Network-based Speech-to-Speech Translation

National Institute of Information and Communications Technology, Japan

Chiori HORI, Sakriani SAKTI, Michael PAUL, Satoshi NAKAMURA
Stand Alone and Client-server S2ST Systems

Stand alone system

Packages the entire speech translation functions into a handheld PC

Japanese

Indonesian

English

Chinese

Client-server system

Japanese speech
“おはようございます.”

English speech
“Good morning.”
Why Network-based?

- Resource limitation in stand alone systems and language pairs are limited.

- ASR/MT/TTS systems for many languages are available and needs to be maintained by each country.

- Broadband network is available.
Common Protocols on Network-based S2ST

- Speech of Language B
- Speech of Language A
- Synthesized Speech

S2ST Client

ASR
Lexicon
Speech data
Language B

MT
Parallel corpus
lexicon
Language B
Language A

TTS
Lexicon
Speech data
Language A

Parallel corpus,
Speech data, lexicon

Data format for
ASR and MT results

Communication protocol
among modules

S2ST Client

ASR
Lexicon
Speech data
Language A

MT
Parallel corpus
lexicon
Language A
Language B

TTS
Lexicon
Speech data
Language B
Lexicon for overall S2ST systems

An example of a lexicon for overall modules in S2ST systems

<table>
<thead>
<tr>
<th>Entry</th>
<th>Language</th>
<th>Attribute</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Japanese</td>
<td>Korean</td>
</tr>
<tr>
<td>Osaka</td>
<td>大阪</td>
<td>오사카</td>
</tr>
<tr>
<td></td>
<td>おおさか</td>
<td>Osaka</td>
</tr>
<tr>
<td></td>
<td>4モーラ0型</td>
<td>Da4ban3</td>
</tr>
<tr>
<td>Tokyo</td>
<td>東京</td>
<td>東京</td>
</tr>
<tr>
<td></td>
<td>とうきょう</td>
<td>Tong1jing1</td>
</tr>
<tr>
<td></td>
<td>･･</td>
<td>･･</td>
</tr>
</tbody>
</table>
Speech Translation using Distributed Service Servers

Example: From Korean to Thai Speech Translation

1. Speech recognition (Korean)
2. Language translation (Korean → Thai)
3. Speech synthesis (Thai)
Asian Network-Based S2ST System
by A-STAR Consortium

1National Institute of Information and Communications Technology (NICT), Japan
2Electronics and Telecommunications Research Institute (ETRI), Korea
3Chinese Academy of Sciences (CASIA), China
4National Electronics and Computer Technology Center (NECTEC), Thailand
5Agency for the Assessment and Application of Technology (BPPT), Indonesia
6Center for Development of Advance Computing (CDAC), India
7Institute of Information Technology (IOIT), Vietnam
8Institute for Infocomm Research (I2R), Singapore
Server Location for Network-based S2ST

Asian Speech Translation Advanced Research (A-STAR) Consortium

- CASIA (China)
- NICT (Japan)
- ETRI (Korea)
- NECTEC (Thailand)
- IOIT (Vietnam)
- CDAC (India)
- I2R (Singapore)
- RPPT (Indonesia)
NICT Network-based S2ST system

Languages:
- Japanese
- Chinese
- English
- Vietnamese
- Indonesian
Network-based S2ST Systems

Client User (Hindi)

Web-based STML Servers (ASR, MT, TTS Engines)

Client User (Malay)

Client User (Vietnamese)

Client User (Japanese)