Proceedings 2003 IEEE International Symposium on
Computational Intelligence in Robotics and Automation
July 16-20, 2003, Kobe, Japan

Analysis of Network Ensemble with Time Lag

Yohei Kobayashi*
Dept. of Computational Intelligence and Systems Sience Tokyo Institute of Technology
www.myk.dis.titech.ac.jp

Yoichi Nagata
Dept. of Computational Intelligence and Systems Sience Tokyo Institute of Technology
www.myk.dis.titech.ac.jp

Yoshihiro Miyake
Dept. of Computational Intelligence and Systems Sience Tokyo Institute of Technology
www.myk.dis.titech.ac.jp

Abstract.

In recent years, the ensemble technology between the distant places with the networks have been proposed. The problem of the technology is the affection of the time lag in the networks. The purpose of this study is to clarify the characteristic of the ensembles between the distant humans. We constructed the network ensemble system, which is possible to set the time lag, and analyze the changes in the various kinds of the time lags. As the results, when with the 90ms or more time lag, the performance becomes failure. Also then with the 50-80ms time lag,they perform with the differentiation of the roles based on the mutual anticipations.

1. Background

In recent years, the progress of network technology such as Internets have caused the various kinds of possibility in the musical medias. One of the example is the method to enjoy music in distant humans; for example, digital transmitter of live performance, session systems, or orchestra systems by using the MIDI (Musical Instrument Digital Interface) [1]. One of the advantage of these technologies is that the distant performers or listeners don't have to use something special instruments and are able to feel like playing or listening the music in the same place more easily. But in the present network technologys, there are limitations of the communications speed because of the time lag between the distant humans.

There are lots of approaches and studies to reduce an influence of the time lag to the media transmittion [2]. For examples, in the researches of force sensing, one of the targets is how to recover and reconstruct the time structure in the medias under time lag and jitter [3]. Especially the force sensing is strongly related to the real-time interactions, so it is known that the time lag causes to the deterioration of the quality. Similarly, the musical performance is one of the medias that strongly related to the real-time process, and it is necessary to preserve and recover the media from the time lag.

In the previous studies and technologies, several kinds of methods were used to send a musical performance under the time lag. But, one of the most important things is these kinds of transmittion method is able to fall into two types. One is one-sided and the other is interactive. For example, to send a musical performance to listener, although it is necessary to send it in real-time under the time lag, the sever-client systems are able to approach the solution [4]. Similar to this, the server-client systems are used on the musical session systems by the separation of the playing parts [5,6]. However, to construct the face-to-face ensembles on the network, the interactions between performers are necessary. But in such kind of situation, the affection of the time lag occurs directly to the interaction and the ensemble is broken down under the strong affection [7]. Thus, the such support are needed to in such kinds of situations, but no method have been developed yet. Therefore in this research, for the first step to construct a real-time interaction on the networks, we analyze how the performers play the music with the time lag. And investigating the mechanism of their performances, we try to apply it to machine and overcome the time lag problems in the networks.

2. Former researches

There are several kinds of researches to make an ensemble on the network. In these researches, the ensemble between distant humans are constructed by connecting personal computers by the ISDN networks and sends the MIDI message each other through the network. In this system, it is useful in the case of short distances, however it is difficult to use in the long distance because of the real-time problem [7].

On the other hand, there is an approach to send musical information correctly on the network under the time lag. In these researches, the systems are constructed to use MIDI instruments form plural workstation [4,5,6]. In this systems, they added the timestamp to the data of performance and send to the server. It control the time from the datas of timestamp and send it to the MIDI instruments. In these methods, it is able to synchronize between the instruments, but these kinds of server-client systems only send the informations of the performance one-sidedly. Thus in these systems, there is no interaction between performers, and different from realization of face-to-face ensemble.

Therefore, there are two problems to realize network ensembles. One is the affection of the time lags and it will destroy the real-time interaction in the performance. And the other is the problem of server-client systems and it will destroy the interaction between performers.

3. Experiment

We constructed a new network ensemble system, which is able to set the time lag between the two distant performers. By using this system, we analyze the influence of time lag to the human performance (Fig. 1). We do ensemble by using this system, and measure how humans' performance changes in the variety of time lag. In this figure, the input

0-7803-7866-0/03/$17.00 ©2003 IEEE

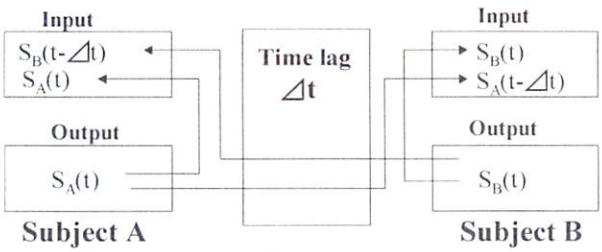

Fig.1 Conceptual figure of measurement system

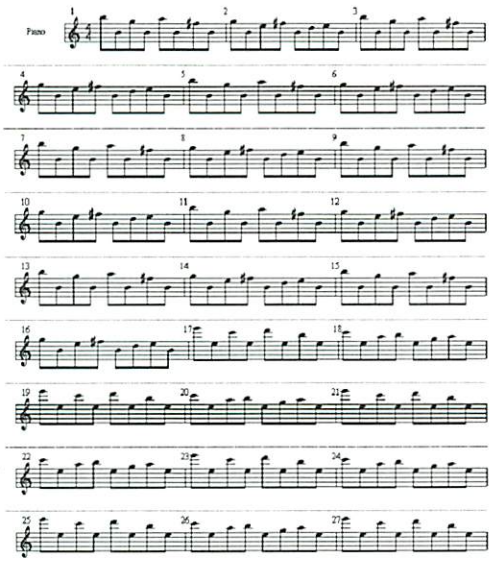

Fig.2 Musical score

is the sound that each subject hears personality and the output is the sound that each subject makes by their performance. $S_A(t)$ means the sound which comes out from Subject A. This sound reaches to Subject A directly and reaches to Subject B with the delay of the time lag Δt. So, in this situation Subject A hears not only sound of $S_A(t)$ but also $S_B(t-\Delta t)$, and Subject A hears not only sound of $S_B(t)$ but also $S_A(t-\Delta t)$ in the same time.

Each subject is separated by the board and the direct interaction is intercepted. Each subject is asked to synchronize with partner but mainly they are able to play freely. The timing of the key hitting is recorded through a session.

3.1 Experimental method

The performance of each subject was recorded by the application, which works on PC (SONY VAIO, Windows XP) and application sounded the MIDI instrument (Roland SC-8850) through headphone. Each subject played MIDI keyboard (Roland PC-300) by hearing the performance of himself and the partner. We used three performers who were special educated in music more than 10 years. They are more than 20 years old and all male right handed. The music score is demonstration music composed by YAMAH as shown in Fig. 2 and its length is 35sec. Each subject has enough skill to perform the contents of the score and well practiced. The experiment was done every pair of subjects and time lag is settled to 11 kinds of length (0,10,20,...,100,(ms)). In this research, our analysis was done for situation without the jitter, so the value of the

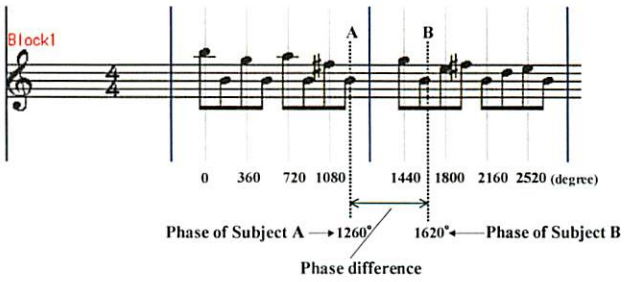

Fig.3 Phase and phases difference

time lag is not changed through one session.

3.2 A definition of an index

In this system, recorded data is the timing of the key hitting, and its time resolution is 10ms. The application has the score information and able to calculate the performing location in the score at the time. We define the performing location as a phase, and difference of performing location between two performers as a phase difference, by following the former research [8]. We convert 1 beat as 360 degree. So in Fig.3, the phase of Subject A is 1260 degree, the phase of Subject B is 1620 degree, and the phase difference between them is 360 degree. We used the music only constructed by 8th note, so the minimum resolution of phase is 180 degree, but complemented linearly.

In this research, reason of usage of phase difference (not time difference) is to make clear the difference of performing location. It is able to define by subtract from the performing location of one subject to the other subject. Fig. 4 shows the concept of the phase difference. In this graph, horizontal axis indicates time, and vertical axis indicates performing location. In this figure, we can define the phase difference (vertical axis) from the time difference (horizontal axis). We define this difference as "global phase difference".

Against this global phase difference we also define "local phase difference".

When Subject A and Subject B make ensemble, Subject A hears not only his performance but also performance of Subject B. It is the same and symmetrical to Subject B. In this situation, if there is a time lag, the Subject A hears the sound of Subject B that already performed before. Imagine that Subject A and Subject B performed symmetrical in the same time, under the large time lag. In this situation, the global phase difference is 0. But subjectively in each performer, there is a difference of timing between sounds come from Subject A and come from Subject B because of time lag. So in this situation, although in globally the timing of each performance is synchronized, the subjective timing of each performance is different. Thus, we defined "local phase difference" against to "global phase difference". Fig. 4 is a concept of local phase difference. In this situation, we make an attention to the phase difference between Subject A's direct phase and Subject B's delayed the phase. So this is local phase difference of Subject A.

3.3 Method of recording and analysis

We calculate average and S.D. (Standard Deviation) of phase difference. These values are the indicators of the smoothness and the harmonics of the ensemble. The aver-

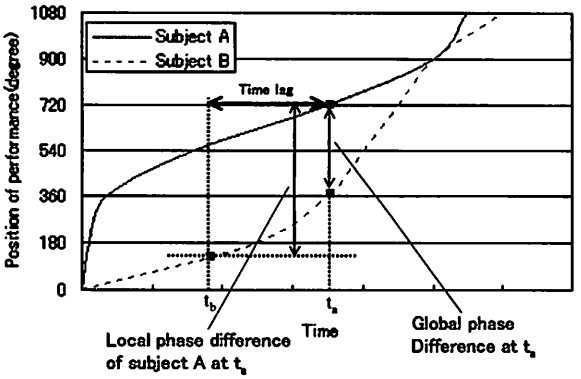

Fig.4 Conceptual figureof Phase difference

age of the phase difference indicates the relationship between performers. If this value is 0, it means their relationship is symmetrical, and if this value is different from 0, it means their relationship is differentiated to leader and follower. The S.D. of phase difference indicates how they synchronize each other. If this value is 0, it means that their synchronization is perfect and there is no difference between the performance of Subject A and B.

4. Results

4.1 Temporal development of global phase difference

Fig. 5 is the result of the temporal development of global phase difference of one pair.

If the time lag is smaller enough, the temporal development of phase difference fluctuates near 0. But when time lag is larger than 60ms, the range of phase difference value is shifted to positive. And when time lag is larger than 80ms, the fluctuation of phase difference increased rapidly and this means the failure of harmonious ensemble.

Fig. 6 shows the average and S.D. of global phase difference. Average of phase difference increase little in time lag of 0-20ms and stable till 50ms. But when time lag reached to 60ms, average of phase difference increase rapidly though S.D. doesn't change significantly. And after 80ms, both average and S.D. change significantly. Average converge again to 0 and S.D. increase rapidly.

The reason why average converge again to 0 is fluctuation of phase difference is too large as shown in Fig. 5(e), (f) and positive and negative phase difference canceled each other. And S.D. increases with the fluctuation of phase difference. From these factors, it is able to understand that in this stage ensemble is broken down. The fact that such phenomenon occurs after a certain time lag means that there is a limitation to consist harmonious ensemble.

The drastically change of average and S.D. at the failure of ensemble is explained as above, but also there is a big change of condition in the harmonious ensemble. It is range of over 60ms and temporal development of phase difference shifts to positive and only average increases drastically. In this stage, there are some changes of ensemble condition. In next chapter, we will show analysis by using local phase difference and discuss about the drastic in-

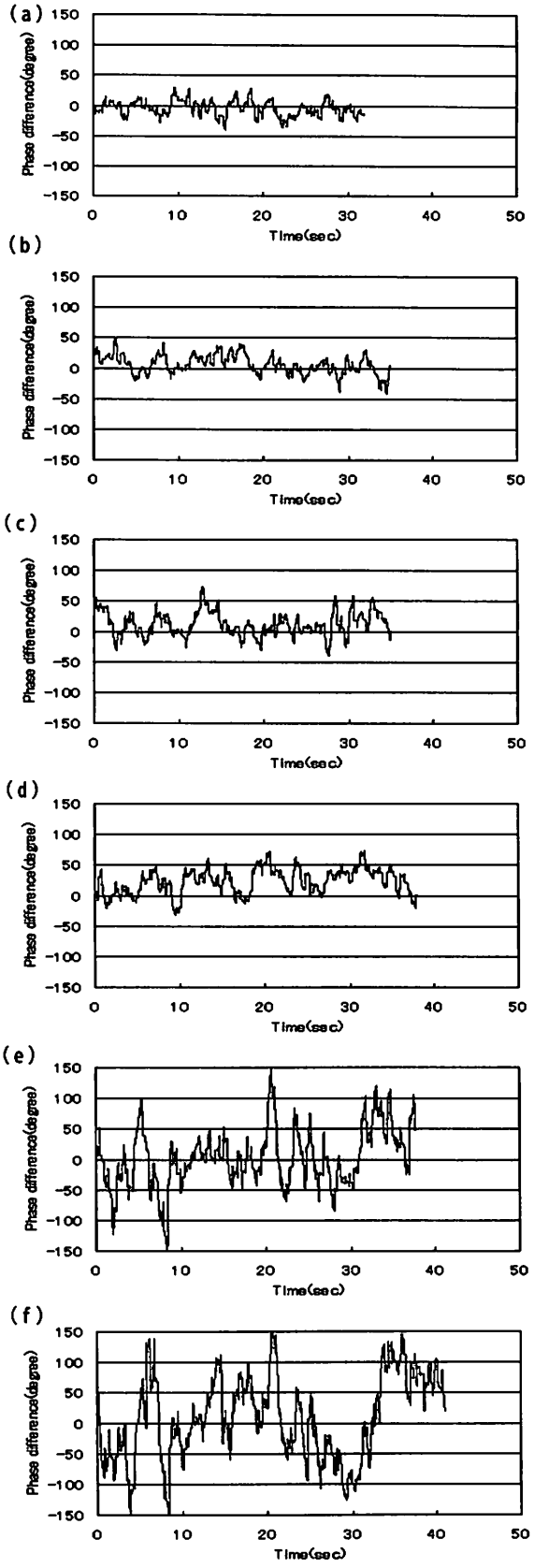

Fig.5 Time deveropment of global phase difference Time lag is (a)0,(b)20,(c)40,(d)60,(e)80,(f)100(msec)

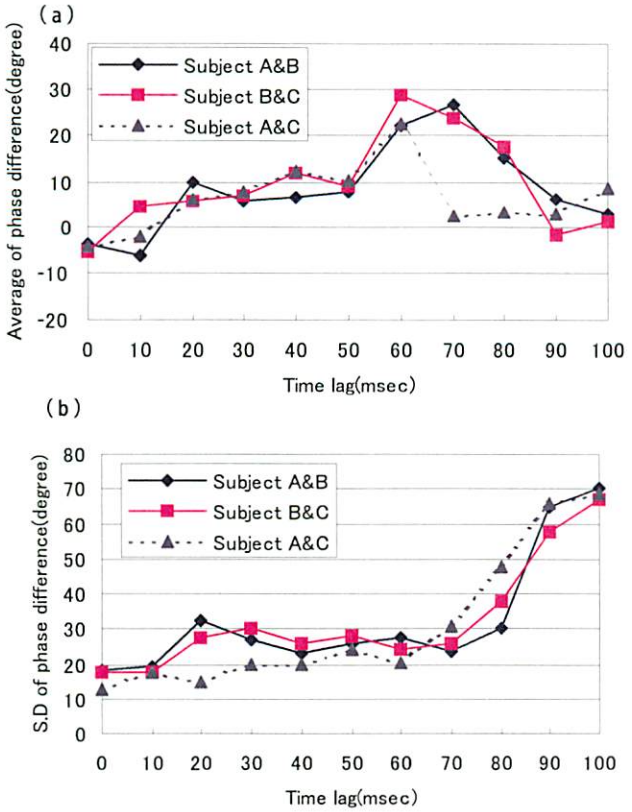

Fig.6 Time deveropment of global phase difference
(a)Average of phase difference
(b)S.D of phase difference

crease of average.

4.2 Temporal development of local phase difference

Fig. 7 shows the result of temporal development of local phase difference in the same pair as explained above. In this graph, positive plot means that the performing location is forward than the other.

In the stage of small time lag (0-40ms), temporal development of local phase difference fluctuates nearly to 0 symmetrically. But after more than 60ms, it is still symmetrical but shifted to positive and the value of each performer is different. After this stage fluctuation of phase difference increases and average returns to 0 again.

The average and S.D. of phase difference are shown in Fig. 8 and Fig. 9 respectively. In the stage of small time lag (0-40ms), both average and S.D. of phase difference keep nearly to 0. But after more than 60ms, the phase difference of one side of the pair shifts to positive than the other side and average increases. This means that one side of the pair performs quite forward to the other side. This difference is not fixed before the ensemble. For example, in Fig. 8a Subject A performs forward to the partner, but in Fig. 8c the same Subject A performs afterward to the partner. So this differentiation is determined in the relation of performers. This is able to be considered as some kinds of change is occurred in ensemble condition.

5. Discussion

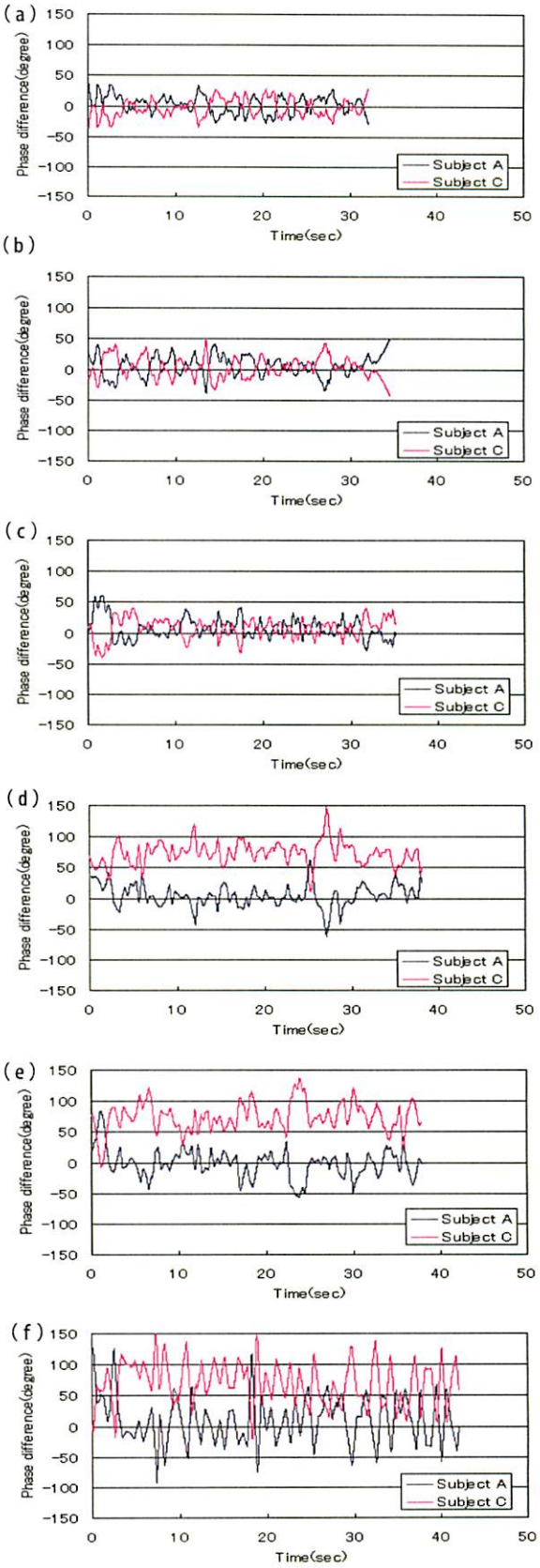

Fig.7 Time deveropment of local phase difference
Time lag is (a)0,(b)20,(c)40,(d)60,(e)80,(f)100(msec)

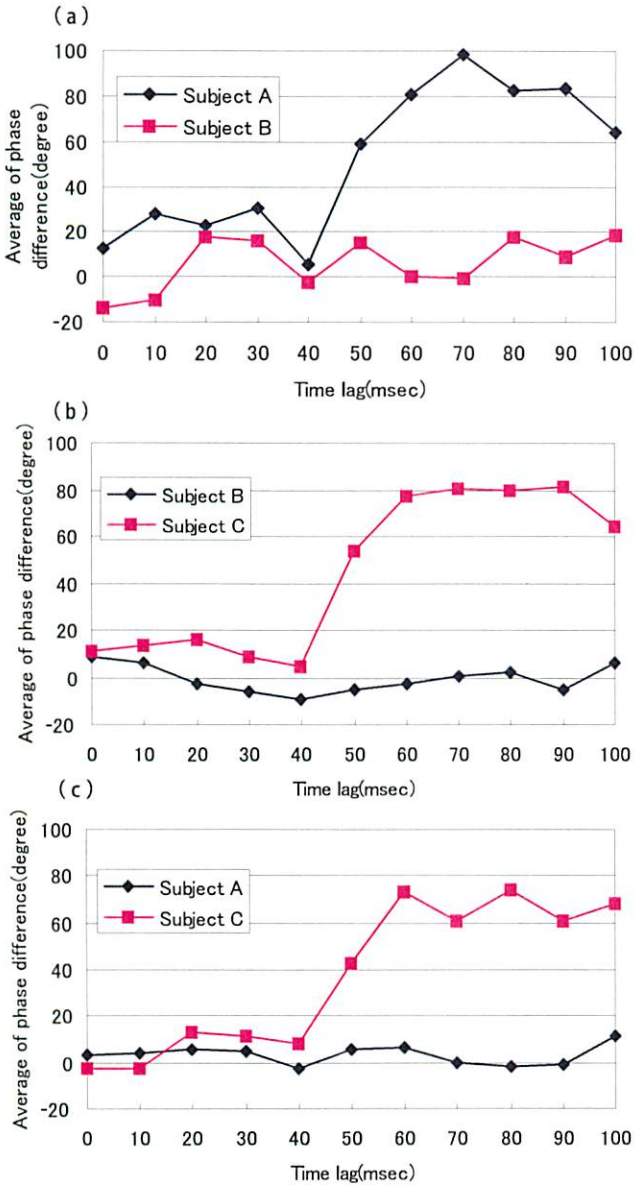

Fig.8 Time deveropment of local phase difference
(a):Subject A&B (b):SubjectB&C (c):Subject A&C

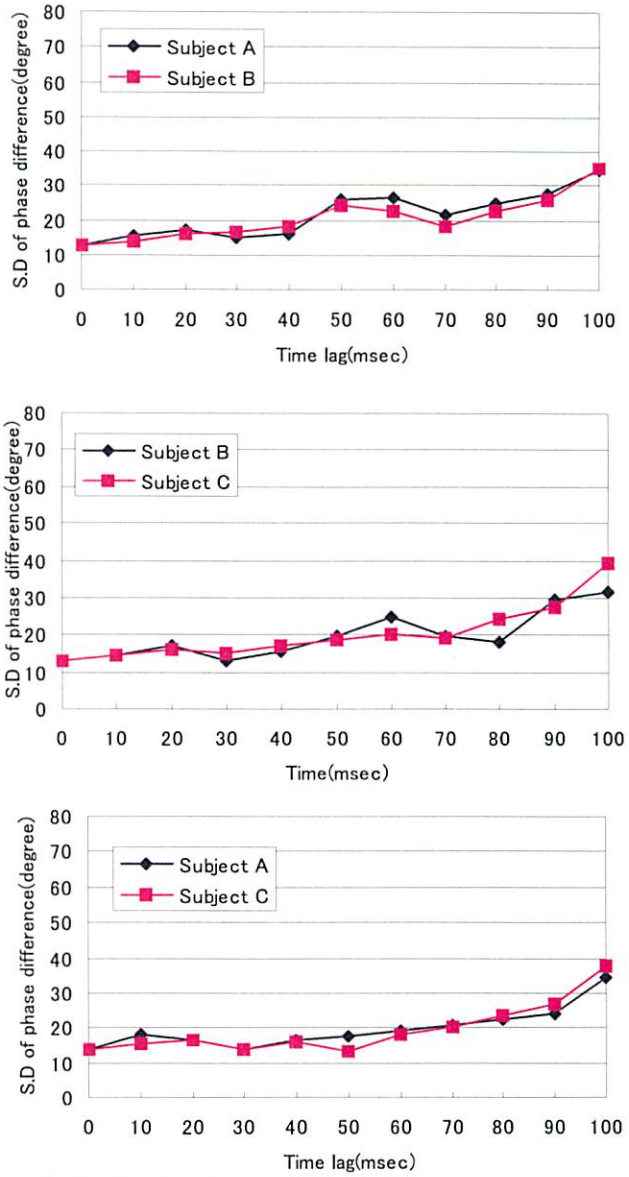

Fig.9 S.D of local phase difference
(a):Subject A&B (b):SubjectB&C (c):Subject A&C

5.1 Summary of the result

A summary of global phase difference is as following. At the range of 0-50ms, both average and S.D. keeps low values. In the range of more than 60ms, only the average increases significantly and S.D. doesn't change. Finally in the range of 80-90ms, average of phase difference returns to 0 and S.D. increases drastically.

Summary of local phase difference is as following. At the range of 0-40ms,both average and S.D. of both subjects keep low value. After more than 40ms, average of one of the pair increases drastically, and the other keeps nearly to 0. Against this, S.D. of phase difference keeps low value through this range and increase at 80ms.

5.2 Consistency and break down of ensemble

In the range of 0-50ms time lag, there is no significant difference of the S.D. of global phase difference among them. Similar to this, the local phase difference of each subject is symmetrical and its ensemble is the same as 0ms time lag. And also subjectively these time lag are not perceived. But in the rang of over 60ms, though there are no significant differences in S.D. of local phase difference, there is large difference in the average of it. From these results, in this stage, we can predict that one side of the pair performs quite forward to the other and the other side follows to it and correspond to time lag. Especially, the average of local phase difference of follower is near to 0 so subjectively follower performs exactly the same timing as leaders sound.

But in the stage of larger time lag as more than 80ms, it becomes difficult to consist ensemble harmonically and breaks down even though performers differentiate into leader and follower. This result means that leader has to lead follower by predicting the follower's sound and uses some kind of feedback mechanism. But if time lag is too

large to predict follower's sound, leader can't ignore the time lag breaks down of the ensemble.

Therefore, we can recognize there are three stages of ensemble situation in the increase of time lag. First stage is no significant difference to 0ms time lag, and second stage is differentiation of part to leader and follower. Finally ensemble breaks down by limit over of time lag. In this research the important discovery is the differentiation of the parts and a predicable function to correspond to the time lag in second stage. And if the time lag is too large, this prediction is not able to correspond to it. In present network communication, the problem of time lag is as exactly the same as the final stage and communication become chaotic. One of a method to solve this problem is to reduce time lag itself. But in preset technology, there is a limitation of transmittable speed and this approach is difficult to reach to the solution. So our suggestion is to use the prediction as used in second stage, and extend such kind of prediction by using machine and approach to the solution.

5.3 The construction of a network ensemble and communication

To make a network ensemble to practical use, one of methods is to reduce the time lag itself and pull up to a practical level. But in the situation like an ensemble between distant places, the time lag is not avoidable, and other method is needed to solve such a problem. In this situation, we want to suggest the method of human-machine-human system, which machine works as an intermediation between humans. This concept is the application of the human prediction system to machine as explained as preceding paragraph.

In present network communication, one of the main problems is how to connect humans directly. Certainly this problem is important, but human communication is not simple interaction. In this research, it was clarified that there is a prediction in human communication and such kind of mechanism is important for creative communication. We consider this kind of communication as a mutual anticipation and, we should connect humans based on this concept. So we want to suggest, human-machine-human communication system, to improve present network communication system.

6. Conclusion

In this research, we analyzed how time lag in the network affects to the ensemble condition. We discovered there are three kinds of stages under the time lag. First stage is no significant difference to 0ms time lag, and second stage is differentiation of part to leader and follower. Finally ensemble breaks down by limit over of time lag. In the second stage, the differentiation of role is observed and a kind of prediction was suggested. We supposed that this kind of prediction could be applied to the machine and able to support the ensemble by using as an agent.

References

[1] Yoichi Nagashima, Computer & Music, KYORITSU SHUPPAN CO.

[2] W.S.Con WR.Steinmetz,Human perception of jitter, IEEE JSAC,Jan.1996

[3] Mitsuhara Iijima, Soju Mtsumoto, Ichiro Fukuda and Kaaoru Sezaki, QoS Measurement for Media Sync hronization of Haptic Netw ork Service, IEICE, B-11-1, 2000.9E

[4] O.Nielsen:MIDI and audio via ISDN,proc. of ICMC, pp. 451-454,1994

[5] K.Hirata and T.Aoyagi:Music server,Proc. of ICMC, pp.123-136,1989

[6] T.Aoyagi and K.Hirata:Music sever system-distributed music system on local area network,Journal of Information Processing, Vol.15,No.1,pp. 1-9,1992

[7] Akira Takayama, Naoyuki Chiba, Keizo Miki and Ichihiko Shuto, Remote Ensemble System, ASJ, MA89-10,1989.7

[8] Y.Kobayashi,Y.Miyake"New ensemble system based on mutual entrainment", SICE Annual Conference 2002 (SICE2002),2002