- Is this a coincidence or not?
- Could ask the pianist (but might be problem in suggesting an answer beforehand). Also they might not remember or be totally conscious of the borrowing (such as accents in language). Or there could be a third performer between them.
- Ideally a model would be used to calculate a probability of significance.

Mazurka in B minor 30/2

Phrenology

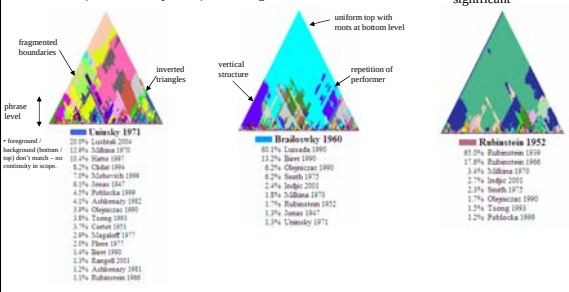


Significant or Not?

Example of non-significant matches (or at least very weak)

Example of possibly significant matches

Match to same performer always significant

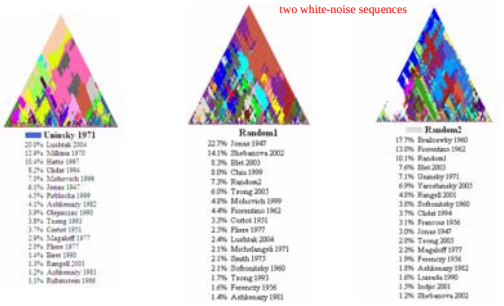


(Mazurka 30/2)

Purely Random Matching

• Plot has to show some match at all points…

two white-noise sequences

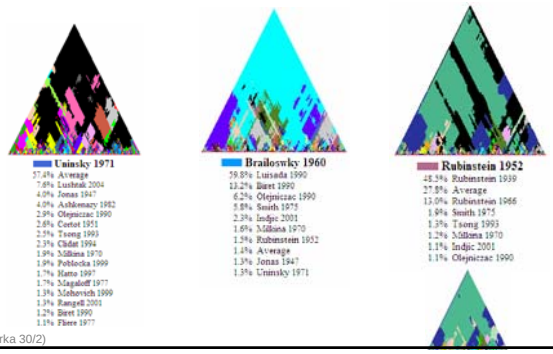


• Many performances are equidistant to Uninsky performance

(mazurka 30/2)

Including Average Performance

helps but does not solve significance question

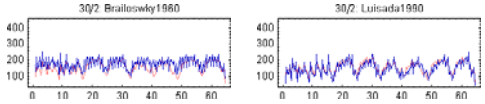


(mazurka 30/2)

What Is Significant?

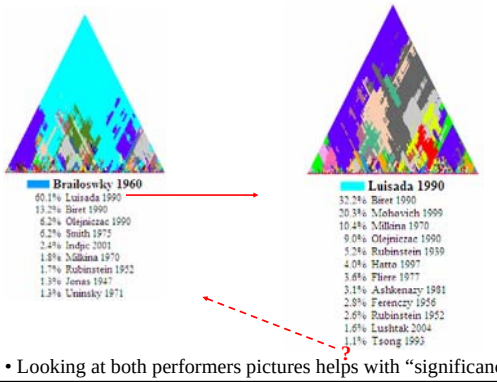
• No direct link found on web between Brailowsky and Luisada (such as teacher/student).

• Strong match in Brailowsky to Luisada probably due to large amount of mazurka metric pattern:



• Phrasing shapes very different, so match is both significant (high frequencies match) and not significant (low frequencies don't match).

Best Matching Not Mutual

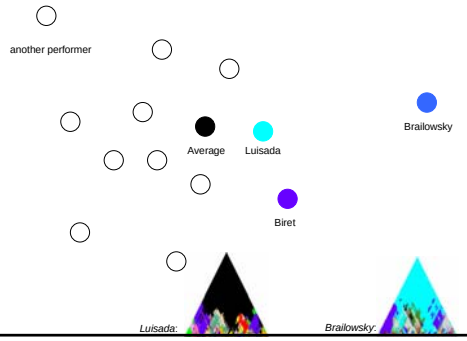


• Looking at both performers pictures helps with "significance"

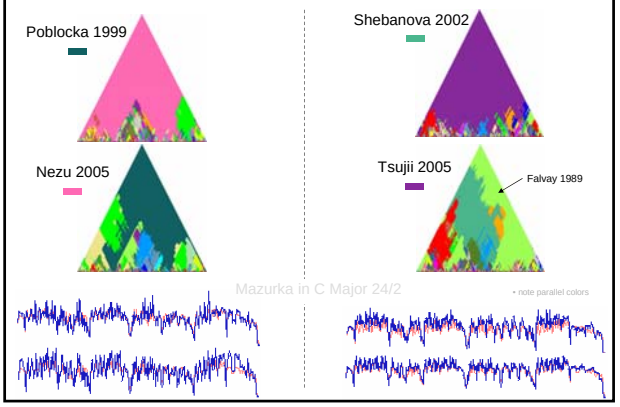
Performance Map Schematic

• Brailowsky has the strongest mazurka meter pattern

• Luisada has the second strongest mazurka meter pattern



Strong Interpretive Influences



Performance Ranking

• Given a reference performance,

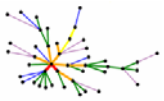
-- which other performance matches best
-- which other performance matches second best
…
-- which other performance matches worst

0^{th}-Order Similarity Rank

• Use large-scale correlation to order performances similarity to a target

performance	rank	correlation	
Rubinstein 1952	0	1.00	← target
Average	1	0.905	← best match
Rubinstein 1966	2	0.899	
Milkina 1970	3	0.876	
Poblocka 1999	4	0.836	
Tsong 1993	5	0.814	
Biret 1990	6	0.807	
Mohovich 1999	7	0.805	
Hatto 1997	8	0.799	
Indjic 2001	9	0.798	
Rubinstein 1939	10	0.784	
Shebanova 2002	11	0.783	
Luisada 1990	12	0.767	
Magaloff 1977	13	0.751	
Olejniczac 1990	14	0.740	
Biet 2003	15	0.721	
Clidat 1994	16	0.716	
Rangell 2001	17	0.716	
Lushtak 2004	18	0.715	
Chiu 1999	19	0.709	
Tsong 2005	20	0.706	
Smith 1975	21	0.695	
Fliere 1977	22	0.678	
Bralowsky 1960	23	0.644	
Ashkenazy 1982	24	0.642	
Ashkenazy 1981	25	0.637	
Cortot 1951	26	0.633	
Perenczy 1956	27	0.628	
Fiorentino 1962	28	0.615	
Uninsky 1971	29	0.597	
Francois 1956	30	0.577	
Vareshinky 2005	31	0.565	
Sofronitsky 1968	32	0.552	
Michelangeli 1971	33	0.550	
Jonas 1947	34	0.415	← worst match
Random1	35	0.102	} synthetic noise
Random2	36	0.054	
Random3	37	-0.015	

• Performance maps used rank #1 data:



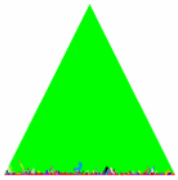
mazurka 30/2

1^{st}-Order Scape Rank

Smith 1975
56.9% Shebanova 2002
11.0% Fliere 1977
10.0% Average
4.7% Magaloff 1978
2.5% Bralowsky 1960
2.3% Tsong 1993
2.1% Ashkenazy 1981
1.9% Jonas 1947
1.6% Chiu 1999
1.4% Biret 1990
(mazurka 68/3)

• Area represented by best choice in scape plot.

• Hatto effect causes problems:



Indjic 2001
98.4% Hatto 1997

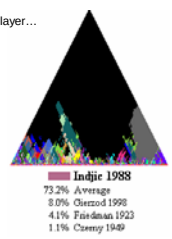
• Who is #2 for Indjic?

Scape Layers

Indjic 1988
99.7% Hatto 1988

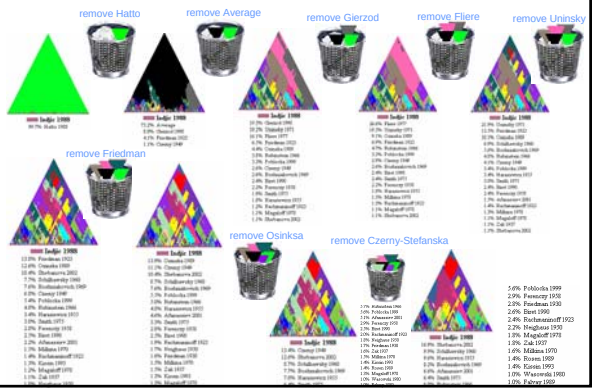
• Hatto takes up nearly 100% of Indjic's scape

• So remove Hatto layer…



Indjic 1988
73.2% Average
8.0% Gierzod 1998
4.1% Friedman 1923
1.1% Czerny 1949

Peeling the Layers



2nd-Order Scape Rank

• Rank score = area represented by best choice in scape plot, but peel away previous best choices.

performance	0-rank		1-rank		2-rank	
Indjic 2001	0	1.000	0	1.000	0	1.000
Hatto 1997	1	0.989	1	0.989	1	0.963
Average	2	0.861	2	0.00174	2	0.700
Poblocka 1999	4	0.796	3	0.00168	3	0.246
Rubinstein 1952	3	0.798	18	0.000225	4	0.156
Milkina 1970	5	0.791	30	0	5	0.188
Rubinstein 1939	12	0.714	13	0.000281	6	0.186
Mohovich 1999	6	0.704	17	0.000225	7	0.130
Leisada 1990	7	0.769	27	0	8	0.140
Biret 1990	8	0.763	4	0.00118	9	0.211
Rubinstein 1966	9	0.758	21	0.0000563	10	0.145
Olejniczac 1990	11	0.720	9	0.000394	11	0.137
Ashkenazy 1982	26	0.633	23	0	12	0.114
Ashkenazy 1981	27	0.628	5	0.000901	13	0.188
Francois 1956	10	0.721	7	0.000619	14	0.119
Uninsky 1971	17	0.697	20	0.000225	15	0.149
Fliere 1977	24	0.646	26	0	16	0.118
Smith 1975	28	0.627	22	0.0000562	17	0.092
Clidat 1994	18	0.677	6	0.000788	18	0.087
Tsong 1993	14	0.706	14	0.000281	19	0.076
Shebanova 2002	13	0.707	32	0	20	0.058
Blet 2003	15	0.706	24	0	21	0.064
Lushtak 2004	17	0.676	16	0.000225	22	0.066
Chiu 1999	20	0.672	12	0.000281	23	0.060
Magaloff 1977	16	0.698	28	0	24	0.077
Tsong 2005	21	0.659	33	0	25	0.070
Brailovsky 1960	23	0.649	10	0.000338	26	0.081
Ferenczy 1956	22	0.654	8	0.000586	27	0.084
Rangell 2001	25	0.635	31	0	28	0.100
Sofronitsky 1960	33	0.579	19	0.000225	29	0.054
Cortot 1951	29	0.605	25	0	30	0.067
Fiorentino 1962	31	0.567	11	0.000338	31	0.062
Yaroshinsky 2005	30	0.575	34	0	32	0.079
Michelangeli 1971	32	0.506	29	0	33	0.002
Jonas 1947	34	0.387	15	0.000225	34	0.101
Random2	36	0.034	36	0	35	0.044
Random1	35	0.047	35	0	36	0.044
Random3	37	-0.008	37	0	37	0.026

• Still slightly sensitive to the Hatto effect

(mazurka 30/2)

3rd-Order Scape Rank

• Start with the 2nd-order rankings, selecting a cutoff point in the rankings to define a background noise level in the scape.

• Then one-by-one add each of the non-noise performances into the scape along with the background noise performances. Measure the area covered by the non-noise performance (only one non-noise performance at a time).

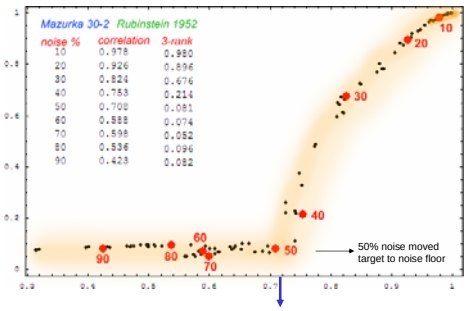
3rd-Order Rankings

Rubinstein 1952	0	1.000	0	1.000	← target
Average	1	0.984	1	0.933	
Rubinstein 1939	10	0.784	2	0.862	← same performer
Rubinstein 1966	2	0.899	3	0.849	← same performer
Milkina 1970	3	0.875	4	0.831	
Poblocka 1999	4	0.835	5	0.700	
Mohovich 1999	7	0.805	6	0.664	
Tsong 1993	5	0.814	7	0.636	
Biret 1990	6	0.806	8	0.622	
Indjic 2001	9	0.797	9	0.621	
Hatto 1997	0	0.790	10	0.616	
Olejniczac 1990	14	0.739	11	0.541	
Shebanova 2002	11	0.702	12	0.509	
Leisada 1990	12	0.766	13	0.426	
Magaloff 1977	13	0.750	14	0.338	
Lushtak 2004	18	0.715	15	0.277	
Tsong 2005	20	0.705	16	0.217	
Blet 2003	15	0.720	17	0.195	
Ferenczy 1956	27	0.627	18	0.169	← 50% noise floor
Ashkenazy 1981	25	0.637	19	0.133	
Fliere 1977	22	0.677	20	0.127	
Fiorentino 1962	28	0.615	21	0.116	
Clidat 1994	16	0.716	22	0.104	
Smith 1975	21	0.694	23	0.101	
Rangell 2001	17	0.715	24	0.100	
Ashkenazy 1982	24	0.641	25	0.094	
Chiu 1999	19	0.700	26	0.094	
Yaroshinsky 2005	31	0.564	27	0.094	
Cortot 1951	26	0.633	28	0.088	
Brailovsky 1960	23	0.643	29	0.073	
Uninsky 1971	29	0.596	30	0.072	
Michelangeli 1971	33	0.550	31	0.071	
Sofronitsky 1960	32	0.552	32	0.070	
Francois 1956	30	0.577	33	0.069	
Jonas 1947	34	0.415	34	0.066	
Random1	35	0.102	35	0.046	
Random2	36	0.054	36	0.046	
Random3	37	-0.015	37	0.026	

50% noise floor (2-rankings below "noise floor")

Proportional Noise

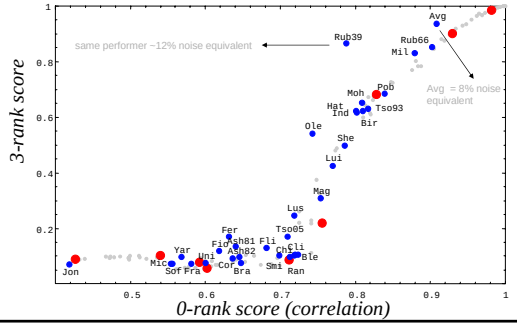
• Gradually add noise to target



0.72 (= 3-rank noise floor)

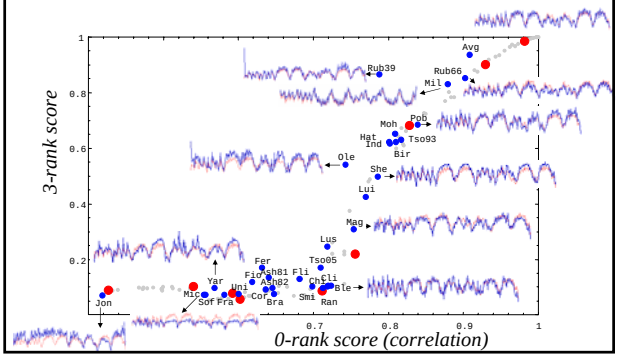
Real Data (1)

Mazurka 30/2
Target: Rubinstein 1952

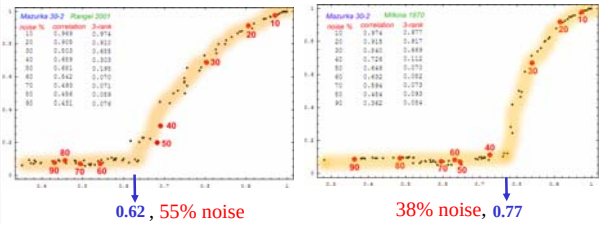


Real Data (2)

Mazurka 30/2
Target: Rubinstein 1952



Distance to the Noise Floor



0.62 , 55% noise

38% noise, **0.77**

• Metric for measuring individuality of interpretation?

• 3-Rank scores more absolute than correlation values
  • noise floor is always about a 3-rank score of 10%

• 3-Rank scores less sensitive to local extrema

Cortot Performance Ranking

• Master class recording contains 48 out of 64 measures (75%)

Con. Artists Rankings

0-Rank:	**3-Rank**
1. Average	1. Average
2. Rangell 01	2. Rangell 01
3. Milkina 70	3. Mohovich 99
4. Mohovich 99	4. Rubinstein 39
5. Shebanova 02	5. Milkina 70
…	…
32. Masterclass	31. Masterclass

Match to other Cortot near bottom of rankings.

Masterclass Rankings

0-Rank:	**3-Rank**
1. Poblocka 99	1. Average
2. Average	2. Rubinstein 52
3. Rubinstein 52	3. Luisada 90
4. Tsong 93	4. Poblocka 99
5. Tsong 05	5. Hatto 94
…	…
33. Con. Artist	35. Con. Artist

Match to other Cortot near bottom of rankings.

(comparing 35 performances + average)