

## インターネット接続したテレヘッドによる 聴覚テレプレゼンス\*

○ 平原達也、森川大輔（富山県立大学 工学部）、岩谷幸雄（東北大学・通研）

### 1 はじめに

テレヘッド(*TeleHead*)とは受聴者の頭部運動に追従して静粛かつ高速に運動するダミーヘッドで、その両耳に装着したマイクロフォンでほぼ完全な動的バイノーラル信号を収録できる。この動的バイノーラル信号をテレヘッドから離れた場所でヘッドホン受聴すると、テレヘッドの置かれた場所に自分の耳が在るかのように感じ、聴覚テレプレゼンスを実現できる[1, 2]。本稿では、インターネットを経由して富山県立大(射水市)と東北大通研(仙台市)の間をテレヘッドで相互に結び聴覚テレプレゼンスを実現した方法について述べる。

### 2 テレヘッド・オーバー・IP

システム全体の構成は Fig.1 に示すとおりである。

#### 2.1 テレヘッド

富山県立大のテレヘッドは四号機(*TeleHead IV*)、東北大通研テレヘッドは五号機(*TeleHead V*)である。いずれも受聴者の頭部運動に追従してダミーヘッドを旋回させるタイプである。*TeleHead IV* の追従遅延時間は約 120 ms、*TeleHead V* の追従遅延時間は約 100 ms である。ダミーヘッドは、MRI で計測し

た受聴者の頭部形状から光造形装置で成形したエポキシ製のものである。頭部姿勢情報は磁気方式のモーションセンサ (Flock of Birds あるいは Fastrak) を利用して検出し、RS-232C 経由でテレヘッド制御用コンピュータに取り込まれる。

#### 2.2 音響信号の伝送系

音響信号のインターネット伝送には、AAC コーデックを内蔵した IP テレビ会議システム (SONY, PCS-XG55) を使用した。

まず、2 台の IP テレビ会議システムを研究室内の LAN で TCP/IP 接続したときの音響信号の伝送遅延時間は約 370 ms であった。使用した LAN のパケット遅延時間は 1 ms 以下なので、この 370 ms はコーデックでの符号化と復号化に要した時間と考えられる。

次に、富山県立大と東北大通研の間で IP テレビ会議システムを WAN (インターネット) 接続し、音響信号をループバックさせたときの伝送遅延時間は 1.267 s であった。したがって、片道の伝送遅延時間は約 630 ms となる。そのときに ping で測ったパケット伝送時間は片道 23 ms であったが、WAN を介した大容量データの TCP/IP 伝送には約 260 ms (630 - 370 ms) かかっていたことになる。

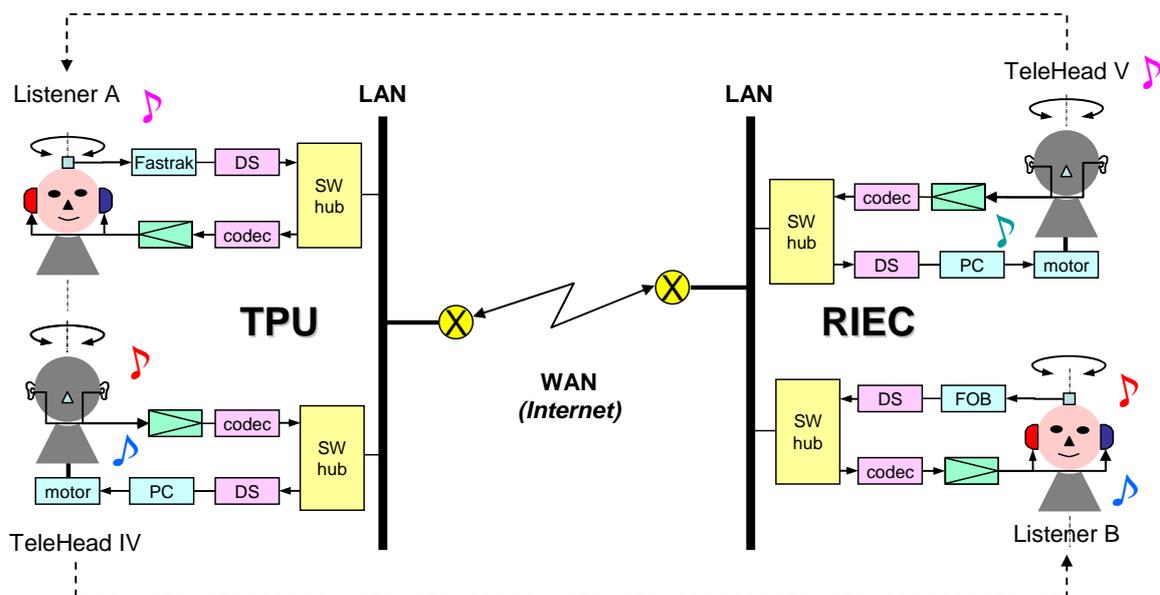


Fig.1 テレヘッドをインターネット経由で接続した聴覚テレプレゼンスシステム

\* Auditory telepresence via *TeleHead* over IP, by HIRAHARA Tatsuya, MORIKAWA Daisuke (Toyama Prefectural Univ.), and IWAYA Yukio (RIEC, Tohoku Univ.)

バイノーラル信号を伝送する場合、左右チャンネルがそれぞれ正確に信号スペクトルの強度を符号化し、伝送し、復号することとともに、左右チャンネル間の時間関係を保持して信号を符号化し、伝送し、復号することが必要である。そこで、研究室内の LAN を介して 2 台のテレビ会議システムを IP 接続し、テレヘッドの両耳に装着した 2 つのマイクロフォンで収録した音響信号と、送信側・受信側合計 2 つのコーデックを経て復号された音響信号とを比較することにより、ILD と ITD がどの程度保持されているかを検証した。収録した信号はテレヘッドの周囲で移動するヒトが発話した音声である。

Fig.2 は送信側と受信側の ITD の時間変化を描いたものである。両者の間には伝送遅延時間 370 ms のずれが認められるが、それを補正した後の送信側と受信側の ITD の差異は最大  $20 \mu\text{s}$  であった。

Fig.3 は送信側と受信側の ILD の時間変化を描いたものである。ITD と同様に両者の間には伝送遅延時間分だけずれが認められ、それを補正した後の送信側と受信側の ILD の差異は最大 0.5 dB であった。

Fig.4 は送信側の信号スペクトルと受信側の信号スペクトルとの差異、すなわちコーデックの伝達周波数特性を描いたものである。同図より伝達特性は、100 Hz 以下では -6 dB 以下、200 Hz ~ 5 kHz の間では  $\pm 2$  dB、10 kHz で -5 dB、14.3 kHz 以上では -10 dB 以下であることがわかった。

### 2.3 頭部姿勢情報の伝送系

RS-232C を使用する頭部姿勢情報のインターネット伝送には、UDS2100 (Lantronix) を用いた。これはシリアルポート対応機器を UDP/IP 接続するデータサーバー(DS)と呼ばれる装置である。

富山県立大と東北大通研に UDS2100 を設置して、WAN を介して頭部姿勢情報をループバックさせて *TeleHead V* を動かしたときの遅延時間を計測した。その結果、受聴者とテレヘッドの頭部運動の遅延時間は約 140 ms であった。*TeleHead V* の動作遅延は約 100 ms、ping で計測したネットワーク遅延は 45 ms なので、WAN のラウンドトリップ時間だけ遅れて受聴者の頭部姿勢情報がテレヘッド・コントローラに届いていることになる。

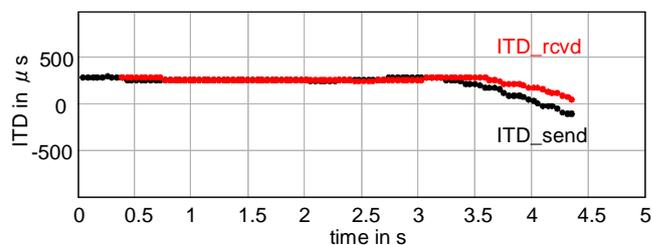


Fig.2 送信側と受信側の ITD の時間変化

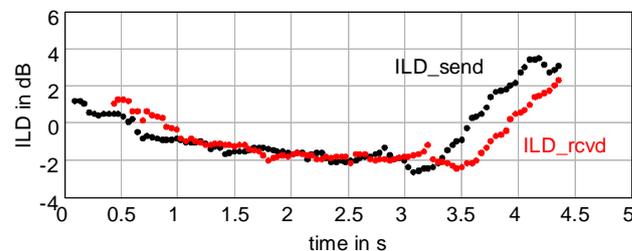


Fig.3 送信側と受信側の ITD の時間変化

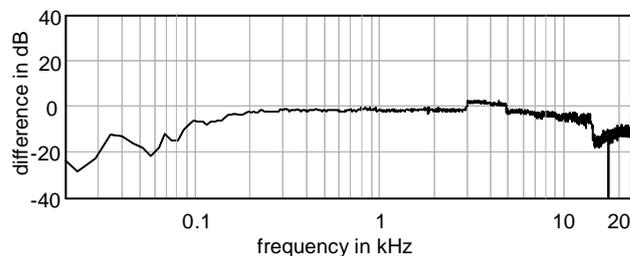


Fig.4 送信側と受信側の信号スペクトルの差

## 3 まとめ

インターネット経由で富山県立大と東北大通研のテレヘッドを相互に接続して、双方の音場を「リアル」に受聴できた。現状では、頭部姿勢情報と音響情報とは同期が取れていないし、音響情報には大きな遅延がある。しかし、頭部運動を反映した動的バイノーラル信号 [3] を用いているので、立体音場を「リアル」に知覚できる。実時間音声コーデックを利用して音響信号の遅延時間を短縮することと、受聴者が感じる立体音場の「リアルさ」を評価することは今後の課題である。

### 謝辞

本研究は東北大学電気通信研究所における共同研究プロジェクト研究による。

### 参考文献

- [1] 平原達也, 戸嶋巖樹, 植松尚, "頭部の 3 次元運動に追従するダミーヘッドシステム-テレヘッド (*TeleHead*)-," 人工知能学会研究会資料 SIG-Challenge-0216-8, 45-52, 2002.
- [2] 平原達也, "テレヘッドを通じた空間音響事象の知覚," 音講論, 1427-1428, 2009.09.
- [3] 平原達也, "バイノーラル信号による音像定位技術 —動的バイノーラル信号の効用—," 騒音制御 **33**(3), 204-211, 2009.