

頭部運動を許容する 立体音響再生技術の開発について

関西大学 環境都市工学部 建築学科
建築環境工学第一研究室 音環境ゼミ
建 10-25 亀川 直人
指導教官 豊田 政弘 助教

目次

1	序論	- 2 -
1.1	研究背景	- 2 -
1.2	目的.....	- 3 -
2	使用機材・ソフトウェア	- 4 -
2.1	ヘッドホン	- 4 -
2.2	Kinect	- 5 -
2.3	Trivisio Colibri.....	- 10 -
2.4	PureData	- 12 -
2.5	Mmr OSC Controller	- 13 -
3	研究方法・結果	- 15 -
3.1	センサーの種類.....	- 15 -
3.2	情報処理の流れ.....	- 17 -
3.3	実装結果	- 20 -
4	総括	- 23 -

付録

1 序論

1.1 研究背景

現在でも、コンサートホールや音楽スタジオ、スタジアムや競技場等の建築物の音響設計において大きな課題がある。それは、受聴者が最も聴き易く、かつ臨場感を感じるような最適なスピーカー配置、あるいは楽器等の配置を決定することである。

例えばオーケストラのコンサートのように、受聴者が基本的に舞台一方向に頭部を固定している場合もあるが、スタジアムや競技場等では頭部を運動させることがある。その場合、頭部運動により音の聴き取りが困難になるといった問題が発生する可能性がある。そのため、そのような問題までも考慮に含めたスピーカーや楽器等の配置が、より良い音響設計には必要不可欠であると考えられる。

1.2 目的

前節で述べた研究背景から、受聴者の頭部運動を許容したスピーカーや楽器等の配置を設計段階で仮想的に再現できる立体音響再生技術を開発し、より良い音響設計の実現に貢献することを本研究の最終目的とする。

現在まで大谷(参考文献[1])は、音像定位（音源位置を検知すること）のメカニズムに基づいた「空間的に音を聴く」ことを再現できる高精度な聴覚ディスプレイの実現を目指して、頭部運動を許容する立体音響再生技術の開発を行ってきた。PC、ヘッドホン、ヘッドトラッキングセンサー、タッチパネルのタブレット端末の機材と、通信及び信号処理ソフトウェアを使用し、ユーザがヘッドホンを装着した頭部を運動させた場合に、音源から発せられる音が運動に追従して聴こえ方が変化し、かつその音源の位置をユーザが自由に操作可能なシステムを提案している。しかしながら、受聴者一人の頭部運動と音源 1 チャンネルのみを対象としており、また建築物の構造や、受聴者の人数、それに伴う音の反射や吸音、透過といった事柄は考慮していない。

そこで本研究では大谷のシステムを発展させ、上記で述べた最終目的に向けて更に建築的実用性・応用性を高めるために、自由に操作可能な音源を 1 チャンネルから制限無しのマルチチャンネルにすることを考える。それにより、歌手や演奏者ごとに個別録音したオーケストラや複数の放送用スピーカー等のマルチチャンネル音源の仮想的な配置・再生を自由に行うことが可能となる。

2 使用機材・ソフトウェア

2.1 ヘッドホン



図1 ゼンハイザー「HD 650」(参考 url [1])

ゼンハイザー「HD 650」(図1)は高い解像度を持ち、本研究において精細な立体音響再生を可能にするオープンエア型ダイナミックヘッドホンである。表1に製品概要について説明しておく。

表1 ゼンハイザー「HD 650」製品概要(参考 url [1])

型式	ダイナミック・オープン型
周波数特性	10～39,500Hz
インピーダンス	300Ω
音圧レベル	103dB
質量:	約 260g (ケーブル重量除く)
接続ケーブル	ケーブル長 3.0m(両だし)、6.3mm ステレオ標準プラグ(ストレート型)
付属品	3.5mm 変換アダプター

2.2 Kinect



図2 Kinect 本体 (参考 url [2])

Kinect (図2) とは、Windows PC とより自然な会話を実現するためのデバイスである。ユーザとコンピューターの画期的な対話方法を可能にする次世代のインターフェイス、「Natural User Interface (NUI)」と呼ばれる。また、Kinect はアイデア次第で幅広い分野でイノベーションを起こせるデバイスであると思われる。以下にその活躍が期待される分野例を示す (表2)。

本研究においては、リアルタイムでヘッドトラッキング (受聴者の頭部運動をセンサーで読み取り)を行い、Windows PC にその情報を送信する。

表2 Kinect の活躍分野 (参考 url [3])

・ゲーム	・商品統計 (コンビニエンスストア等)
・リビングルーム (住宅)	・アパレル
・医療現場	・デジタルサイネージ
・リハビリテーション	・アミューズメント施設
・見守り支援	・危険区域における構造調査
・教育	・ヒューマンエラー防止
・人流計測	

Kinect に内蔵されている各種センサー（図 3）と、各種センサー等から Windows PC への処理の流れ（図 4）に示す。

図 3 Kinect センサー（参考 url [4]）

図 4 Kinect 情報処理の流れ（参考 url [3]）

以下、Kinect に搭載されているセンサーについて詳しく説明する。まずは距離センサーの特徴について表 3 に示す。

表 3 Kinect—距離センサーの特徴（参考 url [3]）

• 赤外線を使用して物体との距離を計測
• アプリケーションはセンサー視野内の (X, Y, Z) を取得
• X, Y: ピクセル座標
• Z: ミリメートル単位
• 周辺環境による影響
• ライティング (強い光、逆光): ノイズの原因
• 物体の性質 (反射、吸収): 透過、飽和の原因
• ただし一般的な環境では問題にならない

【スペック】

次に距離センサー及び映像センサーについて表 4 に示す。

表 4 Kinect—スペック（参考 url [3]）

解像度 (深度)	320x240 (既定)
解像度 (カラー)	640x480 (既定)
フレーム レート	～30 フレーム/秒
深度センサー	
物理範囲	約 0.8 ～ 4.0 メートル
実用範囲	約 1.2 ～ 3.5 メートル
近距離	約 0.4 ～ 3.0 メートル
その他の機能	指向性マイク アレイ
	オーディオ エコー除去
	傾斜モーター

距離センサー及び映像センサーの視野角は、図 5 と図 6 に示すように、水平視野は約 57 度、垂直視野は約 43 度（可動範囲 ± 27 度）となっている。

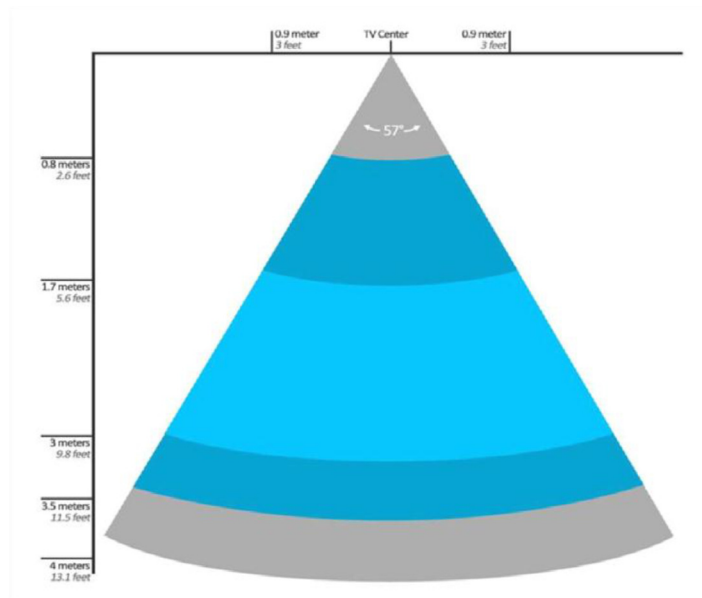


図 5 Kinect—水平視野角（参考 url [3]）

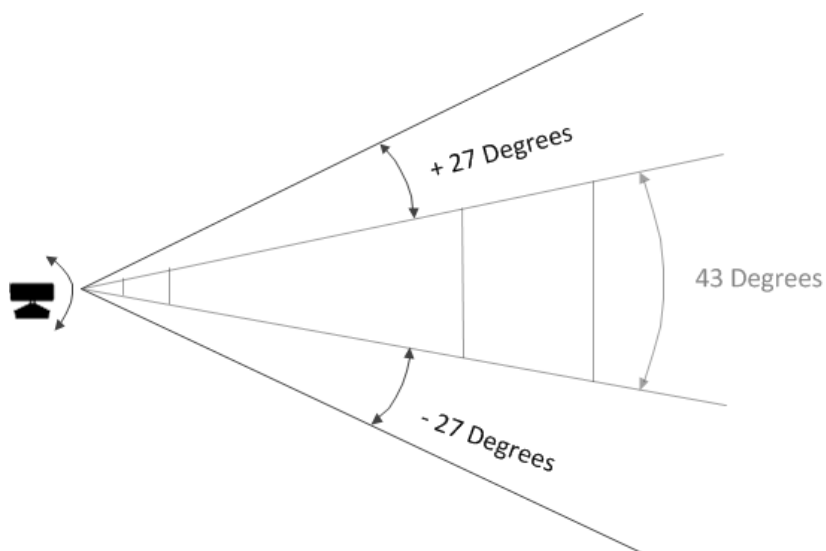


図 6 Kinect—垂直視野角（参考 url [3]）

Kinect のカメラは、視野に入った「人」の骨格を認識する。また骨格認識には、体重、体型や体格、肌の色、性別は影響しない。表 5 及び図 7 に骨格認識の詳細について示す。

表 5 骨格認識詳細 (参考 url [3])

・ 全身の 3 次元の関節座標をリアルタイムに追跡
・ 20 箇所の関節
・ 頭, 肩の中央, 左肩, 右肩,
・ 左ひじ, 左手首, 左手のひら, 右ひじ, 右手首, 右手のひら,
・ 背骨の中心, 腰の中央, 左足の付け根, 右足の付け根,
・ 左ひざ, 左足首, 右ひざ, 右足首, 左足, 右足
・ 2 人同時に骨格追跡
・ 認識そのものは 6 人 同時まで可能

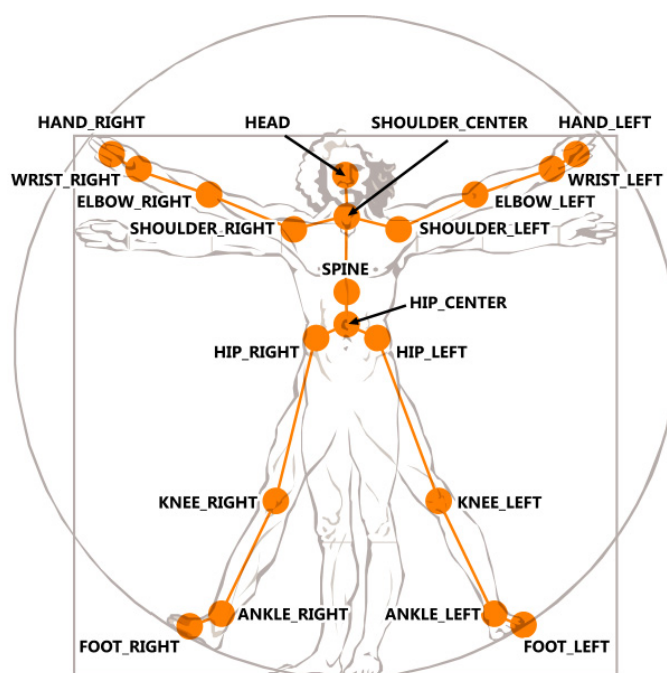


図 7 骨格認識－20 箇所の関節 (参考 url [3])

2.3 Trivisio Colibri



図 8 Trivisio Colibri (参考 url [5])

加速度計・ジャイロスコープ・磁気センサーTrivisio Colibri (図 8) は、ドイツ Trivisio Prototyping GmbH 社製の磁気加速度計である。そして Trivisio Colibri は慣性計測装置(IMU)であり、加速度・角度・磁場を計測するための最先端 3 軸センサーを持っている。また、内蔵の温度センサーは、他のセンサーに対する温度による影響を軽減する。さらに Trivisio Colibri は、センサーの生データ及びキャリブレーションされた浮動小数点データの両方を出力可能である。どのセンサーからのデータも有効、無効にすることができ、サンプリング周波数を 10 から 100 Hz まで変更することも可能である (参考 url [5])。表 6 に Trivisio Colibri の特徴、表 7 に仕様を示す。

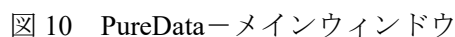
表 6 Trivisio Colibri—特徴 (参考 url [5])

・ 温度センサー内蔵
・ 低消費電力、超軽量
・ 3 軸 MEMS 加速度計
・ 3 軸 MEMS ジャイロスコープ
・ 3 軸インダクタンスセンサー
・ USB インターフェイス (バーチャル COM ポート)
・ マシン及びヒューマンフレンドリーインターフェイス
・ 位置トラッキング用拡大 Kalman フィルターを適用した Windows 及び Linux 用 API
・ 堅牢、防水、高精度アルミケース

表 7 Trivisio Colibri—仕様（参考 url [5]）

加速度計	スケール $\pm 16g$ 分解能 13 bit
ジャイロスコープ	スケール $\pm 1500^{\circ}/s$ 分解能 13 bit
磁気センサー	スケール $\pm 1000\mu T$ 分解能 $0.0263\mu T(10\text{ Hz}) \sim 0.842\mu T(250\text{ Hz})$
温度センサー	精度 $\pm 0.5^{\circ}C$ 0 $\sim 70^{\circ}C$ の範囲
動作速度	100 Hz
位置精度	ピッチ/ロール: 0.5° ヨー: 1°
消費電力	USB から 5V、40mA
PC 接続	2 メートル USB ケーブル（最長 5m 可）
寸法	30 x 30 x 13 mm
重量	22 g（ケーブル抜）

PureData とは、Miller Puckette (参考 url [6]) が開発したデータフロー型プログラミングによるマルチメディア制作環境である。また、演算機能をもった「オブジェクト」を相互に接続することで、さまざまなデータ (数値・音・映像など) に対する処理の流れ (フロー) を「パッチ」と呼ばれる図式として視覚的・対話的に構成できる (参考 url [7])。

[illegible]

2.5 Mrmr OSC Controller

Mrmr OSC controller とは 10base-t Interactive（参考 url [8]）が開発した、OSC（Open Sound Control）をサポートするアプリケーションである。OSC とは iPad と Windows PC 間で音楽演奏データをネットワーク経由でリアルタイムに共有するための通信プロトコルである。本研究では立体音響空間での直感的な音源の操作に利用する。

本研究では Mrmr OSC controller を iPad（図 11）にインストールして使用する。表 8 に仕様を簡略に示す。



図 11 iPad（参考 url [9]）

表 8 iPad—仕様（参考 url [9]）

幅	169.5mm
高さ	240mm
厚さ	7.5mm
重量	469g
ディスプレイ	9.7 インチ 2048x1536 ドット（Retina ディスプレイ） 264ppi
Wi-Fi	802.11 a /b /g /n (2.4GHz /5GHz、MIMO)

そして図 12 に示すのは一例である。

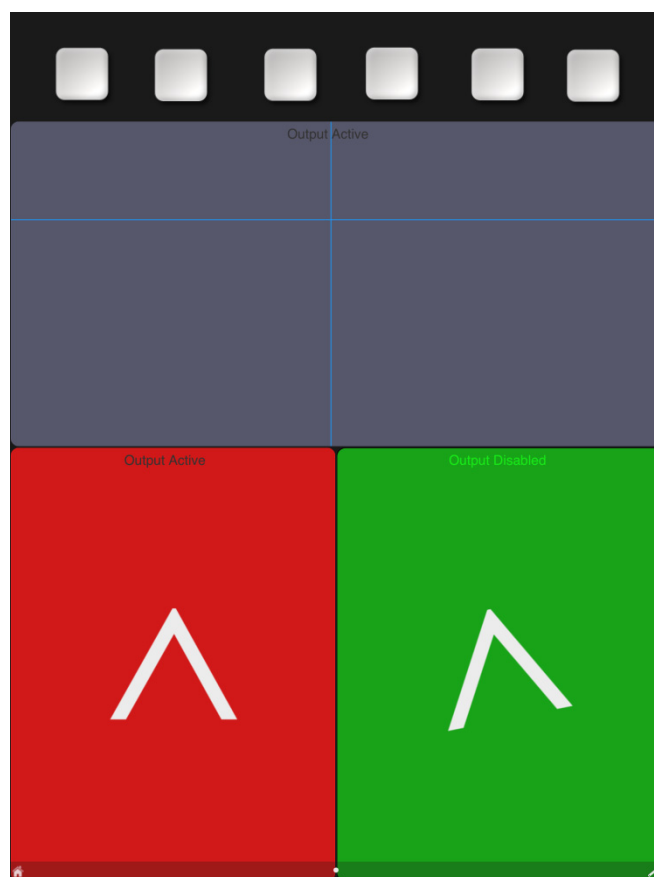


図 12 Mrmr OSC controller—一例

3 研究方法・結果

3.1 センサーの種類

本研究方法は、センサーの種類により 2 パターンに分けられる。以下に示す Kinect 使用の場合（図 13）と、Trivisio Colibri 使用（図 14）の場合である。

【Kinect 使用の場合】

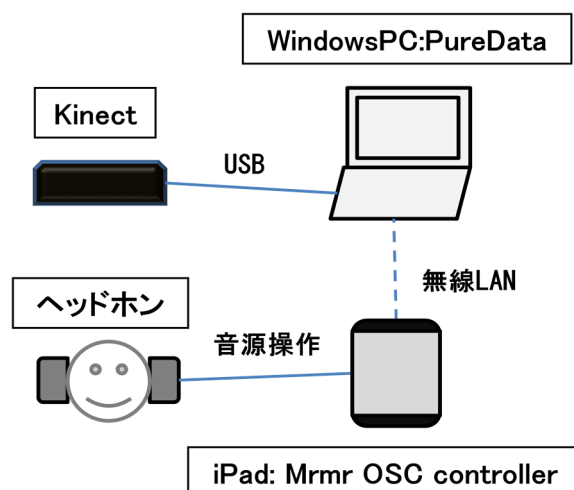


図 13 研究方法－Kinect 使用

【Trivisio Colibri 使用の場合】

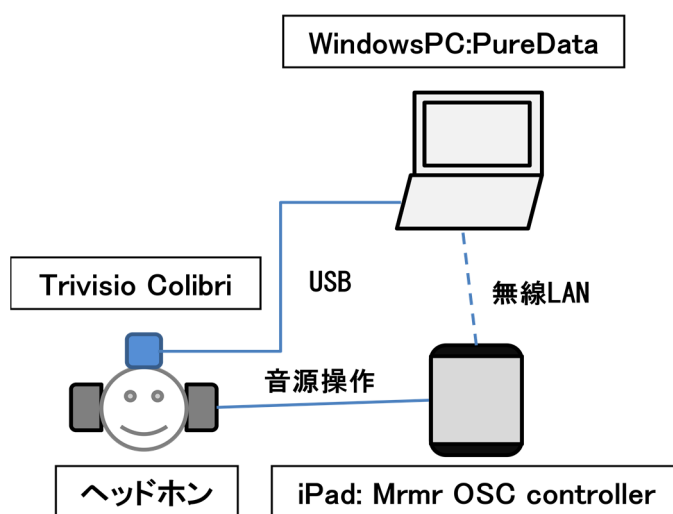


図 14 研究方法－Trivisio Colibri 使用

本研究方法の特徴は、マルチモーダルインタラクション、すなわち頭部運動とタッチパネルを使用した視覚的かつ直感的な操作に対応した立体音響再生である。

また、Kinect 使用の場合には水平な設置スペースが必要であることや、ユーザとの距離が適切でない場合と頭部の回転角が大きい場合には認識されなくなるという問題があるが、Trivisio Colibri 使用の場合はヘッドホンに装着するだけで良い。

以上のことから Trivisio Colibri の使用が望ましいと思われるが、Trivisio Colibri は Kinect に比べてコストが掛かる点が挙げられる。

3.2 情報処理の流れ

本研究方法において、どのような情報処理が行われているのかを詳細に説明する。図 13 と図 14 で示したように、ヘッドトラッキング情報を取得するセンサーは Kinect または Trivisio Colibri である。Kinect 使用の場合には、認識した頭部のセンサーからの距離 (Z 軸) と座標 (X・Y 軸)、ピッチ・ヨー・ロールの回転角の情報を取得する。一方 Trivisio Colibri 使用の場合には、ピッチ・ヨー・ロールの回転角の情報を取得する。そしてそれぞれの取得した情報は、USB を介して Windows PC の PureData へ送信される (図 15、図 16)。

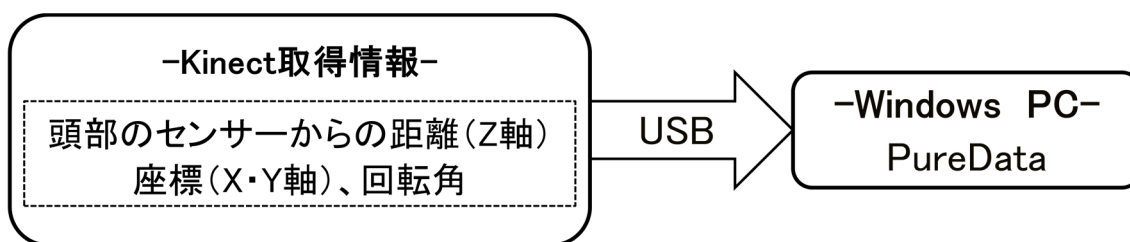


図 15 情報処理の流れ－Kinect－

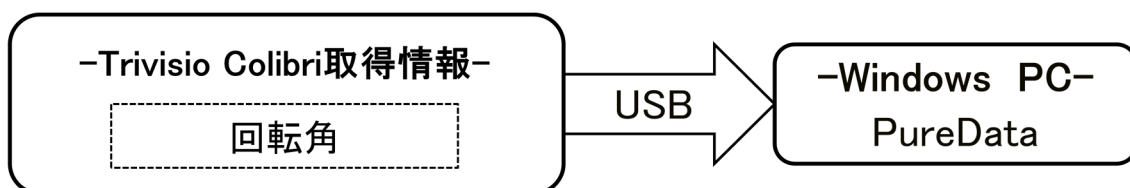


図 16 情報処理の流れ－Trivisio Colibri－

また、iPad の Mrmr OSC controller では音源の再生・停止、切替え、座標情報を無線 LAN で Windows PC の PureData へ送信する (図 17)。

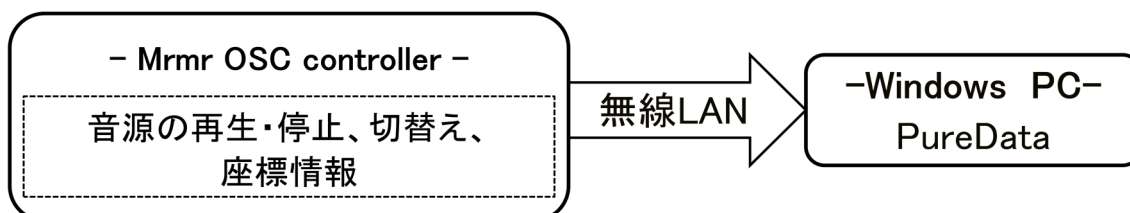


図 17 情報処理の流れ－Mrmr OSC controller－

そして Windows PC の PureData で、それらの受信した座標と回転角の情報に対応した HRTF（頭部伝達関数）を別途用意したデータベースから呼び出し、音源データとの畳み込み演算処理が行われる。その結果、立体的に構成された音響情報がヘッドホンに出力され、PureData のウィンドウには音源の座標、頭部の回転角情報が視覚的に表示される（図 18）。

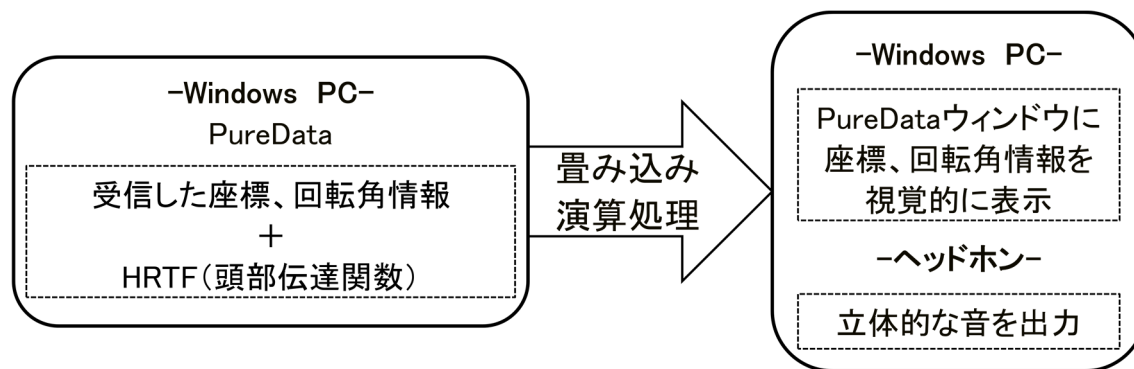


図 18 情報処理の流れ－Windows PC (PureData)－

なお HRTF（頭部伝達関数）とは、音源から発せられた音が両耳に到達するまでの伝達特性のことである。

図 19 に、図 13 と図 14 の Kinect、Trivisio Colibri、iPad(Mrmmr OSC controller)、Windows PC(PureData)で構成された情報処理の流れを簡略に示す。

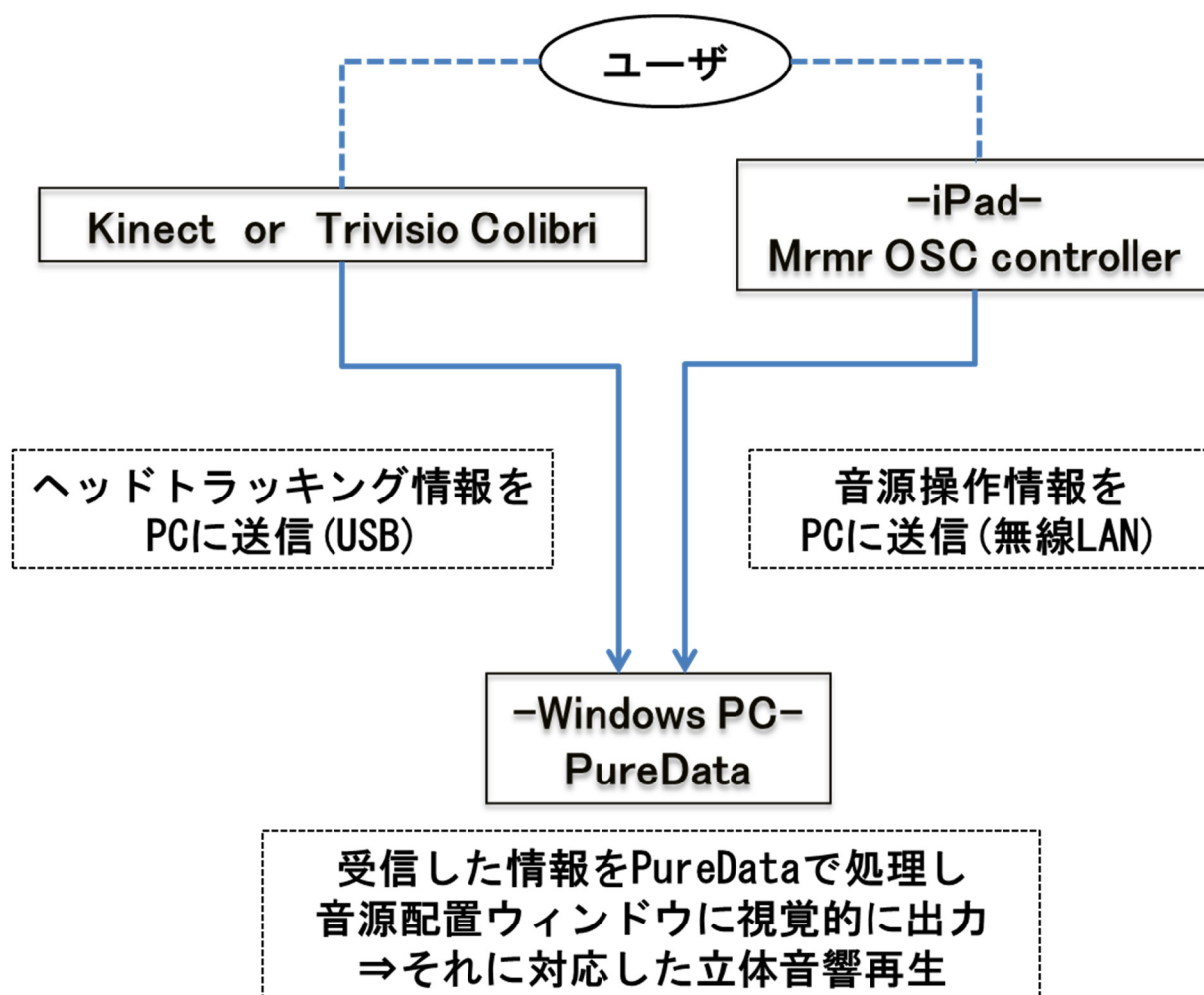


図 19 情報処理フローチャート

3.3 実装結果

今回開発したシステムの PureData パッチを付録に記載する。これにより、大谷のシステムを改良し、マルチチャンネル音源の仮想的な配置・再生を自由に行うことが可能となった。

具体例として 8 チャンネル音源を想定したものを以下に示す。PureData における音源配置とヘッドトラッキング情報を視覚的に出力したウィンドウ（図 20）、Mrmr OSC controller の音源操作画面（図 21）の実装結果を示す。また、それぞれの画面の概要説明を記載しておく。



図 20 PureData—音源配置・ヘッドトラッキング情報画面

【図 20 画面概要説明】

- ・ 1～8：個別録音音源
- ・ 円：頭部平面図
- ・ 半径の線：頭部正面方向
- ・ 四角のエリア：立体音響空間

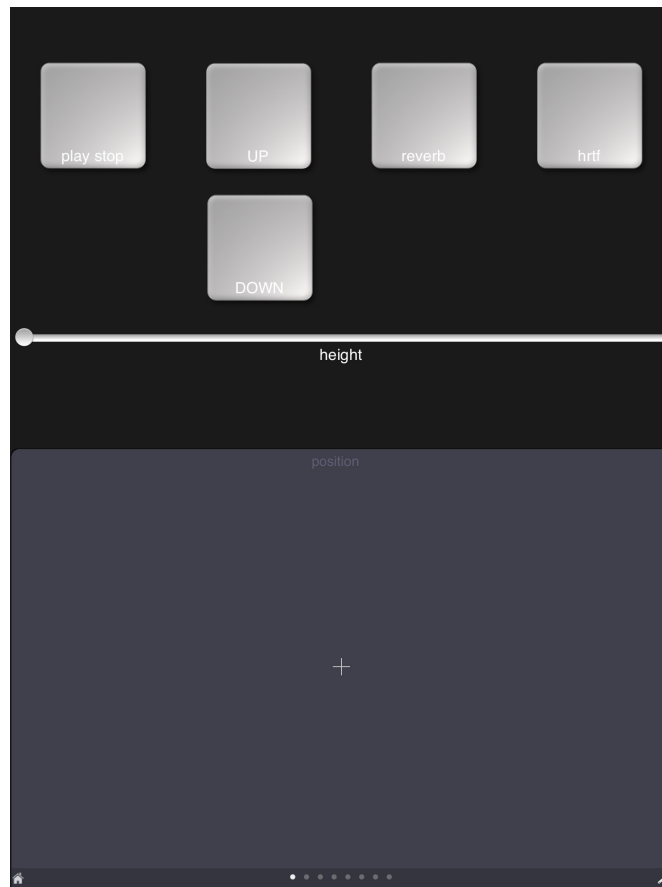


図 21 Mrmr OSC controller－音源操作画面

【図 21 画面概要説明】

- ・ play stop（左上）：音源の再生・停止
- ・ UP と DOWN（play stop 横）：音源の切替え
- ・ height（中央上）：音源の上下方向の操作
- ・ 画面下半分のエリア：図 20 のエリアと連動した立体音響空間
- ・ +：音源の水平方向の操作

以上のような、画面を見ながら操作する。なお、音源を再生した状態での音源移動も可能である。

【音源配置シミュレーション例】

以下に音源配置のシミュレーション例（図 22、図 23）を示す。音源の個数や再生される音源の種類の変化により、様々なシミュレーションが期待される。

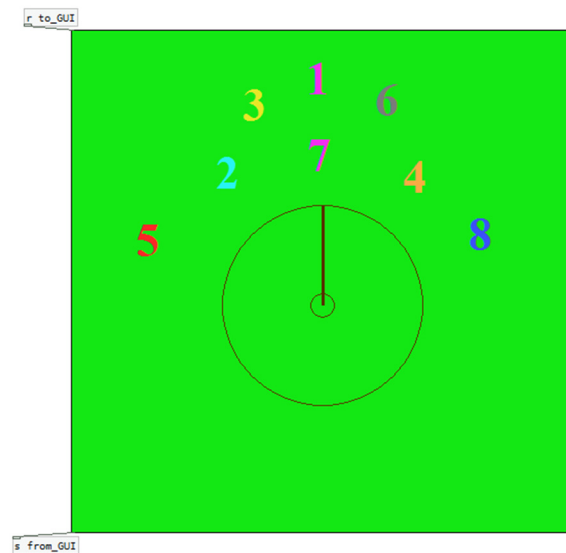


図 22 音源配置のシミュレーション例 1

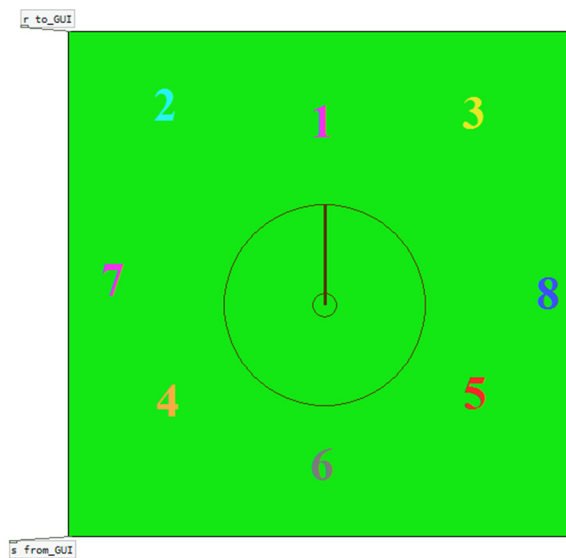


図 23 音源配置のシミュレーション例 2

4 総括

今回、マルチチャンネル音源の仮想的な配置・再生を自由に行うという目的を達成した。次の課題として、図 20 で示した **PureData** の立体音響空間内に仮想的な壁面を設置し、音の反射や吸音、透過を考慮した立体音響再生の実現を目指すことが挙げられる。最終目的達成に向け、今後もさらなる発展が期待される。

【参考文献】

[1]大谷真 「立体音響技術」 一般社団法人日本音響学会第 129 回技術講習会 東京
(2013/10/30)

【参考 url】

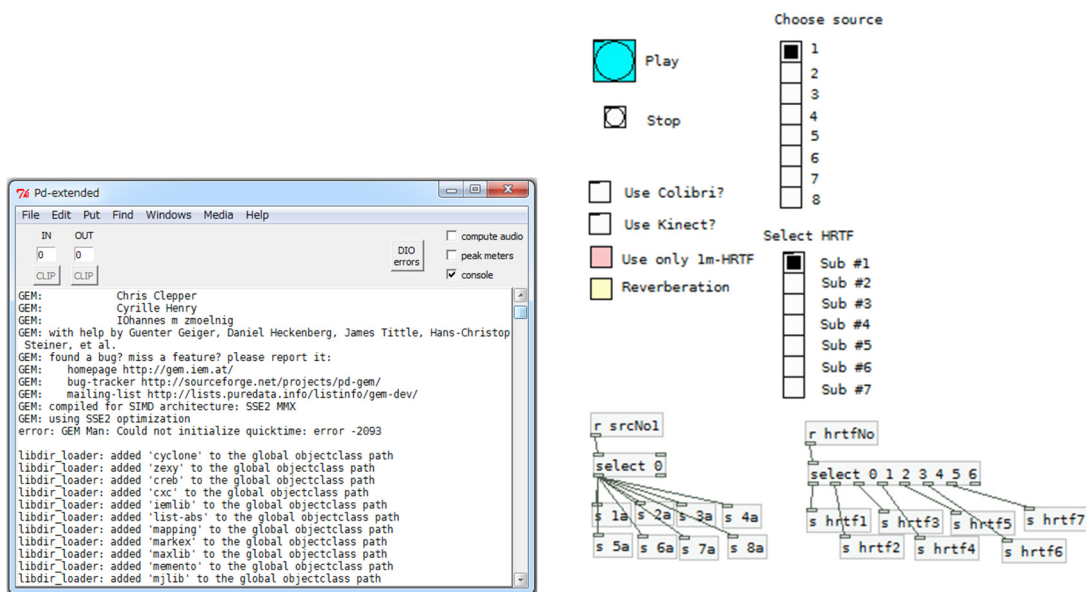
- [1] <http://www.sennheiser.co.jp/sen.user.Item/id/7.html> (2014/01/11)
- [2] <http://www.winbeta.org/sites/default/files/KinectForWindows.jpg> (2014/01/11)
- [3] <http://www.microsoft.com/en-us/download/default.aspx> (2014/12/10)
- [4]<http://www.unidux.co.jp/~media/Images/unidux/maker/manufacture/Microsoft/product/k4w.ashx>
(2014/01/11)
- [5]<http://www.nihonbinary.co.jp/Products/VR/PositionTracking/colibri/index.html> (2014/01/11)
- [6] <http://msp.ucsd.edu/> (2014/01/11)
- [7] <http://www.myu.ac.jp/~xkozima/lab/avmed-pd1.html> (2014/01/11)
- [8] <http://10base-t.com/> (2014/01/11)
- [9] <http://www.apple.com/jp/ipad-air/> (2014/01/11)

【謝辞】

研究を進めるにあたり、お忙しい中ご指導ご鞭撻をして頂いた豊田政弘助教、
快く研究データをご提供して頂いた信州大学の大谷真准教授には、心より御礼
申し上げます。

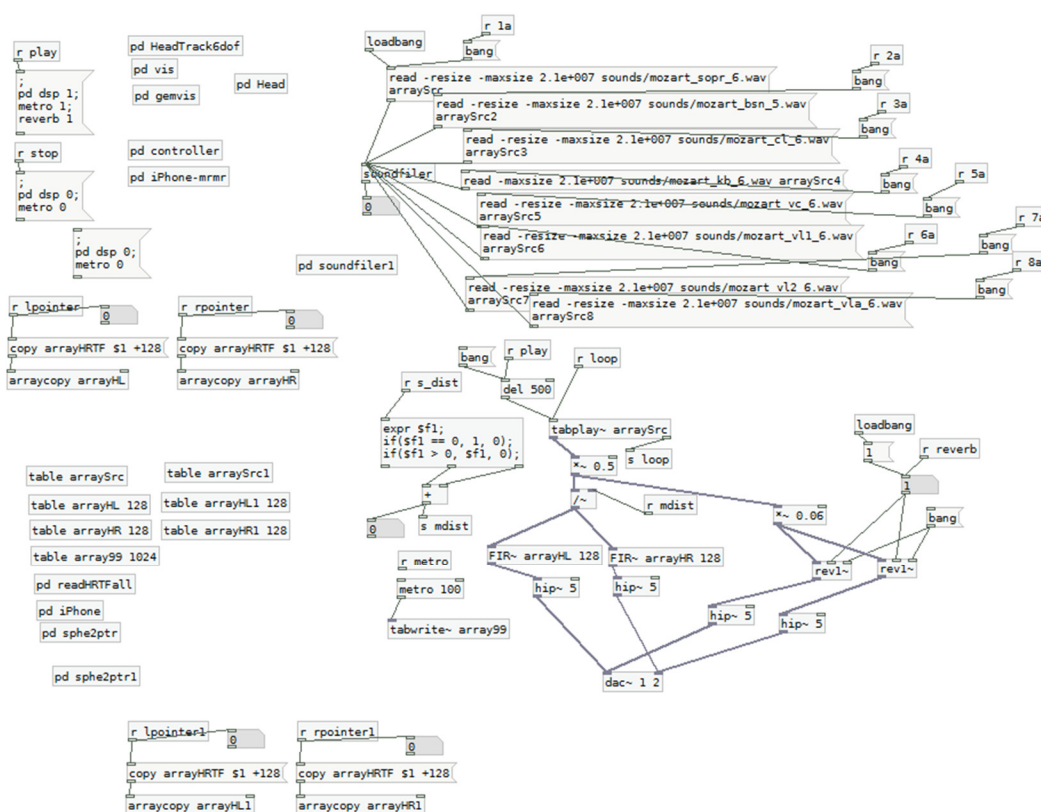
また音環境ゼミの皆様にも、健全で恵まれた環境の中で研究を行えたことを
深く感謝致します。

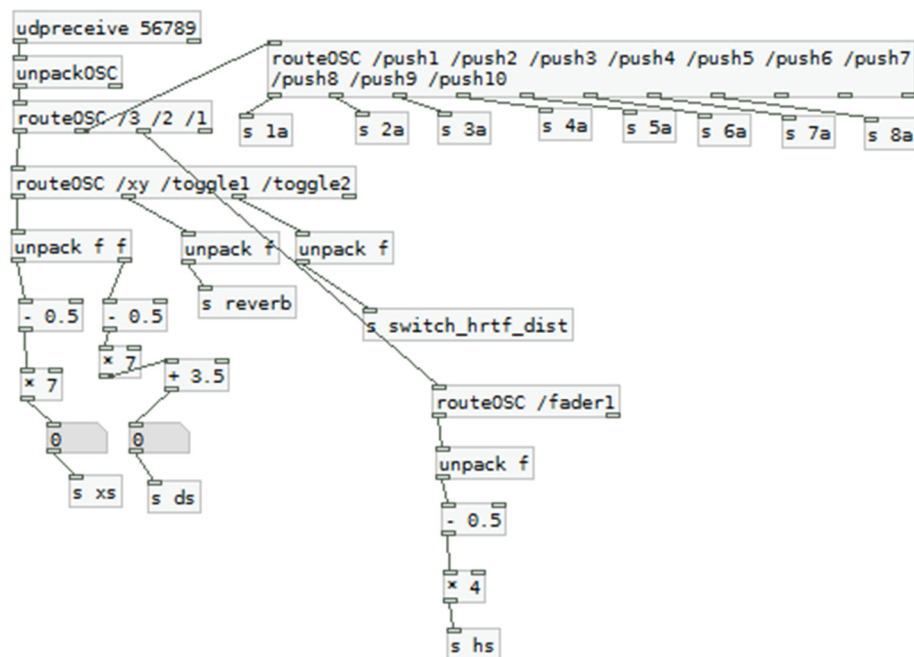
【付録】



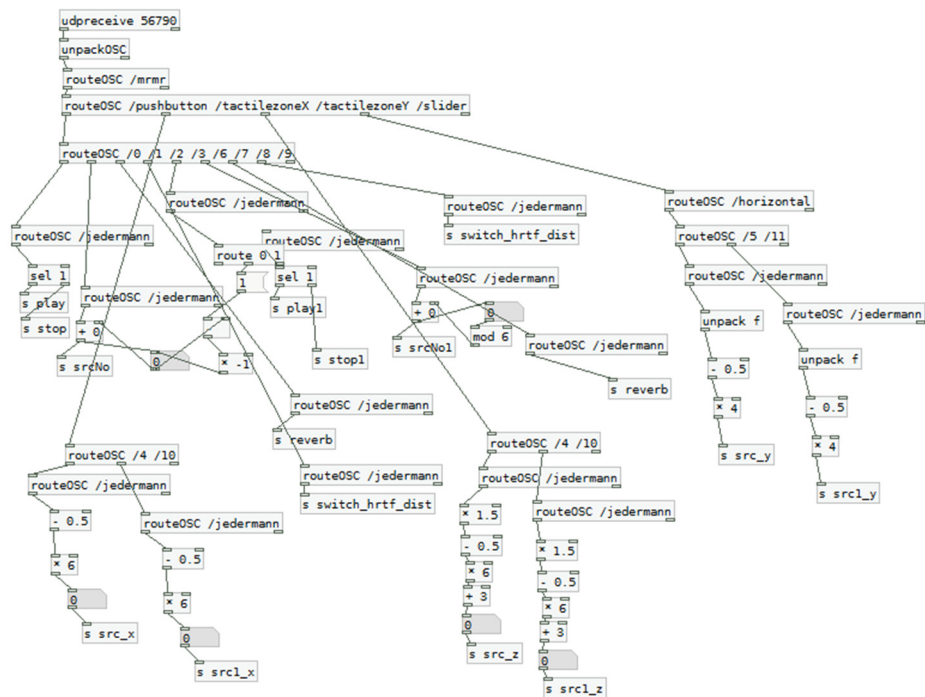
PureData—メインウィンドウ

音源切替え表示画面

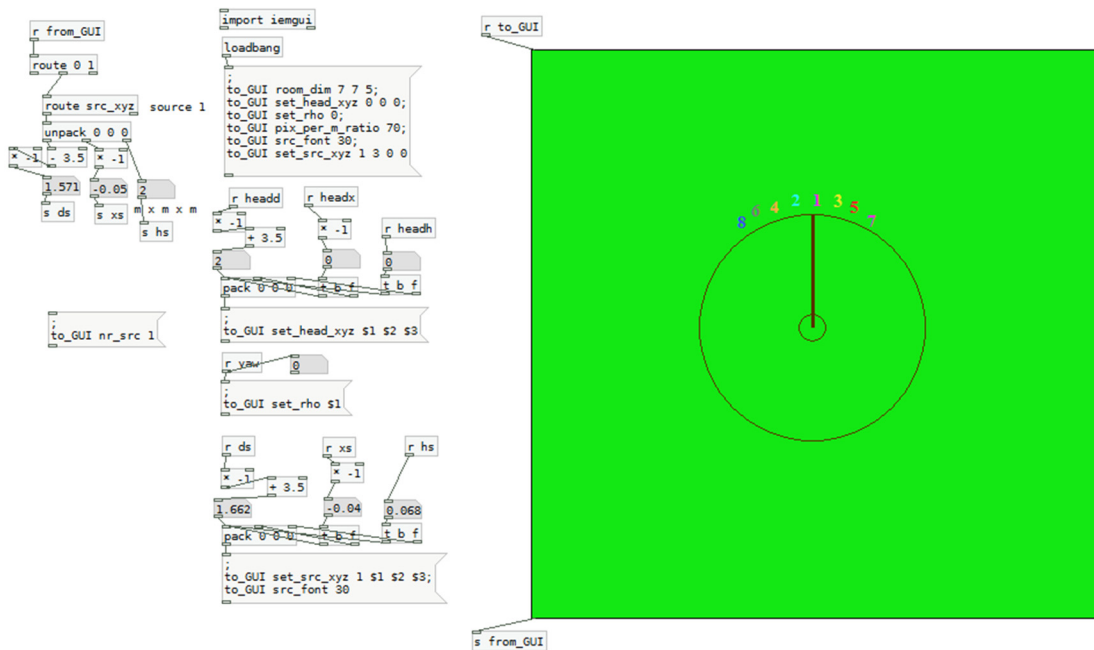




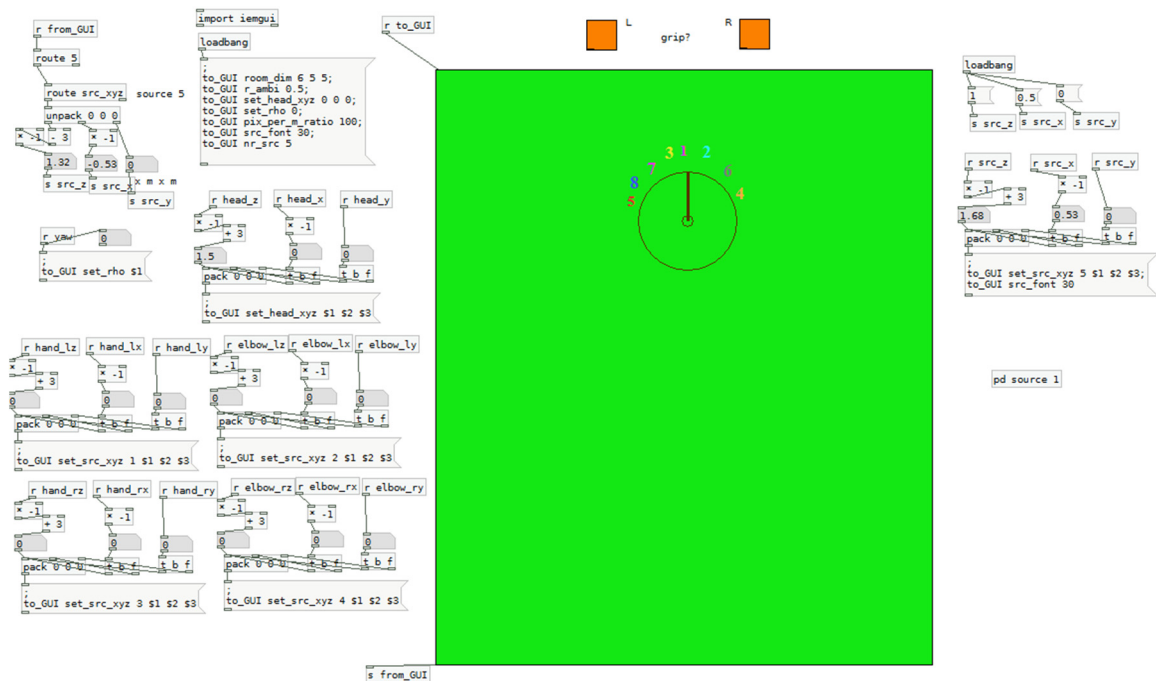
iPad から受信した情報の処理画面 1



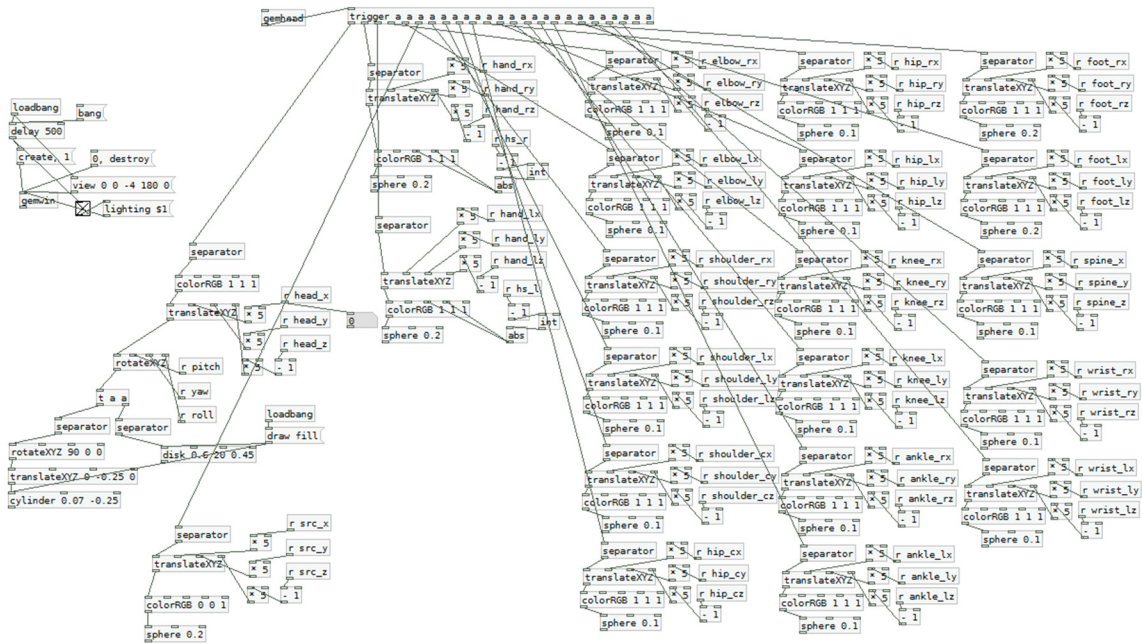
iPad から受信した情報の処理画面 2



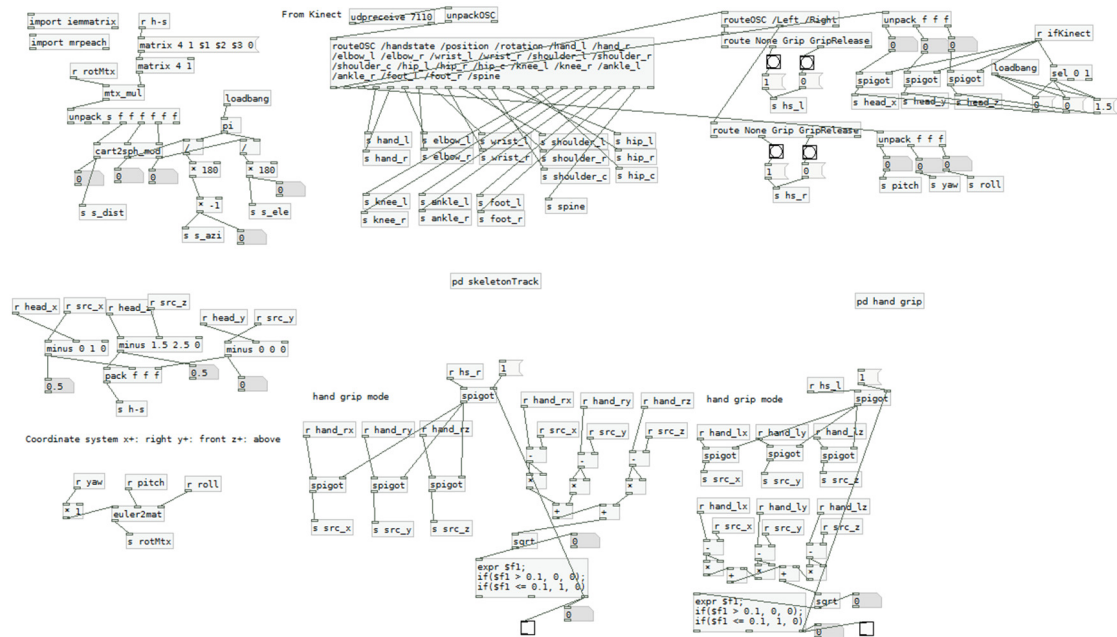
Trivisio Colibri 使用ー音源配置・ヘッドトラッキング情報画面



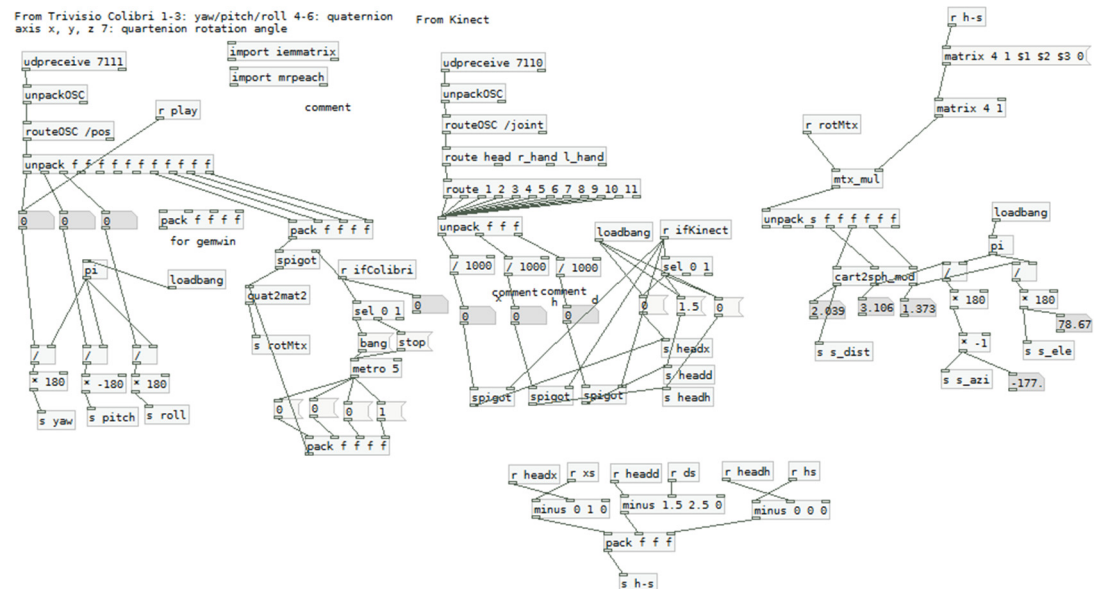
Kinect 使用ー音源配置・ヘッドトラッキング情報画面



Kinect から受信した情報の処理画面 1



Kinect から受信した情報の処理画面 2



Trivisio Colibri から受信した情報の処理画面



HRTF（頭部伝達関数）データ呼び出し処理画面

