A Jazz Session System for Interplay among All Players

— VirJa Session —
(Virtual Jazz Session System)

School of Science and Engineering, Waseda University

Masataka Goto   Isao Hidaka   Hideaki Matsumoto

Yosuke Kuroda   Yoichi Muraoka

No.1   *1996/08/23 ICMC96*

---

1. Introduction

☐ What is Important in Jazz Session?

- **Interplay**

  **All players improvise together
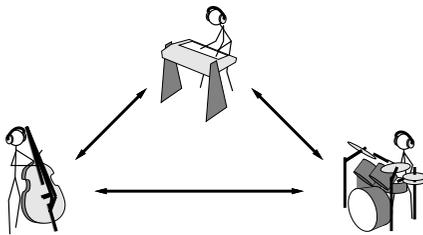 while reacting to other players' performances
  with no leader-follower relationship**

- **Multimodal Interaction**

  **Communicate by musical sounds and
  additional visual information such as gestures**

No.2

☐ Our Goal

- **Build a computer system that enables
  interplay among humans and computers**

- **Simulate the actual interaction
  that occurs among human players**

No.3

☐ Related Work

- **Jazz session / Automatic accompaniment**

  (Nishijima et al., 1992   Kondo et al., 1993   Wake et al., 1994
  Dannenberg, 1984   Vercoe, 1984   Baird et al., 1989
  Rowe, 1993   Horiuchi et al., 1993   Hidaka et al., 1995)

  Generate several accompaniment parts together
  Only react to the soloist's performance
  Fixed leader-follower relationship

- **Mention interaction among computer players**
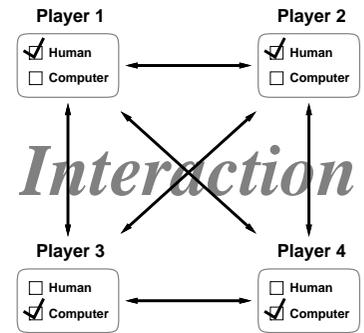
  (Kanamori et al., 1993)

  Only focus on the music-listening process
  Not generate musical performance

No.4

---

2. Jazz Session Model for Interplay

☐ Jazz Session Model

- **All players can
  listen and react to
  all other players**

- **Interact without fixed
  leader-follower
  relationship**

- **Computer players
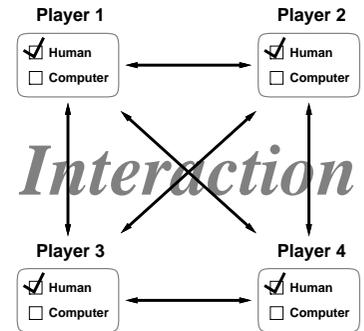 play a kind of solo**

No.5

☐ Imagine: All Players are Human Players

- **Useful for remote sessions in which
  players are not in the same physical location**
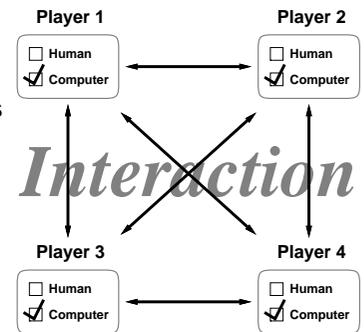
- **Exchange musical
  information through
  computer networks**

No.6

☐ Imagine: All Players are Computer Players

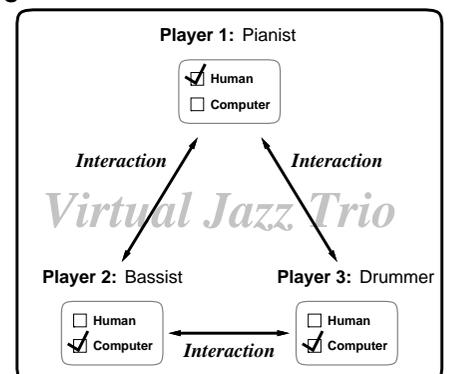- **Effective when various designers implement
  different computer players with various
  characteristics**

- **Designers' substitutes
  interact with
  each other**

No.7

☐ Jazz Session Model for Piano Trio

- **Computer players
  are executed
  as separate
  processes**

- **Player 2 listens
  to Player 3
  as well as
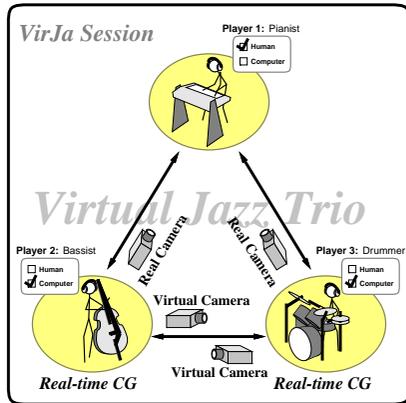 Player 1**

- **Player 2 and 3
  interact with
  each other**

No.8

3. VirJa Session (CG and Camera)

◻ VirJa Session

- **All players communicate both**

  by listening to other players

  by seeing each others' bodies and gestures

VirJa Session

Player 1: Pianist
☑ Human
☐ Computer

Virtual Jazz Trio

Player 2: Bassist
☐ Human
☑ Computer

Player 3: Drummer
☐ Human
☑ Computer

Real Camera   Real Camera

Virtual Camera

Real-time CG   Virtual Camera   *Real-time CG*

◻ Features

- **Multimodal interaction among all players using both auditory and visual information**

  Human Player: See computer players on CG

  Computer Player: See human player by camera
                    See computer player via network

- **For players:**

  Feel the presence of other players as if they were actually playing together

  Cooperate well through visual information

- **For audiences:**

  Feel the presence of all players as if they were attending a live concert

---

◻ Visualization of Computer Players on CG

- **Motions**

  (1) Playing musical instruments according to his sounds

  (2) Keeping time to musical beats
       by foot-tapping or rocking his body

  (3) Making two kinds of gestures

  (4) Nodding to show that he understood a gesture

  (5) Turning his eyes to another player

◻ Scenario

- **Song form: combination of song parts**

  theme ⇨ piano solo ⇨ bass solo
       ⇨ four verses (piano / drums) ⇨ theme (reprise)

◻ Multimodal Interaction

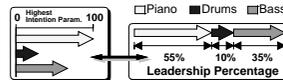- **The number of repetitions of each song part is determined dynamically by interaction**

- **Gestures**

  Leaning to the left or right

  Direct the player on the selected side to take up solo

  Pointing to his head

  Direct the other players to return to the theme

- **Computer player frequently turns his eyes to another player to watch for gestures toward the end of each song part**
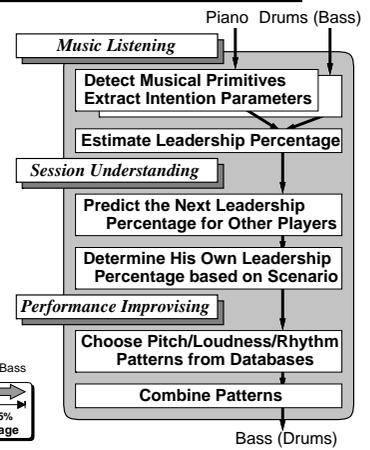
  Computer player nods to the player who made gesture

---

4. Computer Bassist and Drummer

Piano   Drums (Bass)

Music Listening

◻ Music Listening

- **Musical primitives**

  Chord note, Tension note, Scale note
  Substitute chord note, Theme note,
  Loud note, High note, Many notes

- **Intention parameters**

  Excitement, Rhythm emphasis,
  Tension, Chord emphasis
  Chord substitution, Theme reprise

- **Leadership percentage**
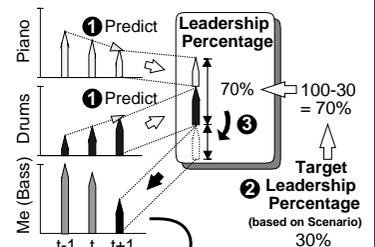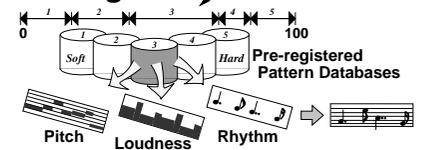
Detect Musical Primitives
Extract Intention Parameters

Estimate Leadership Percentage

Session Understanding

Predict the Next Leadership Percentage for Other Players

Determine His Own Leadership Percentage based on Scenario

Performance Improvising

Choose Pitch/Loudness/Rhythm Patterns from Databases

Combine Patterns

Highest Intention Param. 0 ... 100
☐ Piano ■ Drums ▨ Bass
55% 10% 35%
Leadership Percentage

Bass (Drums)

◻ Session Understanding

- **Consider the whole musical relationships among players**

- **Determine how much he tries to lead the session**

Piano   ❶ Predict

Drums   ❶ Predict

Me (Bass)

t-1 t t+1

Leadership Percentage

70% ⟵ 100-30 = 70%

❸

❷ Target Leadership Percentage (based on Scenario) 30%

◻ Performance Improvising

- **Select databases**

- **Combine patterns**

0 | 1 | 2 | 3 | 4 | 5 | 100

Soft ... *Hard* Pre-registered Pattern Databases

Pitch   Loudness   Rhythm

---

5. Implementation

Ⓜ Music Information   Ⓑ Beat Information
Ⓖ Gesture Information   Ⓐ Animation Information
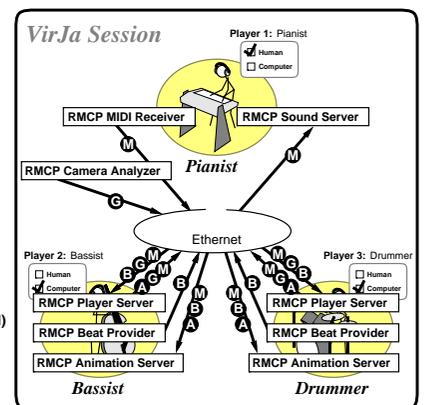
◻ Environment

- **SGI Indigo2 Impact x 1 Extreme x 2**

- **Ethernet**

- **MIDI**

- **RMCP**

  (Remote Music Control Protocol)

  Communication protocol on the UDP/IP based on the server-client model

VirJa Session

Player 1: Pianist
☑ Human
☐ Computer

RMCP MIDI Receiver   RMCP Sound Server

RMCP Camera Analyzer   *Pianist*

Ethernet

Player 2: Bassist
☐ Human
☑ Computer

Player 3: Drummer
☐ Human
☑ Computer

RMCP Player Server   RMCP Player Server

RMCP Beat Provider   RMCP Beat Provider

RMCP Animation Server   RMCP Animation Server

Bassist   *Drummer*

◻ RMCP Servers and Clients

- **RMCP MIDI Receiver**

  Take the human player's performance as input

- **RMCP Player Server**   *(computer player's brain)*

  Understand and improvise musical performance

- **RMCP Beat Provider**

  Keep the tempo of the whole performance

- **RMCP Sound Server**

  Output sounds of the computer players' performances

- **RMCP Animation Server**   *(computer player's body)*

  Display each computer player through CG animation

- **RMCP Camera Analyzer**   *(computer player's eyes)*

  Recognize the human player's gestures through camera

6. Experiments and Results

◻ Conditions

- **Four-beat jazz standard**   *'Take the "A" Train'*
- **Theme, Chord progression, Key signature provided**
- **Tempo: 187-230 M.M., constant**

◻ Results

- **Achieve a jazz session in which all players interacted without the fixed leader-follower relationship**
- **CG animation gave the pianist a greater feeling of being at a live performance**
- **Achieve multimodal interaction using sounds and gestures**

No.17

◻ CG output (bassist and drummer)

No.18

---

7. Conclusion

◻ Summary

- **VirJa Session: Vir tual Ja zz Session System**

- **All players   listen to the other players see each others' bodies and gestures**

- **Improvise without fixed leader-follower relationship**

- **Interact using both musical sounds and gestures**

- **Implemented on distributed workstations**

No.19

◻ Future Work

- **Upgrade the system**

    Follow tempo changes

    Support other configurations:
        piano trio in which all players are computer players
        sessions with different numbers of players

- **Remote jazz session**

    Our implementation facilitates remote jazz session

- **Freeware distribution (JAVA / VRML version)**

    http://www.info.waseda.ac.jp/muraoka/members/goto/PROJ/virja.html

No.20