Proceedings of 20th International Congress on Acoustics, ICA 2010

23–27 August 2010, Sydney, Australia

Nonlinear Assessment of Korean Temple Bell using Reccurence Quantification Analysis

Kuekjae Lee, Hoon Heo, Mingu Lee Sang Ha Park and Koeng-Mo Sung

Applied Acoustic Lab., Institute of New Media and Communications, Seoul National University, Seoul, Republic of Korea.

PACS: 43.60.Wy, 43.66.Ki

ABSTRACT

Recurrence plot provides a graphical representation of the recurrence patterns of time-series and the quantifiction of which gives comparable feature values. Here we highlights the application of them to derive structural information of Korean temple bells. With the help of decision making models, we set up the experiments to find RQA features whose values are consistent to the preference rankings of subjective listening tests.

INTRODUCTION

Korean temple bell is a religious percussion instrument which are manufactured more than 1200 years. In spite of long history of korean temple bell, manufacturing skills are based on experiences of craftmans and even their skills are not handed down to the decendents. Therefore, the sounds of korean temple bell have wide spectra which causes preference tendecies of listeners, Moreover, characteristics such as beat phenomenon and inharmonic frequency ratio of fundamental to hum in are known as the virtue of the sound of Korean temple bell which are avoided in manufacturing the western bells.

The purpose of the study on finding preference tendencies of the sounds of korean temple bells and relates them to physically meaningful features. Since beat phenomenon and inharmonicities are not explained by linear time-series analysis. we adopted recurrence plot (RP) and recurrence quantification analysis (RQA) which are nonlinear time-series analysis. While conventional analysis based on frequency domain analysis starts from decomposing time-series into each frequency components, RP inherently investigates the relationship of components as a whole and visualize the hidden structures of time-series in 2D space. This effort can be taken a step further by the quantification of recurrence plot elements. Quantified RQA features have benefits in comparing structuredness in the sounds of Korean temple bells.

In this paper, we review the definition and properties of PQ and PQA in section 1, In section 2, we set up the experiments using decision making model which enables features as criteria of decision making procudure. then modeled rankings are compared to the preference listening tests. We close our study after discussing the meaning of RQA features which are consistent with the preference rankings of listening tests in section 3 .

RECURRENCE QUANTIFICATION ANALYSIS

Embedding Theorem

The first step to analyze nonlinear dynamic system from one-dimensional scalar time-series signals is on the reconstruction of multidimensional data using delay coordinate embedding method [3]. That is, one-dimensional scalar time series signal $\{x(i), i = 1, 2, \dots\}$ are rearranged into multidimensional vector $X_i = (x(i), x(i+L), \dots, x(i+(m-1)L), \dots)$ using time delay L and embedding dimension m. This procedure is based on Takens' embedding theorem [4] which shows that topological features consist of multivariable are restorable from one dimensional measurements of multidimensional system, This is essential for system restruction and for analysis when information of multidimensional nonlinear dynamic system is restricted to one dimensional meansurements.

Time delay is usually chosen where autocorrelation or mutual information of one-dimensional time-series signal are minimized. However, there is no certain theoretical method for determining embedding dimension. Moreover, there's some reports that it is not a significant parameter in lower dimensional systems [**Iwanski1998**, 2]. Hence it is properly determined by experiment [1, 5].

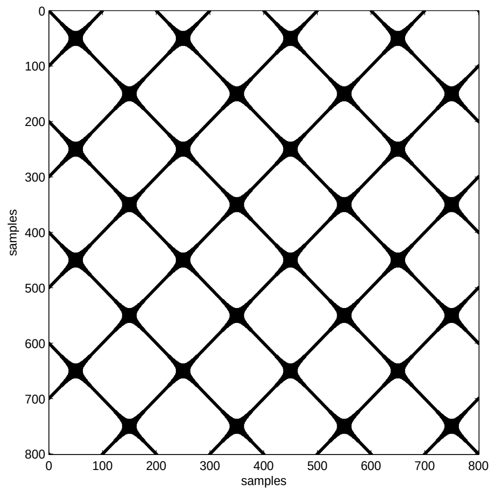
Recurrence plot (RP)

Recurrence plot is a sort of nonlinear scalar analysis, which visualizes repetitive pattern of multidimensional restored data X_i from one-dimensional time-series signal by delay coordinate embedding. Closeness between X_i and X_j is determined from distance measure and threshold ε.

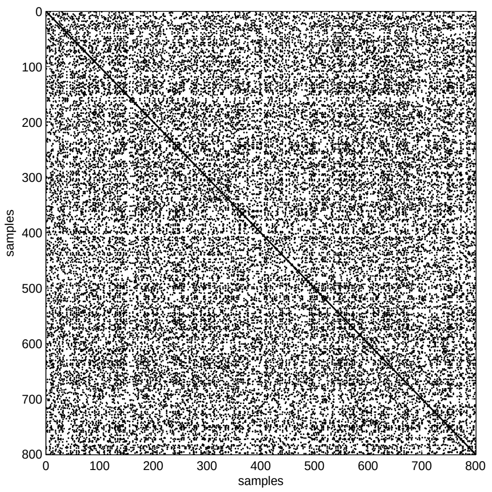
$$R_{i,j}^{m,\varepsilon} = \theta(\varepsilon - ||X_i - X_j||) \qquad (1)$$

In equation 1, Heaviside function θ returns 1 if distance between each X_i is less than ε and 0 if otherwise. When $R_{i,j}^{m,\varepsilon} = 1$, X_i is called recurrence point, and RP shows recurrence matrix consist of its elements 0 and 1. Examples of RP of sinusoidal signal and white noise is below. For 800 samples signals, embedding dimension and time delay are chosen as 1 and Euclidean distance with threshold 0.1 is used. Figure 1(a) shows swell-organized recurrence patterns of sinusoidal input signal while that of white noise signal does not show particular pattern as figure 1(b)

There are some advantages in RP since any assumption for time-series signal is not necessary [6] which are necessary for frequency domain analysis. Besides, while the frequency domain analysis starts from decomposing signals into components using transformation such as FFT and wavelet, recurrence plot analyze the overall structure that all components organized.



(a) sinusoidal



(b) white noise

Figure 1: examples of recurrence plot

Recurrence quantification analysis (RQA)

Recurrence quantification analysis (RQA) is first suggested by Zbilut & Webber in 1994 [6]. RQA acts as a complement of RP since visualized structuredness of time-series of RP needs to be quantified. From now on, several features explains about structuredness and statistics of RP are suggested and utilized. Explanation of features are omitted here to interprete them after exprements in section 2.

EXPERIMENTS

Preference listening test

There are two kinds of ranking methods based on pairwise comparison and multistimulus comparison. Evaluating more than 10 stimulus simultaneously is not recommened in ITU-R BS.1534 standard for subjective listening tests Moreover participants in listening test are not experts in Korean temple bells. Therefore we acquire the preferential information in form of parisowse comparison s of all pairs of objects, which needs lesser cognitive efforts than that of multistimulus comparison. 380 sets of paiswise comparisons contains all combination of 20 stimulus are presented to 14 participants by 5 scale pairwise comparison programs in STEP(Subjective Training and Evaluation Program) developed by Audio Research Lab.

Each results of listening test of participants is stored as a pairwise comparison table (PCT). Even though PCT is comparable to other ranks, we consider that paisewise comparison results of non-experts are noise-contaminated from false-determination. So we exploit the outranking graph using the Net Flow Scoring procedure (NFS) to find rankings of each participant [**Budzynska**]. In outranking graph, an stimulus $a \in K$ is represented as a node. Score from net flow is a difference between ingoing flow $\phi^+(a)$ and outgoing flow $\phi^-(a)$ which are represent relative strength and relative weakness of node respectively.

$$\phi^+(a) = \sum_{b \in K} \pi(a,b), \quad \phi^+(a) = \sum_{b \in K} \pi(a,b) \tag{2}$$

$$\phi(a) = \phi^+(a) - \phi^-(a) \tag{3}$$

where $\pi(a,b)$ is a element of PCT. Then, relations between stimulus such as preference and indifference are determined by comparing net flow scores.

$$\text{a P b iff } \phi(a) > \phi(b) \tag{4}$$

$$\text{a I b iff } \phi(a) = \phi(b) \tag{5}$$

where P and I means prefered and indifferenced respectively.

Feature extraction

Stimulus are recorded at 44.1 kHz sampling rate and average 20 dB decay time is over 20 seconds. And RQA feature is known to sensitive to the change of systems. Therefore we needs process recorded stimulus beforehand to concentrate on localized time-series and reduce the analyzed data size. First, lowering sampling rate is a method. In our experiments sampling rate is lowered to 16kHz in the limit of minimizing the loss of spectral components of stimulus and the change the characterisitc of RQA features.

Next, we focus on two different parts of stimulus; (1) beginning part and (2) time envelope of stimulus. Frequently beginning part of stimulus determinines the image of bell sound. So RQA features are extracted only on the duration of 0.5 second starting from oneset time. Time envelope is selected since it contains information of beat and decay time of korean temple bell. time envelope of stimulus are produced by simple envelope detection algorithm consists of half-wave rectifier, lowpass filter and downsampling.

Following RQA features are extracted from beginning part and time envelope of stimulus. these features are computed in matlab by cross recurrence toolbox developed by Marwan et al . All of features are not meaningful here, some of them are explained in section 3.

- REC : recurrence rate
- DET : determinism
- L_{avg} : average diagonal length
- L_{max} : length of longest diagonal line
- ENT : entropy of diagonal length
- LAM : laminarity
- TT : trapping time
- V_{max} : length of longest vertical line
- T1 : recurrence time of 1st type
- T2 : recurrence time of 2nd type

Additionally we extract features from frequency domain analysis.

- f_h : hum frequency
- f_0 : fundamental frequency
- c_h : weighted centroid for the whole stimuli
- c_a : weighted centroid at the beginning part
- c_s :weighted centroid at the stationary part
- Δc : $c_a - c_s$

- b_0 : beat period of fundamental
- b_L : beat period of loudest partial
- T^0_{20dB} :20dB decay time of fundamental
- T^L_{20dB} : 20dB decay time of loudest partial
- f_b/f_0

Feature selection

In this subsection, we compare individual preference rankings of listening test to rankings constructed from each feature one by one. If preference tendencies of individual similar to the rankings constructed from a feature, they gives a vote to that feature. Then we select the features whose have many votes and high similarities. Since this procedure based on one by one comparison, it has different meaning that the procedure in next subsection. There are two points to be concern, one is how to construct rankings from each feature and the similarity measures.

Constructing ranking from a feature is considered as an unicriterion decision making problem [**Brans1985**].

$$\max\{f(a)|a \in K\} \tag{6}$$

where $f(\cdot)$ is a criterion used in decision making which are RQA features here. Since it is well-stated problem, it is possible to construct a complete graph for all a and $b \in K$. even without the help of net flow scoring method (NFS).

Kendall's rank correlation τ is used to check the similarities between pseudo rankings obtained from RQA feature and individual preferrence rankings of listening tests. It is based on computing the distance between pairwise comparison table obtained from two different ranking systems.

$$\tau = 1 - 4\frac{d_k(R^1,R^2)}{m(m-1)} \tag{7}$$

In equation 7, Kendall's distance $d_k(R^1,R^2)$ is defined as

$$d_k(R^1,R^2) = \frac{1}{2}\sum_{i,j=1}^{m} |r^1_{ij} - r^2_{ij}| \tag{8}$$

Computed rank correlation values from 1st place to 3rd place are listed in below tables. Even though several features dominates othe features, it needs to make a scoring method to definitely see the order of features. there might be many possible ways to score and we just simply score the features using the sum of correlation and number of votes. Only 1st,2nd,3rd place votes are considered in scoring.

$$s = v_1\tau_1 + v_2\tau_2 + v_3\tau_3 \tag{9}$$

where v_i is the obtained number of votes in ith place.

To eliminate the effects of insignificant correlation in votes, rank correlation of votes lesser than 0.05 significant level are disregarded in voting. Significant test of Kendall's rank correlation is computed on the distribution of all possible outcomes of ranking $N!$. In more than 10 stimulus cases simple approximation is possible

$$\sigma^2_\tau = \frac{2(2N+5)}{9N(N-1)} \tag{10}$$

$$Z_\tau = \frac{\tau}{\sigma_\tau} \tag{11}$$

where Kendall's rank correlation τ are normalized to Z_τ. then significant level is computed by error function of Z_τ.

Table 1: Votes for RQA features at beginning part

	1st place votes	2nd place votes	3rd place votes
REC	2 (-0,44)	2 (-0.45)	0
DET	3 (0.41)	3 (0.37)	2 (0.31)
L_{avg}	0	0	2 (0.32)
L_{max}	1(0.38)	2 (0.28)	2 (0.3)
ENT	0	0	1 (0.24)
LAM	0	1 (0.41)	1 (0.34)
TT	2 (0.40)	0	2 (-0.33)
V_{max}	0	1 (0.32)	1 (0.37)
$T1$	2 (0.46)	2 (0.42)	1 (0.32)
$T2$	4 (0.46)	3 (0.36)	2 (0.38)

Table 2: Votes for RQA features at time envelope

	1st place votes	2nd place votes	3rd place votes
REC	0	1 (0.23)	0
DET	2 (-0.61)	4 (-0.48)	0
L_{avg}	2 (-0.46)	3 (-0.48)	3 (-0.45)
L_{max}	1 (0.25)	0	0
ENT	5 (-0.43)	1 (-0.41)	0
LAM	0	0	0
TT	0	0	4 (-0.51)
V_{max}	1 (0.27)	0	0
$T1$	0	0	1 (-0.23)
$T2$	4 (0.46)	1 (-0.24)	0

Multicriteria Decision Making Model

Next, we compare the averaged preference rankings of listening test and rankings constructed from decision making model using selected features. Unlike feature selection, we compare averaged ones. All of the individual net flow of listening tests have zero means but variance is different. Therefore net flow of individuals are normalized to have equal variance and summed to form an averaged preferencek ranking.

it is more mimic the decision making of humans. Even though it is already uttered that verification is not strict because we used listening test result to select features, integrated form of decision making shows some dependencies between RQA features used as criteria. That is, independently chosen features are examined as a model of decision making model. In the assumption that preference ranking oraganization of humans are multicriteria deicision making problem, we construct multicriteria decision making model(MCDM). It aims to verify the correlation between the ranking obtained from preference listening test and the ranking obtained from selected RQA features

Multicriteria decision making problem is stated as [**Brans1985**]

$$\max\{f_1(a), f_2(a), f_3(a), \ldots |a \in K\} \tag{12}$$

Since several criteria is concerned in decision making procedure, the essence of MCDM is how to combine preference functions of different criteria into a prefernce index. We use sum of preference functions extended by Gaussian like function.

$$P_h(a,b) = \begin{cases} 1 - e^{-\frac{(f_h(a)-f_h(b))^2}{2\sigma_h^2}} & \text{if } f_h(a) \geq f_h(b) \\ 0 & \text{if } f_h(a) \leq f_h(b). \end{cases} \tag{13}$$

$$\pi(a,b) = \sum_h \omega_h P_h(a,b) \tag{14}$$

where standard deviation σ_h is only parameter to be determined and weighting factor ω_h is determined by prior knowledge. Since we are interested not in constructing a prediction model but in verification of validity of RQA features, ω_h is equal to selected features and variance of itself are used to normalize

Table 3: Votes for features from frequency domain analysis

	1st place votes	2nd place votes	3rd place votes
f_h	3 (0.41)	1 (0.43)	0
f_0	1 (0.49)	3 (0.34)	0
c_h	7 (-0.40))	0	1 (0.39)
c_a	0	7 (-0.38)	0
c_s	0 (-0.43)	0	2 (-0.3)
Δc	0	0	2 (-0.38)
b_0	0	0	0
b_L	1 (-0.25)	0	0
T^0_{20dB}	0	0	2 (-0.29)
T^L_{20dB}	0	1 (-0.24)	3 (0.17)
f_b/f_0	0	0	0

the strength of criteria. Once preference index are obtained, multicriteria decision making problem becomes unicriterion problem. That is, net flow score is obtained and prefered and indifferenced are determined by that score.

Table 4: Tables should be centred.

	beginning part	time envelope	frequency analysis
MCDM	0.49*	-0.32	0.32
*PC*1	0.31	-0.33	-0.14

Compared results are listed in table **??** using kendall's coefficient . To see the unitilies of MCDM, rank correlation against scores obtained from 1st component of Principle component analysis(PCA) of variance normalized RQA features are noted additionally. In the cases of RQA features of beginning part and features obtained from frequency analysis, MCDM based one is better than PC1 because MCDM reduces the effect of outlier by using gaussian like function in equation 13. In results, ranks obtained from MCDM using RQA features of beginning part shows highest correlation with preference listening test. In the case of time envelope, similaritiy is reduced compare to similaritiesobtained in feature selection. This means there are no general tendencies whether in features or in individuals. This similar to the case of features from frequency analysis. Figure 2 shows that rank correlations similar but different aspects of noises are different. That is, in the case of features from frequency analysis, some of stimulus are aligned into negative correlation while there's no correlation in time envelope case. In addition to the fact that there are few votes for the envelope features such as 20 dB decay time and beat frequency, this means that participants in listening test does not concerns envelope information of stimulus to makes his/her prefence rankings.

DISCUSSION

It is difficult to say whether a feature is relevant to preference tendencies or not from above results. But by investigating the meaning of features obtained high scores, further interpretation is possible. Here we only concerns the features determinism and recurrence time of type 2 and centroid and fundamental frequency and hum frequency.

Determinism is defined as

$$DET = \frac{\sum_{l=l_m in}^{M} l p^{\varepsilon}(l)}{\sum_{i,j}^{M} R_{i,j}^{m,\varepsilon}} \tag{15}$$

where $p^{\varepsilon}(l)$ is the frequency distribution of the lengths of diagonal line l. iThe diagonal line l which parallel to the main diagonal in RP. diagonal line in RPcorresponds to when X_i and X_j are close then X_{i+k} andX_{j+k} are also close together for a series of k's. Determinism relates not only to average value of diagonal length, but also to recurrence rate in denominator. Recurrence rate itself does not represent the structuredness of time-series since it only counts the number of recurrence points in RP. One example is that of white noise whose value is just sensitive to threshold. Therefore determinism has information about diagonal structureness in same number of recurrence points. In the beginning part of stimulus, a mixed form of noise-like attack and periodic signals are presented. So determinism is related to the ratio of periodic components to noiselike components. In voting results, other statistics related to the diagonal line which parallel to the main diagonal such as entropy and L_{max} are got nothing, and the recurrence rate shows negative correlations.

Recurrence times are in two kind of type [**Gao2000**]. one is the Poincare recurrence times. Let us arbitrarily choose a reference point X_1 on the reconstructed trajectory,$\{X_i, i = 1,2,\ldots,N\}$, and consider recurrences to its threshold ε:$B(X_1) = \{X : ||X - X_1|| \geq \varepsilon\}$. Denote the subset of the trajectory that belongs to $B(X_1)$ by $S_1 = \{X_{t_1}, X_{t_2}, \ldots, X_{t_i} \ldots\}$. These are the Poincare recurrence points. Then dots will be placed at points $(1,t_i), i = 1,2,\ldots$. From set S_1, we can define the Poincare recurrence times by $\{T_1(i) = t_{i+1} - t_i, i = 1,2,\ldots\}$ then we uses the mean value $\tilde{T}_1$. If the threshold ε is not too small, we can have a recurrence times $T_1 = \tau$ where τ is sampling time. Therefore it needs to find recurrence time of different type avoiding find adjacent recurrence points which called sojourn points. So recurrence times of type 2 are computed by negliging the sojourn points. For a periodic signal, $T_2(i)$ simply gives an estimation of the periodicity of the signal.

CONCLUSION

In this paper, nonlinear assessment of Korean temple bell is performed by RP and RQA. We make a set of experiments using decision making model to find features relevant to preference listening test. Then further interpretation of features based on their definition are conducted. Since participants are not experts in Korean temple bell, related features are not of inharmonicity or beat frequencies.

Future work advance in the direction of preference listening test of given criteria and further refinement of stimulus for feature extraction. And other kind of nonlinear analysis method is to be applied to find meaningful features.

REFERENCES

[1] A. Keller. "Computational Prediction of Song Genre with Recurrence Quantification Analysis Contents". *Relation* 10.1.33 (2008), p. 1085.

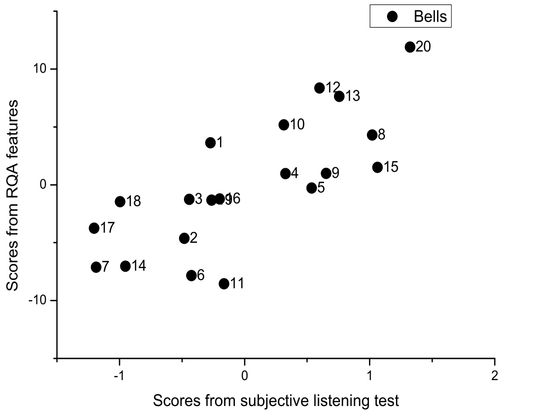
[2] TK March, SC Chapman, and RO Dendy. "Recurrence plot statistics and the effect of embedding". *Physica D: Nonlinear Phenomena* 200.1-2 (2005), pp. 171–184.

[3] T. Sauer, J.A. Yorke, and M. Casdagli. "Embedology". *Journal of Statistical Physics* 65.3 (1991), pp. 579–616.

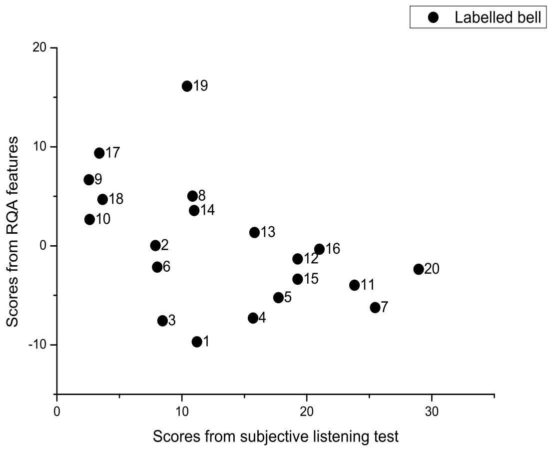
[4] F. Takens. "Detecting strange attractors in turbulence". *Dynamical systems and turbulence, Warwick 1980* (1980), pp. 366–381.

[5] L.L. Trulla et al. "Non linear assessment of musical consonance". *EJTP* 2.8 (2005), pp. 22–34.

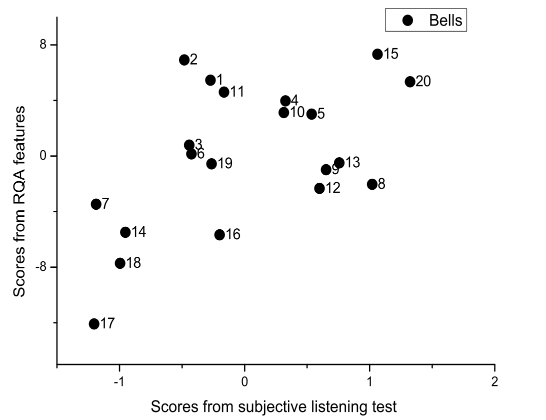
[6] C.L. Webber Jr and J.P. Zbilut. "Recurrence quantification analysis of nonlinear dynamical systems". *Tutorials in contemporary nonlinear methods for the behavioral sciences* (2005), pp. 26–94.



(a) RQA features of beginning part

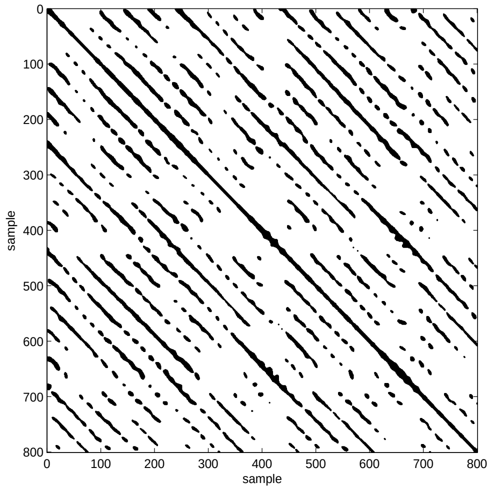


(b) RQA features of time envelope

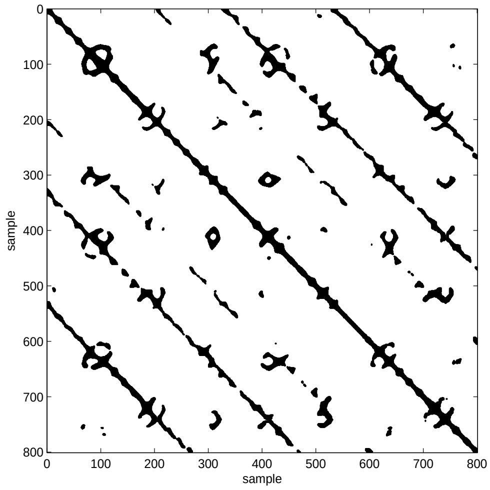


(c) Features from frequency analysis

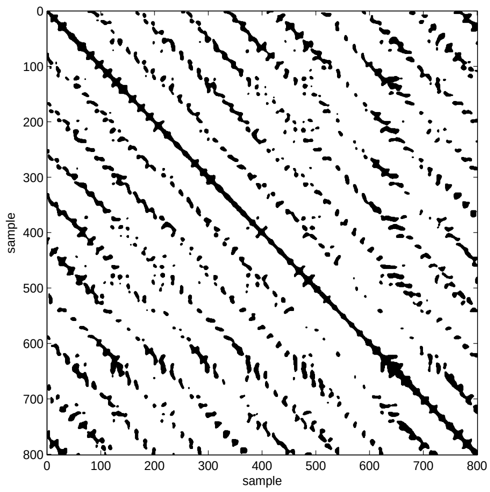
Figure 2: Scatterplots of preference listening tests results and modeled features



(a) bell 20



(b) bell 15



(c) bell 7