

多人数テレビ会議システムにおけるキャリブレーションの自動化と映像表示の工夫

中野大輔† 富野剛† 福井登志也† 市村哲† 松下温†

東京工科大学†

1. 背景

会議における経費・移動時間の削減・ヒューマンリソースの偏りを防ぐためにテレビ会議が利用されることが増えている。しかし、従来のテレビ会議システムには、ハードウェアのコストが非常に高い、顔映像が小さく話者が分かりにくい、また会議の雰囲気をつまみにくいといった問題点があった。

2. 昨年度の研究成果

著者らは、昨年度の研究で2地点間、10人程度の会議室を想定し、画像処理と市販されているPC・カメラ・マイクを用いることでランニングコストを抑え、多人数が参加できるシステムを構築した。本システムの全体の流れを図1に示す。

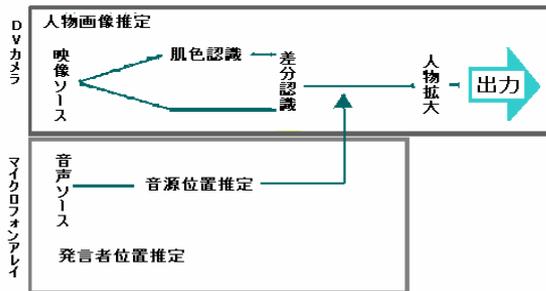


図1 システムの構成



図2 TV会議システム映像

本システムは二台のDVカメラの入力映像と、4本のマイクロフォン（マイクロフォンアレイ）から得られた話者位置推定データを統合して、会議参加者および話者の検出、拡大をし、上部に発言者を（三人まで）拡大したものを、下部に会議室全体を映したものを会議用映像としてもう一方の会議地点のディスプレイに出力するようになっている。（図2）

本システムにより以下の成果が得られた。

- ・ マイクロフォンアレイにより音声の到来方向が推定できることを実証した。
- ・ デジタルフィルタにより音声の到来方向の推定精度を向上させた。
- ・ 参加者10人、360°の話者位置推定実験において、9割～9割5分の高い精度で話者を検出できた。
- ・ 動体検出との組み合わせにより、発言以外の音の誤認識を減少させた。
- ・ ピーク周波数の判定により、紙めくり音の誤認識を減少させた。

3. 昨年度の課題

キャリブレーションとはマイクアレイを中心とした角度をDVカメラの映像上の角度と一致させることを指す。マイクロフォンアレイを中心とした角度とDVカメラからの角度は非線形に対応しているため、キャリブレーションが必要となる。

昨年度までのシステムではDVカメラの全体映像を目視で0°から360°、30°毎にクリックすることでキャリブレーションを行っていた。しかしこの方法では、正確なキャリブレーションは極めて難しく、マイクロフォンアレイを中心とした所定の角度方向に立って目印になると、キャリブレーションのためのクリックをする人が必要になり、時間と手間がかかってしまうという問題が挙げられた。キャリブレーションはシステムの精度に大きく影響を与える作業であるために、今回改善策を施した。

Automation of calibration in large number of people video conference and device of image display

†DaisukeNakano, TominoTakeshi, ToshiyaFukui, SatoshiIchimura, YutakaMatsushita

Tokyo University of Technology†

4. 提案

本研究では全方位カメラ（図 3）を用いることでキャリブレーションの自動化を行った。

マイクロフォンアレーの中心と全方位カメラの中心を合わせて配置して、全方位カメラ映像からマーカーとなる物体を検出し（図 4(a)）、その x 座標の値から、0 度から 360 度までの 1 度刻みの角度で位置を算出する。それと同期して DV カメラ映像から同一のマーカーを探す（図 4(b)）。同一のマーカーが認識されたら、全方位カメラで算出した角度と、同時刻の DV カメラ上のマーカーの角度を対応付ける。マーカーについては会議の開始前に、0 度から 360 度まで机の周りを赤色のマーカーを持って一周するだけで、キャリブレーションを自動的に行うことが出来るようにした。



図 3 全方位カメラ

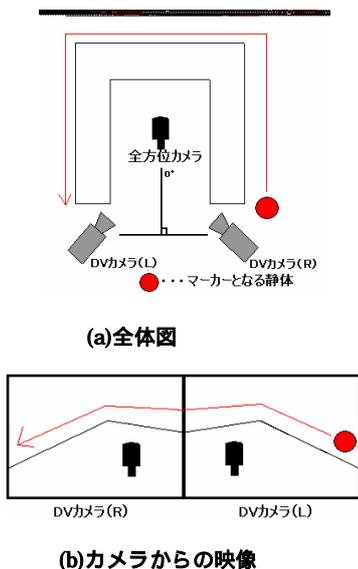


図 4 キャリブレーションの自動化

3. 実装

A) マーカー選択

マーカーとしては会議室であまり用いられない色である必要がある。

ここで、仕事の能率を向上させる上で色は大切な要素の一つとされている。淡い明るい色がオフィスには好まれ、仕事の能率を上げるには心地よい環境が不可欠である。多数の会議室を観察したところ、実際に白やグレーが基調とされており、人の感情を落ち着かせる青などが多かった。一方、人を興奮させる色とされる赤色はほとんど無く、マーカーとして利用するのに適している色であると考えられた。

マーカーが高精度で認識される大きさは、全方位カメラからの距離に応じて表 1 のような結果が得られている。なお本研究では、全方位

カメラからの距離が 170[cm]の時を想定して研究を進めていった。

表 1 マーカーの大きさと距離の関係

マーカーの大きさ [cm]	12	22	30	35	48
	×	×	×	×	×
全方位カメラからの距離 [cm]	45	90	110	170	220

B) 色認識

色認識には HSV 表色系を用いて H（色相）の値が 20 以下かつ 330 以上で、S（彩度）の値が 0.7 以上であるものを抽出した。しかし、全方位カメラでは、外側にあるガラス面に光の反射などによるノイズが入ってしまい、ゴマ塩ノイズと呼ばれるノイズを認識してしまい、正確にマーカーを認識することが出来なかった。

C) ノイズ除去

そこで 2 値化した後に、メディアンフィルタを用いることで細かいノイズの除去を図った。しかし、余り効果を得ることができなかった。これは全方位カメラのノイズは細かいもののみではなく、表面のガラスに反射した赤色が予想以上に大きく、除去しきれない赤色があったためだと考えられる。

そこで代わりに 2 値による収縮によって、ノイズの除去を試みた。収縮とは、ある画素の近傍に一つでも 0 があればその画素を 0 に、そのほかは 1 にする処理である。この処理をかけることで細かいノイズを除去することができた。そして、少し大きめのノイズに対しては、収縮処理を重ねてかけることによって少々広い範囲で赤色があっても除去することが可能となった。

D) 角度算出

全方位カメラの映像を横軸 720 画素に変換し、横軸の座標から角度を算出し、その角度を DV カメラの映像上のマーカーの座標に与えることでキャリブレーションを行った。

4. おわりに

本研究により、キャリブレーションを自動で行うことができ、キャリブレーションの簡易化・精度向上を実現することが出来た。

今後は会議中にもキャリブレーションを行うことができ、参加者数が変動してもリアルタイムで対応できるように利便性を向上させたい。

5. 参考文献

[1]色とチャンネル分解

<http://wingimp.hp.infoseek.co.jp/usegimp/ColorCompose.html>

[2]色と心理学 <http://www.fujipek.co.jp/color/sinri.html>

[3]視覚情報の統合による話者一検出システム—情報処理学会

第 59 回(平成 11 年後期)全国大会 特 1-57-62