

音声フィードバックを用いたカメラ撮影ための動体追跡支援システム

瀬古 圭一 福地 健太郎*

概要. 音声フィードバックによってビデオカメラ撮影の動体追跡を支援する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、ビューファインダで被写体の位置を確認しながら撮影する。ビューファインダを確認しながら撮影するのは、撮影者が静止している場合、容易だが、撮影者が移動している場合は、困難であるし、ビューファインダ以外に注意がいき届かなくなり、障害物に当たる恐れがある。我々は、ビュー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させることで、より簡単かつ安全に動体追跡を行えると考えた。本稿では、提案システム、実装したプロトタイプの詳細と今後の展開について述べる。

1 はじめに

ビデオカメラで撮影する際、思い描いた構図の通りに撮影するには、ビューファインダで被写体の位置を確認する必要がある。しかし、動く物体を追跡しながら撮影する場合、ビューファインダに注視している状態では、ビューファインダ以外に注意がいき届かなくなり、障害物に当たる恐れがある。

最近では、GoPro[1]、CONTOUR[2]などのビューファインダを搭載しないウェアラブルカメラが市場に出回っている。これらのカメラは、軽量で撮影者の頭部や腕など様々な箇所に取り付けられ、広角レンズを装備しており、被写体を撮影画面内に収めるのが容易である。しかし、被写体の位置情報のフィードバックがないため、三人称視点映像を撮るには向いていない。

我々はこれらの問題を解決するため、被写体の位置を音声によるフィードバックによって認知させるシステムを提案する。音声フィードバックでは、撮影者の視線を自由に動かすことが可能なので、安全に動くものを追跡し撮影できる。被写体の移動量によって音声フィードバックを変化させるのだが、移動量の水平方向成分と垂直方向成分でそれぞれ生成手法が異なっている。水平方向成分は、立体音響による仮想的な音源位置が変化し、垂直方向成分は、周波数が変化する。また、被写体を動体追跡するため WiiRemote[5] の赤外線のリモコン機能を用いた。

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 システム

我々の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。音声の生成は、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。

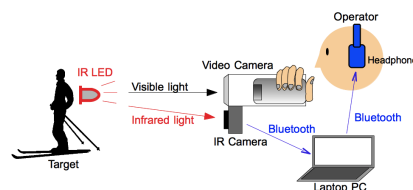


図 1. プロトタイプのハードウェア構成

4 プロトタイプ

図1は実装したプロトタイプのハードウェア構成を示している。前章で述べたように本システムは、トラッキング部分と音声フィードバック部分の2つモジュールから成る。両モジュールの実装について述べる。

Copyright is held by the author(s).

* Keiichi Seko, Kentaro Fukuchi, 明治大学大学院 理工学研究科 新領域創造専攻



図 2. 左：被写体に装着した IR マーカー 右：プロトタイプの実装

4.1 トラッキングモジュール

評価実験の際に、音声フィードバックの効果を正確に測るため、トラッキングの検出ミスをできる限り軽減する目的で、堅牢にトラッキングできる赤外線赤外線カメラで検知する手法を採用した。図 2 の左のように、被写体にバッテリー駆動の赤外線 LED を装着した。赤外線カメラには、赤外線の位置情報を出力してくれる WiiRemote を用い、トラッキングの実装を簡略化した。図 2 の右は、WiiRemote をカメラに接続した様子である。

4.2 音声フィードバックモジュール

WiiRemote から Bluetooth を介して赤外線 LED の位置データをラップトップ PC で受信する。ラップトップ PC では、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。ユーザは予めカメラ画面内の被写体の理想的な位置を決め、その範囲を「Sweet Spot」とする。「Sweet Spot」は図 3 のようにカメラ画面内の平面上で選択できる。音声フィードバックは「Sweet Spot」の中心からの移動量によって変化する。被写体が水平方向成分に移動した場合は、頭部伝達関数を基に実装された立体音響を用い、垂直成分に移動した場合は、音の周波数変化を用いる。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、まるで被写体から音が聞こえてくるかのような表現をする。但し、厳密に被写体から音が出るように表現するわけではない。立体音響に用いる水平方向の角度 θ は「Sweet Spot」の範囲を $|x| \leq p, |y| \leq q$ としたとき、式 (1) のように算出する。また、周波数は MIDI のノートナンバー単位で変化し、ノートナンバー n 、 N_c は 60、 N_m は 30 として、式 (2) のように変化する。

$$\theta = \begin{cases} \frac{\pi}{2p}x & (|x| \leq p) \\ \frac{\pi}{2} & (x > p) \\ -\frac{\pi}{2} & (x < -p) \end{cases} \quad (1)$$

$$n = \begin{cases} N_c + \frac{N_m}{q}y & (|y| \leq q) \\ N_c + N_m & (y > q) \\ N_c - N_m & (y < -q) \end{cases} \quad (2)$$

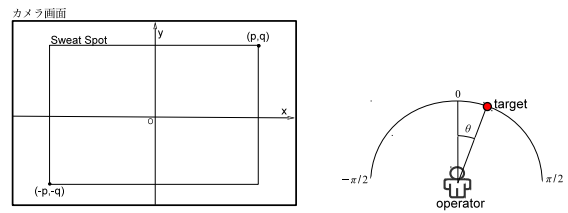


図 3. 左：カメラ画面内の座標平面と「Sweet Spot」の範囲 右：水平方向の立体音響の有効範囲

5 今後の展開

5.1 評価実験

ユーザが静止した状態と動いている状態で、動いている被写体を撮影し、撮影画像から「Sweet Spot」の中心からの距離を検証する。野外でプロトタイプを操作した結果、外乱によるトラッキングのミスが非常に多かった。赤外線を検知する上で、外乱が少ない環境、例えば室内で、ユーザが走行している状況を仮想的に作り出す予定である。具体的な案としては、運転のシミュレーションをしながらの撮影を検討している。

5.2 システムの改善

本システムの実用性を向上するために、トラッキングモジュールの改善が不可欠と考えている。なぜなら、現状のプロトタイプシステムでは、被写体に IR マーカーを事前に装着しなければいけないため、装着可能や被写体に限られてしまうし、撮影画像に IR マーカーが写ってしまう。現状のトラッキングモジュールに代わるものとして、深度計測可能なセンサを用い、物体の位置を検知する方法を検討中である。また、被写体の位置以外に必要な情報、例えば、カメラの傾きを提示するシステムなども検討中である。

参考文献

- [1] GoPro. <http://gopro.com>.
- [2] CONTOUR. <http://contour.com>.
- [3] Y. Kajiware, K. Tajimi, K. Uemura, N. Sakata, and S. Nishida. Clipping- Light: A method for easy snapshots with projection view- finder and tilt-based zoom control. In *AH '11*, Article No. 14.
- [4] Gonzalez-Mora J. L. Seeing the world by hearing: Virtual Acoustic Space (VAS) a new space perception system for blind people. Information and Communication Technologies, 2006. ICTTA '06. 2nd. 1. 837-42.
- [5] Lee J. C. Hacking the Nintendo Wii Remote. *Pervasive Computing*. IEEE, 2008. 7, 3, 39-45