

音声フィードバックを用いたカメラ撮影のための動体追跡支援システムの研究

瀬古圭一^{†1} 福地健太郎^{†1}

ビデオカメラで動く物体を撮影するため、音声フィードバックを用いた支援システムを提案する。ビデオカメラの撮影で、被写体を追跡する場合、カメラファインダで被写体の位置を確認しながら撮影する。カメラファインダを確認しながら撮影するのは、カメラマンが停立している場合、容易だが、カメラマンが移動している場合は、周囲の状況を把握できにくいため、衝突などの危険性がある。我々は、カメラファインダの代わりに音声フィードバックで被写体の位置を認識することで、カメラファインダからカメラマンの視野を開放し、周りの環境を把握でき、安全に動く物体を追跡しながら撮影できるシステムを提案する。本稿では、実装したプロトタイプの詳細とその評価実験について述べる。

キーワード：Sonification, 動体追跡, ビデオカメラ

1. はじめに

一般的にビデオカメラの撮影において、思い描いた構図の通りに撮影するには、カメラファインダを見て、被写体の位置を確認する必要がある。しかし、動く物体を撮影する場合、追跡しながらカメラファインダを注視していると、周囲の環境に注意が及ばなくなり、衝突や転倒の恐れがある。逆に、周囲の環境が気になって、カメラファインダに注意が及ばなくなり、思い描いた構図通りに被写体を撮影できない恐れがある。

最近では、エクストリームスポーツの撮影を主目的としたウェアラブルカメラ、GoPro[1]、CONTOUR[2]が市場に出回っている。これらの製品は、軽量でカメラマンの頭部やスポーツギアに取り付けられ、カメラマン自身の視点、一人称視点の映像を撮影するのに優れている。しかし、カメラファインダがないため、構図を確認する機構がなく、三人称視点の映像を撮るのには向いていない。

我々はこの問題を解決するため、被写体の位置を音声フィードバックによって認識できるシステムを提案し、実装した。音声フィードバックは、カメラマンの視線を自由に動かすことが可能で、動く物体を安全に追跡しながら撮影できる。音声フィードバックのデザインは、当初、立体音響を利用し、被写体の位置に合わせて、仮想音源の位置も変化させる計画だった。しかし、仮想音源の垂直成分の移動が認識しづらかったため、垂直成分の移動は、周波数変化によって表現することにした。被写体が撮影フレームの中心より上に位置する場合は、中心から離れるほど、周波数を高く、逆に中心より下に位置する場合は、中心から離れるほど、周波数を低く変化するようにした。水平成分の移

動は、立体音源をそのまま採用した。斜めの移動、すなわち、両成分を含む移動の場合は、周波数変化も仮想音源の位置変化も両方向へ。また、フレームの中心からの距離を強調して認識させるため、中心から離れば離れるほど、音量を上げるようにした。

提案システムについて評価実験を行った結果、音声フィードバックを用いた撮影はカメラファインダを用いた撮影よりも水平成分に関しては精度が高いユーザが多かったが、その効果には個人差があった。音声フィードバックの垂直成分については精度が高くなかったため、さらに改善案を考察し、次期システムのための予備実験も実施した。

2. 関連研究

カメラのインターフェースの研究では、投影型ビューファインダを採用した ClippingLight [3] が挙げられる。このシステムは、ビューファインダを投影できる平らな面が必要であるので、動く物体を追跡しながら撮影するのに不向きである。また、音声フィードバックによって物体の位置を認知するための研究としては、Seeing the world by hearing[4] が挙げられる。この研究は、視覚障害者のために、頭部伝達関数から得た立体音響によって、室内空間を認識させるシステムを実装している。

3. 提案システム

提案するシステムは、被写体のトラッキングと音声フィードバックのモジュールから成る。トラッキングモジュールは、カメラ画面上の被写体の位置を認識する機能を持ち、音声フィードバックモジュールは、被写体の位置情報から、音声を生成する機能を持つ。ユーザはステレオヘッドフォンを装着し、音声出力を聞き取りながら、被写体の位置を認識し、カメラの向きを調整し撮影する。

^{†1} 明治大学大学院
Meiji University.

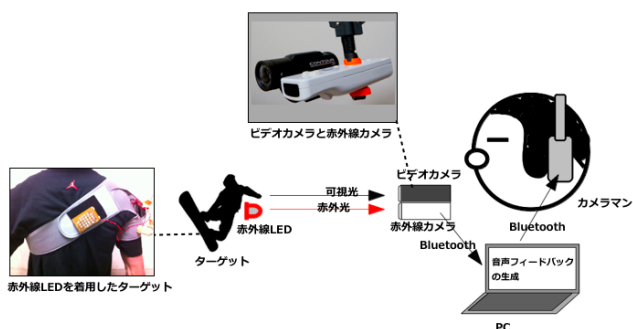


図 3 提案システムのハードウェア構成
 Figure 3 Configuration of a proposed system

4. トラッキングモジュール

トラッキングモジュールは、被写体の位置を認識するための機構である。被写体に赤外線 LED を装着して、赤外線カメラでトラッキングする手法と、ビデオカメラのキャプチャデータから被写体の色相ヒストグラムを参照しながら追跡を行う CAMSHIFT[5]を用いる手法の 2 種類を実装した。

システムの検討段階で、ロバストにトラッキングできる見込みで赤外線カメラを採用していた。しかし、赤外線カメラを採用したプロトタイプの前実験を行ったところ、撮影ターゲットの姿勢が傾くと、赤外線 LED の光の指向性から、トラッキングエラーが頻発した。本研究の評価実験では、音声フィードバックの評価を重視する方針で、トラッキングモジュールは被写体を確実にトラッキングできる環境を用意する必要があった。CAMSHIFT では、色相ヒストグラムを参照するので、暗い中、一定の色味を持つ物体のトラッキングが安定する。そのため評価実験用のトラッキングモジュールに CAMSHIFT を用いた手法を採用し、実験の環境を暗室、撮影ターゲットをプロジェクタで投影したオブジェクトにした。

4.1 赤外線カメラを用いた手法

被写体に赤外線 LED を装着し、赤外光の波長付近だけの光を通すフィルタを付けている赤外線カメラによって、被写体の位置を認識する。

赤外線カメラには Wii リモコン[6]を採用した。Wii リモコンは検出した赤外光の位置座標を算出するプロセッサが内蔵されている。Wii リモコンから音声フィードバックを生成するラップトップ PC へ Bluetooth 経由で赤外光の位置情報を転送する。

4.2 CAMSHIFT を用いた手法

ビデオカメラには SONY HDR-CX7[7]を採用した。ビデオカメラのキャプチャデータを Blackmagic Design Intensity Shuttle for USB 3.0 を経由して、音声フィードバックを生成するラップトップ PC に転送する。被写体の色相ヒストグラムは、撮影前に取得しておき、撮影時には CAMSHIFT を用いて被写体のヒストグラムを参照しながら、被写体の位置

を認識する。コンピュータビジョン向けライブラリ OpenCV[8]を利用して実装した。



図 4 左:SONY HDR-CX7 右:Blackmagic Design Intensity Shuttle for USB3.0
 Figure 4 Left:SONY HDR-CX7 Right:Blackmagic Design Intensity Shuttle for USB3.0

5. 音声フィードバックモジュール

トラッキングモジュールから受信した位置データを元に音声出力データを生成し、Bluetooth を介してヘッドフォンに音声出力を伝送する。

5.1 音声フィードバックデザイン

5.1.1 水平成分の音声フィードバック

被写体がカメラの画面上を水平方向に動いた場合は、立体音響を用い、仮想音源の位置を変化させ、まるで被写体から音が聞こえてくるかのような表現を行う。

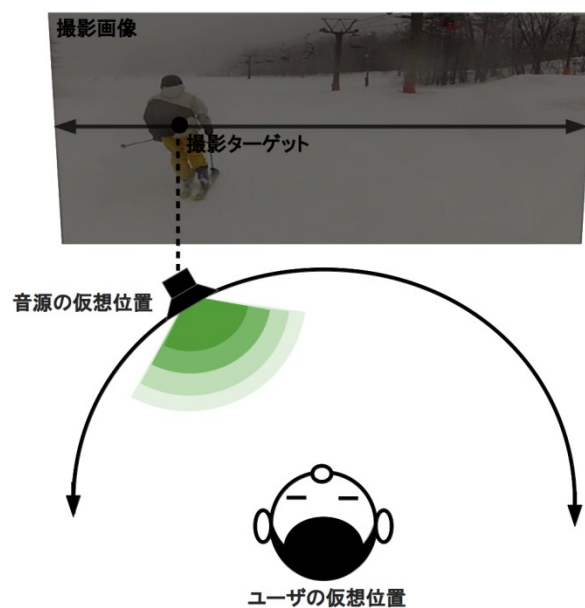


図 5.1 水平成分の音声フィードバック

Figure 5.1 Auditory feedback of horizontal component

図 5.2 のように撮像画像の中心を原点とした座標系で画像の幅が $-X$ から X とした場合、仮想音源の極座標 (r, θ) の θ は式(1)で表すことができる。ユーザの仮想位置から音源の仮想位置までの距離 r は一定の値にした。

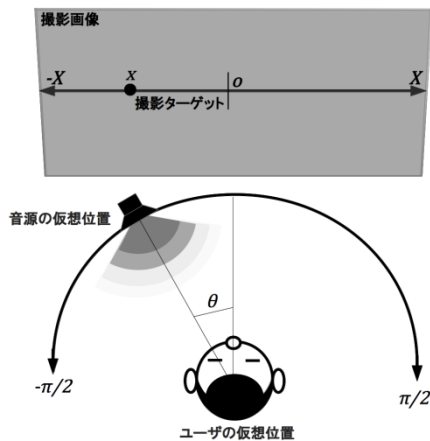


図 5.2 仮想音源の角度
 Figure 5.2 Angle of virtual source

$$\theta = \frac{\pi}{2} \cdot \frac{x}{X} \quad (1)$$

5.1.2 垂直成分の音声フィードバック

垂直成分に移動した場合は、周波数を変化させる。変化幅を 1Hz ごとにしても、変化に気づきにくいため、MIDI のノートナンバー単位で変化させた。被写体が撮影フレームの中心より上に位置する場合は、中心から離れるほど、周波数を高く、逆に中心より下に位置する場合は、中心から離れるほど、周波数を低く変化するようにした。

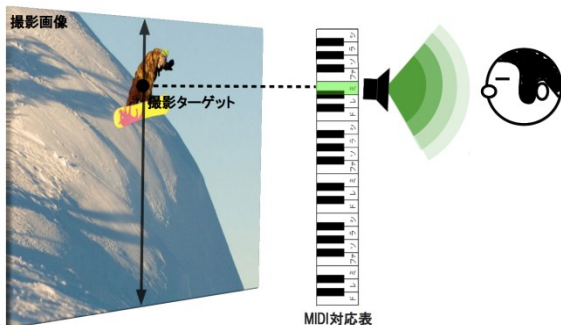


図 5.3 垂直成分の音声フィードバック
 Figure 5.3 Auditory feedback of vertical component

5.1.3 音声変化のタイミング

音声変化のタイミングは、図 5.4 のように 12x12pixel の枠から撮像ターゲットが外に出たときである。また、音声出力のタイミングが異なる 2 種類のタイプを実装した。1 つ目は、連続的に音声が出力していて、音声変化のタイミングで、立体音響や周波数を変化させる。この音声タイプを Continuous Sound と呼ぶ。2 つ目は、音声変化のタイミングで、一時音声を再生するが、次第に減衰させる。この音声タイプを Intermittent Sound と呼ぶ。システムの検討段階では、Continuous Sound のみ採用していたが、予備実験を行った際、被験者のコメントで、「ずっと音声が出力している」と

ストレスが溜り、集中を保てない。」という意見があった。この問題を解決するために Intermittent Sound を提案した。どちらの音声タイプも総合的に評価する必要がある。

○ 撮影ターゲットの中心

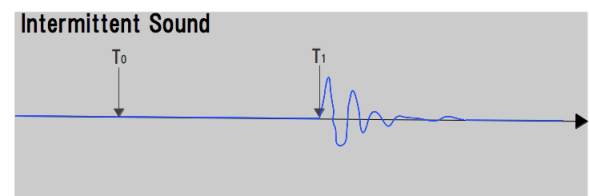
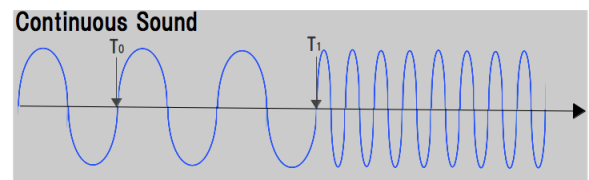
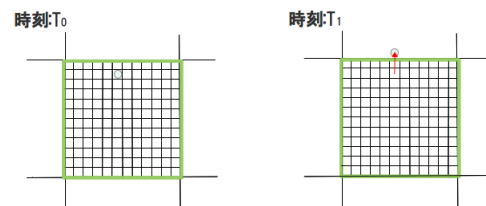
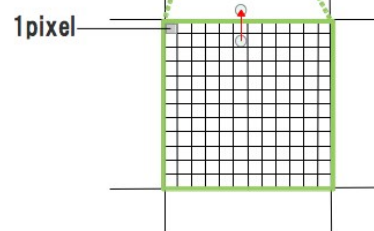
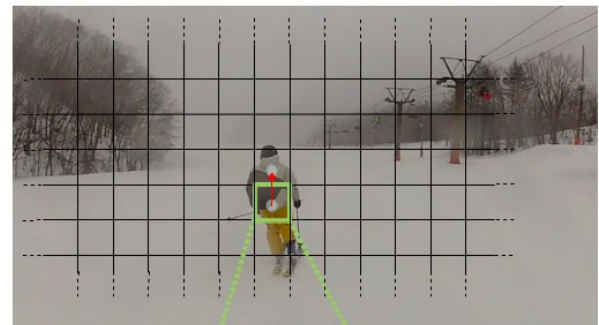


図 5.4 音声変化のタイミング
 Figure 5.4 Timing of sound change

5.1.4 音量

フレームの中心からの距離を強調して認識させるため、中心から離れれば離れるほど、音量を上げるようにした。

5.2 実装

リアルタイムに信号処理を行うことを考え、音声モジュールは PureData[9]を用いて実装した。立体音響のライブラリとして、Earplug-[10]を使用した。また、トラッキングモジ

ュールと音声フィードバックモジュール間のインターフェースは, OpenSound Control[11]を採用した.

6. 評価実験

6.1 概要

提案システムがカメラファインダより正確に撮影ターゲットを撮影できるのか検証するため, 評価実験を行った. 音声フィードバックの効果をより正確に検証するため, 撮像対象を確実にトラッキングできる環境を用意し, 実験を行った. また, 評価指標を撮影フレームの中心から撮影ターゲットまでの距離とした. 但し, 音声フィードバックの生成手法が水平成分と垂直成分とで異なるため, 指標を水平成分と垂直成分に分けて評価し, 各成分の効果を分析した.

6.2 評価設計



図 6.1 評価実験の環境

Figure6.1 Environment of Evaluation Experiment.

被験者 8 人(A, B, C, D, E, F, H)がカメラマンとなり, 撮影ターゲットの撮影を行った. カメラファインダとプロトタイプ 2 の 2 種類の音声フィードバックの計 3 種類のインターフェースでそれぞれ動画を撮影する. また, 被験者は, ランニングマシンの上で撮影し, 停立状態と走行状態の 2 種類の状態で, 撮影する. 3 種類のインターフェースと 2 種類の状態の計 6 種類の条件で撮影を行った. 撮影ターゲットは, 図 6.2 のようなスクリーン上にプロジェクタで投影した赤い円とした. ターゲットのトラッキングを全フレームで確実にを行うために, 実験中は部屋を暗くし, CAMSHIFT で起き得る認識ミスをできるだけ少なくした.

撮影時間は, 各々の条件で撮影を 1 分間行なった. 本実験で使用したカメラが 15fps なので, 900 フレームを評価対象とした.

評価指標は, 予備実験と同様, フレーム内の撮影対象の位置からフレームの中心の距離に設定した. 撮影した総フ

レームの平均の指標が短い程, 高い精度で撮影対象をトラッキングできていると評価する. ただし, 音声フィードバックが水平方向と垂直方向で異なるため, 距離を水平成分と垂直成分に分けて評価を行う.

また実験後に, 被験者に心理アンケートに答えてもらった. 4 つの心理要素, 安全性(1:危険に感じた, 7:安全に感じた), 快適性(1:不快に感じた, 7:快適に感じた), 難易度(1:困難だった, 7:容易だった), 楽しさ(1:楽しくなかった, 7:楽しかった)について質問した.

6.3 システム

6.3.1 撮影ターゲット

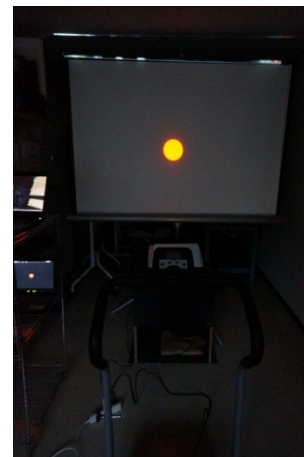


図 6.2 撮像ターゲット

Figure Target

図 6.2 はスクリーンに映った撮像ターゲットの様子である. 撮像ターゲットの動作パターンは, 水平方向に左, 右に動くパターンを 20 秒, 垂直方向に上, 下に動くパターンを 20 秒, ランダムに動くパターンを 20 秒の計 3 パターンをランダムの順序で実行する. 撮像ターゲットのアニメーションは, Processing[12]を用いて実装した.

6.3.2 ルームランナー



図 6.3 ルームランナー稼働中の撮影風景

Figure 6.3 Used treadmill

カメラマンの走行状態を実験環境で再現するため, ルームランナーを使用した. 被験者にはルームランナーの上で

撮影してもらい、走行状態の条件では、ルームランナーを稼働させた。

6.4 結果

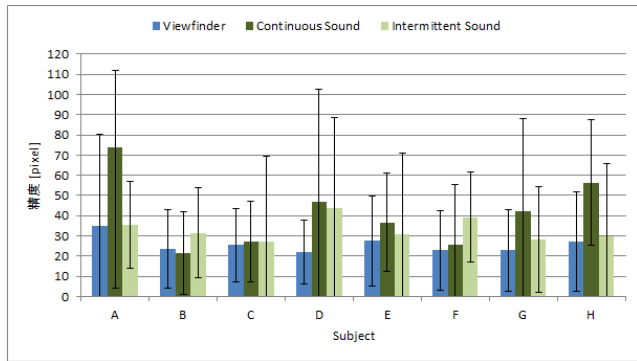


図 6.4 停止状態の水平成分の平均精度

Figure 6.4 Average precision of horizontal component in stand

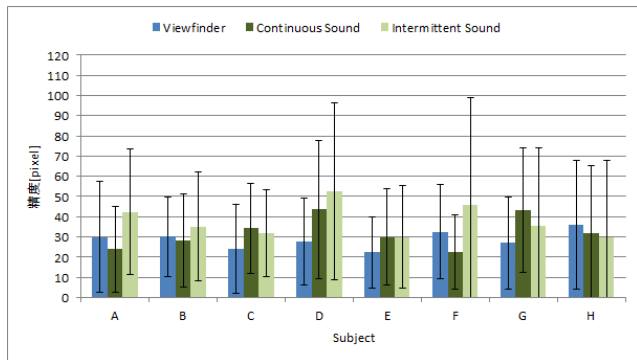


図 6.5 走行状態の水平成分の平均精度

Figure 6.5 Average precision of horizontal component in run

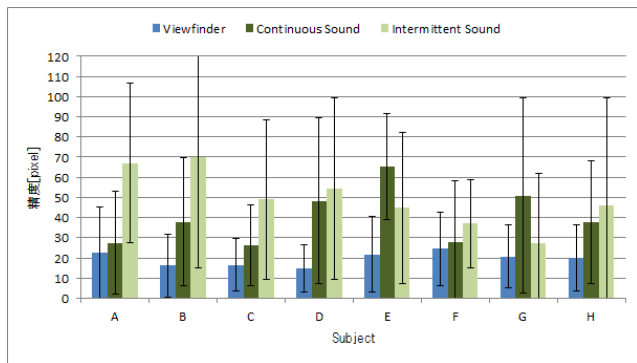


図 6.5 停止状態の垂直成分の平均精度

Figure 6.5 Average precision of vertical component in stand

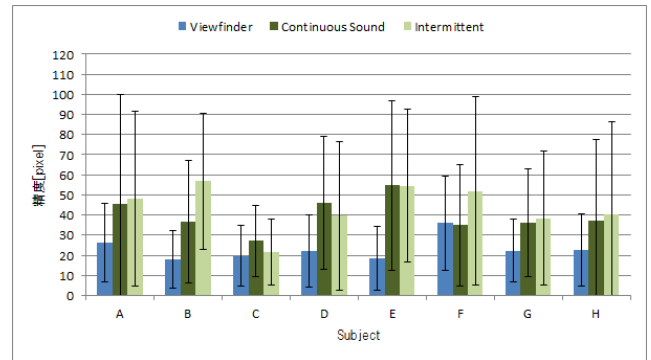


図 6.6 走行状態の垂直成分の平均精度

Figure 6.6 Average precision of vertical component in run

表 6.1 は各々の状態で、最も精度の高かったインターフェースの数を示している。予備実験と同じく、停止状態では、カメラファインダの精度が著しく高い。また、走行状態の垂直成分もカメラファインダの精度が著しく高い。一方、走行状態の水平成分では、カメラファインダの精度が良い被験者が 4 名、音声フィードバックの精度が良い被験者が 4 名と結果が分かれた。

表 6.1 精度の高いインターフェース

Table 6.1 Most Precision Interfaces

状態	成分	ファインダ	Continuous Sound	Intermittent Sound
停止	水平	7	1	0
	垂直	8	0	0
走行	水平	4	3	1
	垂直	7	1	0

6.5 分析

6.5.1 精度の個人差

全ての条件下の被験者間の指標データについて分散分析を行い、ターゲット追跡の精度に個人差があることを確認した。帰無仮説「被験者に関わらずターゲット追跡の精度に有意性はない」とした場合、有意水準 5% とすると有意確率 p は 5% 以下で帰無仮説を棄却できた。被写体のトラッキングの精度には、個人差があると言える。

6.5.2 走行状態の水平成分の音声フィードバック

音声フィードバックで平均の精度が高かった被験者が最も多かったのは、走行状態の水平成分である。この精度が、カメラファインダの精度より有意なのか t 検定を用いて調査した。表 6.2 は、走行状態の水平成分について、精度が最も良かったインターフェースとその他のインターフェースの間で有意差があった数を示している。カメラファインダの平均の精度が最も高かった 4 名のうち 1 名は、

Continuous Sound と有意差がない。よって、この1名についてはカメラファインダと Continuous Sound の精度が同レベルであったことを示している。

表 6.2 走行状態の水平成分の精度が最も高かったインターフェースとその他のインターフェース間での有意差

Table 6.2 Significant difference

平均精度の高いインターフェース	T 検定の対照		
	ファインダ	Continuous Sound	Intermittent Sound
ファインダ (計 8)		3/4	3/4
Continuous Sound (計 3)	3/3		3/3
Intermittent Sound (計 1)	1/1	0/1	

6.5.3 垂直成分の音声フィードバック

走行状態の垂直成分では、一人の被験者が Continuous Sound の精度の平均が最も高かったが、その他の被験者はカメラファインダの精度が最も高かった。走行状態の垂直成分は、水平成分とは異なり、音声フィードバックの精度が低い。

6.5.4 心理アンケート

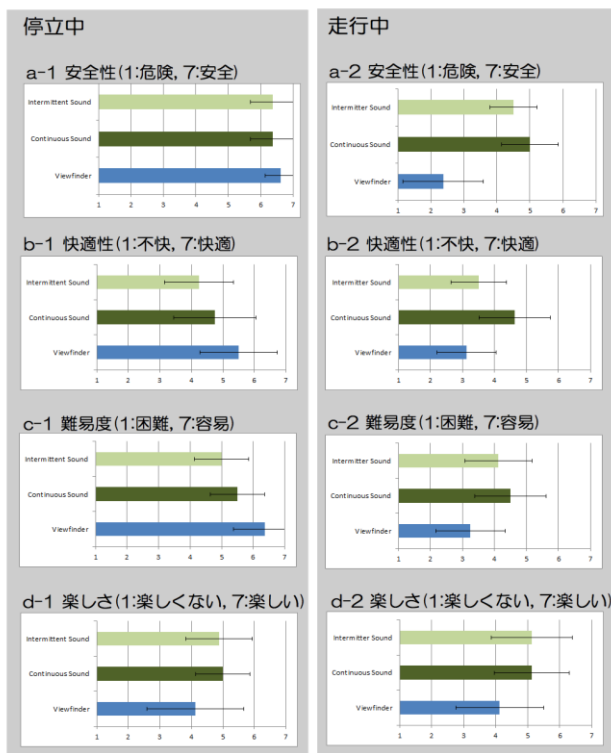


図 6.7 心理アンケート結果
 Figure 6.7 Result of Questionnaire

t 検定で有意水準 5% とし、各々のインターフェースに有意差があるか確かめた。

安全性

停止状態では、どのインターフェース間でも有意差はない。走行状態では、Viewfinder より Continuous Sound, Intermittent Sound のほうが安全で、有意差がある。Continuous Sound と Intermittent Sound 間には有意差はない。

快適性

停止状態では、全てのインターフェース間で有意差はない。走行状態では、Viewfinder より Continuous Sound, Intermittent Sound のほうが快適で、有意差がある。Viewfinder と Intermittent Sound 間には有意差はない。

難易度

停止状態では、全ての組み合わせで有意差はなかった。走行状態では、Viewfinder より Continuous Sound が簡単で、有意差がある。Viewfinder と Intermittent Sound には有意差がない。

楽しさ

停止状態、走行状態の両方で、全てのインターフェース間で有意差はない。

6.6 考察

6.6.1 停止状態でのカメラファインダ

停止状態では、カメラファインダが最も精度が高い。

6.6.2 走行状態のカメラファインダを使う際の心理状態

心理アンケートの結果上、走りながらの条件では、カメラファインダより音声フィードバックのほうが難易度、快適感、安全性で優れていた。特に安全性の結果では、両音声フィードバックでカメラファインダより有意差があった。

なぜカメラファインダを見ながら走ることで危険を感じるのか。その理由のひとつに、ファインダ内と走っている場所の座標系が異なっている点があげられると考えている。ファインダ内の撮像ターゲットの位置は、相対座標（カメラの位置とターゲットの位置からなる相対座標）という意識と、足を制御している絶対座標の意識との間で混乱が生じている可能性がある。また、走行によって、手元が震えることで、ファインダ内の撮像ターゲットの位置に影響を及ぼすという非常に複雑な制御を脳に負担していると考えられる。

問題提起として取り上げた通り、走行しながらのファインダを使用した撮影は、精神面で難しく、ストレスを感じやすく、危険な印象を与えるようだ。

6.6.3 走行状態の水平成分の音声フィードバック

走行状態では、水平成分で半数の被験者は音声フィードバックに優れた精度があった。一方、カメラファインダの精度が優れていた被験者もある。

なぜ被験者ごとに精度の優劣が異なるのか。カメラファ

インダの精度が最もよかった被験者のコメントに「左右の音量の違いが聞き取りにくかった」というものがあった。これは現在のシステムが1つの固定したHRTFを使用しているためだと考えられる。3D オーディオで正確に仮想音源を認識するためには、ユーザごとのHRTFを算出する必要がある。しかし、現在のシステムではユーザごとのHRTFを算出せず、固定の関数を使用しているため、人によっては仮想音源の位置を認知しにくい可能性がある。HRTFを算出し、その関数を使用すれば、水平方向のトラッキング精度が改善できる可能性がある。

6.6.4 垂直成分の音声フィードバック

垂直成分の音声フィードバックはほぼ全ての被験者でカメラファインダより優れた効果は出なかった。「低い音、高い音の音程が区別しにくい」や「低い音のときに上下どちらにターゲットがあるのか考えてしまう」という被験者のコメントがあった。このコメントから、低い音、高い音の"認知"の時間とさらに低い音はターゲット下/高い音はターゲット上という"判断"の時間が水平方向に比べてかかっていると推測できる。認知の時間を少なくするため、上下で音色を違うものにするなど、はっきりとした区別をする必要があるのかもしれない。また判断の時間を短くするのに、上下を連想できる音(例えばTVゲームのキャラクターがジャンプするような効果音なら上、土管に入るような効果音なら下など)などを試してみたい。

垂直方向では3D オーディオの採用をシステムの検討段階で断念したが、ユーザごとのHRTFを算出できれば、仮想音源の認知が正確になり、水平成分で得られたような精度を垂直成分でも得られる可能性がある。

6.6.5 2種類の音声フィードバック

撮像ターゲットを捉える精度は、Continuous Soundのほうが優れた効果があった。しかし、被験者によってはIntermittent Soundのほうが効果的だった被験者もいた。Continuous Soundのほうが好印象だった被験者の意見では、「Intermittent Soundはターゲットが移動したときだけ音が出るため、ちゃんと追跡できている不安だが、Continuous Soundは常に位置を教えてくれるので、安心」、 「Intermittent Soundは音が常に出ていないため前の音程を忘れてしまうが、Continuous Soundは常に前の音程の差が認識しやすい」などの意見があった。Intermittent Soundのほうが好印象だった被験者の意見では、「Continuous Soundは常に音が鳴っているため、少しの音の差では物体が移動したか明確に分からないが、Intermittent Soundは明確に分かる」、「Continuous Soundは常に音が鳴っているため、精神的ストレスが溜るが、Intermittent Soundは変化があったときだけ鳴るのでストレス感が少ない」などの意見があった。Continuous Soundは、常にターゲットの変化が読み取れるメリットがあるが、ターゲットが微妙に移動している場合、その変化を認識できるかはユーザの能力に依存する可

能性がある。Intermittent Soundは、ターゲットが移動したタイミングが明確にわかるが、前の音程との差分が分かりにくい。

改善策として、両方の音声デザインを組み合わせる。例えば、常にContinuous Soundが再生されているが、ターゲットが移動したら、Intermittent Soundが割り込んで再生するようにすれば両者の長所を活かすことができるかもしれない。また、ユーザの好みで音声デザインを事前を選択可能にするなどの改善策も考えられる。

6.7 まとめ

本実験では、走行中の水平成分の音声フィードバックがカメラファインダよりも精度が高い場合があることを明らかにした。しかし、精度が高いユーザには個人差があり、全てのユーザにとって、高い精度で扱えるインターフェースではない。また、垂直成分の音声フィードバックは、カメラファインダよりも精度が低いことが判明している。これらの課題を解決するため以下の3点の改善策を今後試す必要がある。

1. 3D オーディオのHRTFをユーザごとに取得する
2. 垂直成分を上、下を連想しやすい音色にする
3. Continuous SoundとIntermittent Soundを組み合わせせた手法を試す

7. 追加実験

7.1 概要

評価実験の結果から、垂直成分の音声フィードバックの精度が低いことが判明した。この問題を解決するため、改善策のひとつ、垂直成分を上、下を連想しやすい音色にした場合、判断時間が短くなるのか検証する。

7.2 評価設計

被験者3名に上下を意味する音ランダムで聞いてもらい、上が出力された場合は、上キーを、下が出力された場合は、下キーをできるだけ早く押しってもらう。音の出力からキー押下までの時間を反応速度とし、反応速度がより速い音が、判断時間のかからない音だと評価する。音は4種類用意した。

音色 A

サイン波の高周波数と低周波数で、高周波が上を低周波が下を意味する。

音色 B

上下を連想しやすい音色で、スーパーマリオ[13]のジャンプの効果音と土管に入るときの効果音を採用した。採用した理由は、被験者全員がプレイしたことのあるゲームだったため、経験的に上下を連想しやすいと推測したためである。

音色 C

音色Bの再生時間0.2秒以降の振幅をなくした音。再生時間が短くなった場合の反応速度を検証するためにこの音を

用意した.

音色 D

音色 B の再生時間 0.085 秒以降の振幅をなくした音. 音色 C と同様に, 再生時間が短くなった場合の反応速度を検証するためにこの音を用意した.

音色 E

音色 D を Wah 効果[14]で音波形を変化させたもの. 上下を連想できる音色からエフェクトをかけ, 音色を変えた場合に, 反応速度がどう変わるか検証するために用意した.

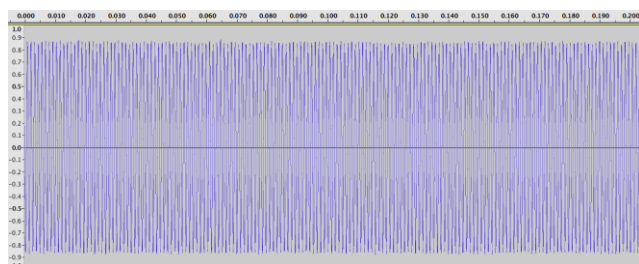


図 7.1 サイン波 高周波数

Figure 7.1 Sine wave high frequency

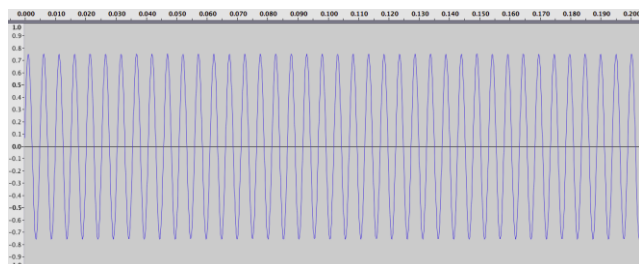


図 7.2 サイン波 低周波数

Figure 7.2 Sine wave low frequency

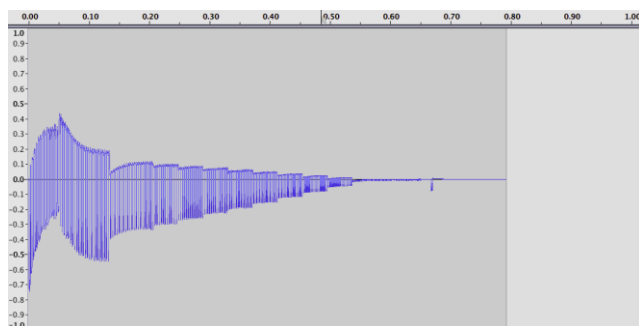


図 7.3 マリオジャンプ効果音

Figure 7.3 Sound of Mario's jump

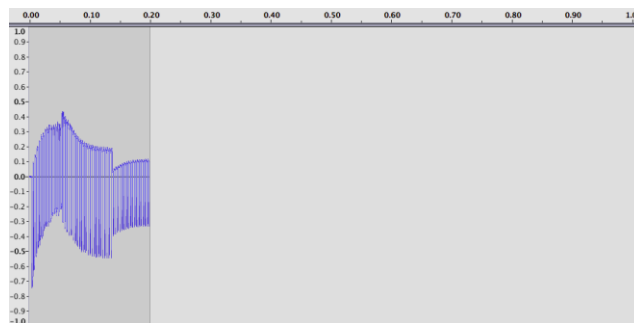


図 7.4 マリオジャンプ効果音(0.2 秒)

Figure 7.4 Sound of Mario's jump(0.2 second)

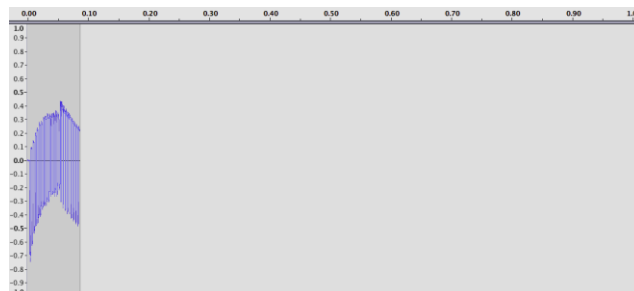


図 7.5 マリオジャンプ効果音(0.085 秒)

Figure 7.5 Sound of Mario's jump(0.085 second)

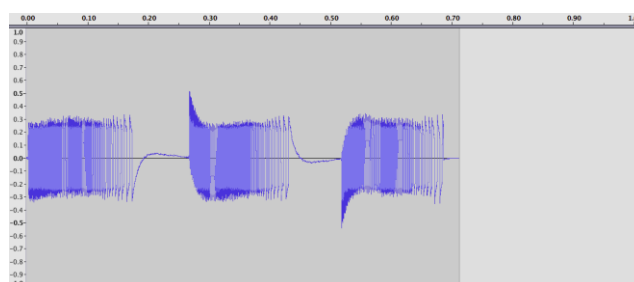


図 7.6 マリオ土管降下 効果音

Figure 7.6 Sound of Mario's dive

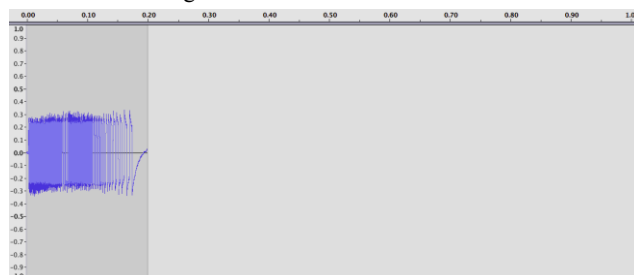


図 7.7 マリオ土管降下 効果音(0.2 秒)

Figure 7.7 Sound of Mario's dive(0.2 second)

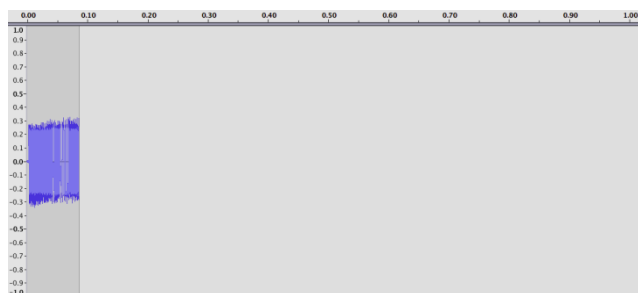


図 7.8 マリオ土管降下 効果音(0.085 秒)
 Figure 7.8 Sound of Mario's dive(0.085 second)

7.3 結果

図 7 は各音色の平均の反応速度を示している。サイン波よりも、上下の連想できる音色のほうが、反応速度が速い。また、再生時間が短いほうが、反応速度が速い。Wah 効果でエフェクトをかけた音色もエフェクトをかける前と比べて反応速度に変化はみられない。

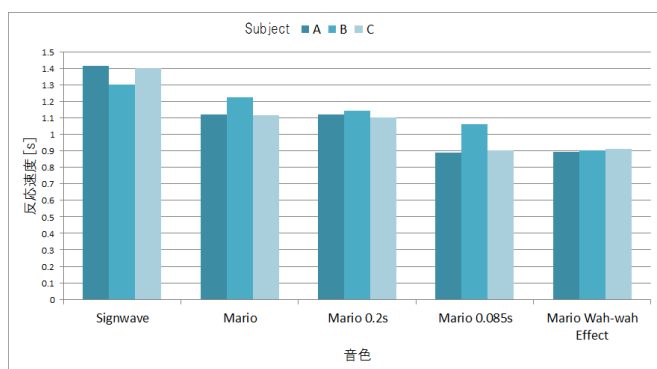


図 7 反応速度
 Figure 7 Reaction Rate

7.4 考察

予想通り、周波数の変化より上下を連想できる音色の変化のほうが、上下の判断にかかる時間が短いことが判明した。また、上下を連想できる音色にエフェクトを加えても、反応速度に影響を与えないことも分かった。今後、垂直成分の音声インターフェースに上下の判断時間が短いものを採用し、動体追跡の精度を高めていきたい。

8. おわりに

本研究では、音声フィードバックを用いて、動体追跡して撮影するための支援システムを提案した。

評価実験では、水平成分の音声フィードバックは、カメラファインダよりも高い精度で、動く被写体を捉えることができる場合があることが判明した。但し、垂直成分の音声フィードバックは、カメラファインダよりも精度が劣ることが分かった。

垂直成分の音声フィードバックの改善を検討するため、追加実験で、様々な音色の反応速度について実験を行った。

謝辞 MS-Word のテンプレートファイルの作成にご協力頂いた皆様に、謹んで感謝の意を表する。

参考文献

- 1) GoPro <http://gopro.com>
- 2) CONTOUR <http://contour.com>
- 3) Yasuhiro Kajiwara, Keisuke Tajimi, Keiji Uemura, Nobuchika Sakata, and Shogo Nishida. (2011). ClippingLight: A method for easy snapshots with projection viewfinder and tilt-based zoom control. AH '11. Article No. 14.
- 4) Gonzalez-Mora J.L. (2006). Seeingtheworldbyhearing: Virtual Acoustic Space (VAS) a new space perception system for blind people. Information and Communication Technologies, 2006. ICTTA '06. 2nd. 1. 837-42.
- 5) Gary R. Bradski. (1998). Computer Vision Face Tracking For Use in a Perceptual User Interface.
- 6) 任天堂 Wii Remote http://www.nintendo.co.jp/wii/features/wii_remote.html
- 7) SONY HDR-CX7 <http://www.sony.jp/ServiceArea/impdf/pdf/3214660011.pdf>
- 8) OpenCV <http://opencv.org>
- 9) Pure Data <http://puredata.info>
- 10) Earplug~ <http://puredata.info/downloads/earplug>
- 11) OpenSound Control <http://opensoundcontrol.org>
- 12) スーパーマリオ <http://www.nintendo.co.jp/ngc/sms/history/sml/index.html>
- 13) Wah 効果 [http://en.wikipedia.org/wiki/Wah-wah_\(music\)](http://en.wikipedia.org/wiki/Wah-wah_(music))