

マルチモーダルセンサ情報を用いた ユビキタス情報場のコミュニケーション活性化

関原拓也[†] 杉山岳弘[‡] 阿部圭一[‡] 竹林洋一[‡]

[†]静岡大学大学院情報学研究科 [‡]静岡大学情報学部

1 はじめに

ユビキタスコンピューティングの研究が各国で盛んになってきた。筆者らは、人と人とのコミュニケーション活性化という観点から、ユビキタス情報場 (UBA) の構築を目指して、文工融合の研究開発をスタートした。広角カメラからの映像、ウェアラブル機器からの音声を含む複数のセンサ情報を、マルチモーダルセンサ情報として活用することで、ユビキタス情報場における人と人とのコミュニケーション解析を行い、社会学的研究を進める。インターネットの世界とは異なり、Face to Face であるユビキタス情報場のコミュニケーションは強力であり、顔の見えるその将来性は大きい。この環境下で、人が自由に動き、会話する状況の理解と構造化を行い、マルチモーダル知識コンテンツ [1] を獲得する。

2 ユビキタス情報場におけるコミュニケーション把握の重要性

2.1 ダイナミックなユビキタスコミュニケーション

ユビキタス環境である講義や会議といった場の、映像や音声情報をコンテンツにする研究はなされているが [2, 3]、講義や会議では、参加者は座った状態であったり、会議の進行役が決まっているなど、参加者全員が自由に動ける環境ではない。参加者が自由に移動できる環境だからこそ、人と人との繋がりが親密になり、そのような環境でしか現れない会話内容や、表現の仕方といったものも出てくる。ユビキタス・ウェアラブルコンピューティングシステムの進化とともに新しいコミュニケーション形態も生まれるため、この環境下でのコミュニケーション場の状況と、人の反応や会話の内容をコンテンツとして残していく価値は高い。この環境では、社会心理学研究も行う予定であり、また他にも、スポーツな

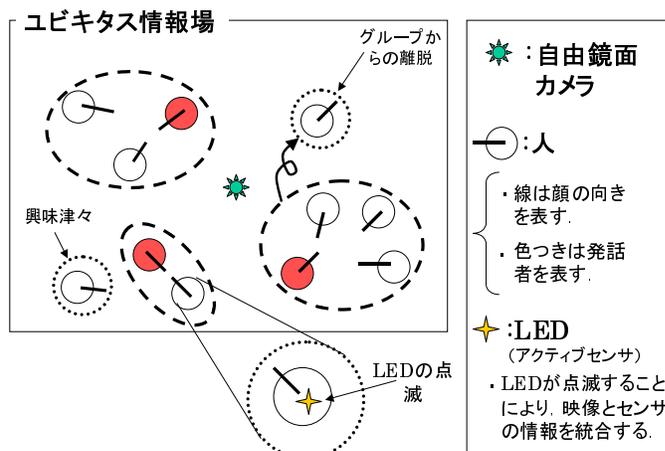


図 1: Face-To-Face Communication Unit 構成

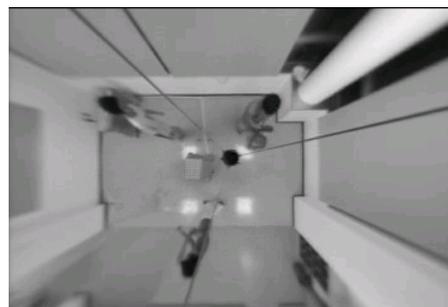


図 2: 自由鏡面カメラからの映像

ど様々な場で活用できると考えている [4]。

2.2 コミュニケーショングループ情報の重要性

人が自由に動けるユビキタス情報場において、すべての個人情報をそのままコンテンツにするのでは、その膨大な情報の中のどこに要点があるのか把握できない。会話を意味のある内容として残していくためには、個人という情報の単位を、コミュニケーションを行うグループという情報の単位として考える必要がある。ユビキタス情報場におけるコミュニケーションの形態は、双方向のやり取りであり、また誰もが主役になれる形態である。このようにダイナミックかつリッチなコミュニケーションのグルー

A Method for Activating Communication in Ubiquitous Space Using Multimodal Sensor Information
Takuya SEKIHARA[†], Takahiro SUGIYAMA[‡],
Keiichi ABE[‡], Yoichi TAKEBAYASHI[‡]

[†]Graduate School of Information, Shizuoka University

[‡]Faculty of Information, Shizuoka University

ブ形態を、Face-To-Face Communication Unit(以下 F2FCU)と呼ぶことにする。よって本研究では、ユビキタス情報場において、F2FCU を特定することを目的とする。F2FCU 情報は、その情報を客観的に参加者にフィードバックすることによって生じるコミュニケーションの変化、また、場面に応じて、参加者側から F2FCU 構成のコントロールをしていけるといった側面から、新たなコミュニケーションの形態が生まれ、コミュニケーションの活性化に繋がると思われる。

3 F2FCU の特定法

F2FCU 構成を図 1 に示す。本節では、ユビキタス情報場において、F2FCU を特定していく上での手法について述べる。F2FCU は、同じ話題を共有している集団であると定義し、まず、F2FCU を特定するために必要な情報について述べる。次に、その情報を得るために必要なマルチモーダルセンサについて述べる。

3.1 F2FCU 特定に必要な情報

人と人がグループを意識する時は、対話している場合が多い。また、対話以外にも、お互いの位置関係、顔の向きといった情報がグループ構成に影響を与える。以上のことを考慮した上で、F2FCU の特定を行うには、以下の項目の情報を得る必要があると考えた。

- 人物の位置
- 顔の向き (誰を見ているか)
- 発話開始時間、持続時間
- 個人の特定

これらの情報を得るには、様々なセンシングシステムの開発が必要である。

3.2 利用するマルチモーダルセンサ

ウェアラブル機器からの情報獲得と、コミュニケーション活性化のユビキタス情報環境デザインの一つから、コミュニケーション状況の理解に関する研究を進める。以下に、各センサの位置付けと働きをまとめる。

センサ マルチモーダルセンサヘッドセット [5] に地磁気ジャイロセンサを搭載し、このウェアラブル機器により、顔の向きを特定する。

映像 ヴィストン社製自由鏡面カメラによって、人物の位置を特定する。このカメラを設置用に改良を加え、部屋の中心の天井に届くように設置した。自由鏡面カメラの映像を図 2 に示す。このカメラは広角カメラでありながら、撮像面に、魚眼カメラのような円形歪みがない特徴がある。ユビキタス情報場において、人が自由に出入りできること、リアル

タイムに処理ができるということ、また、人物の追跡に失敗してもそれを検出できることを考慮し、頭の色情報を使用した。毎フレームごとに人物の頭の位置を算出し、それを前後のフレームで対応づける方法により、複数人物の追跡を行う。

音声 音声の内容を記録し、音声の内容理解を行っていたのでは、時間が掛かり過ぎ、システム全体に時間的ボトルネックを発生させてしまう。よって、ウェアラブル機器に搭載したマイクからは、発話の開始時間、持続時間といった、短時間で処理できる情報を得る。

アクティブセンサ ウェアラブル機器から得られるセンサ情報、音声情報は、局所的な個人の情報であるのに対し、映像から得られる情報は、人がどこに存在するかという空間的な情報である。この二つの情報の対応付けには、ウェアラブル機器のアクティブセンサを用いる。図 1 のように、ウェアラブル機器に搭載したアクティブセンサとしての LED が点滅し、それをカメラで捕らえることで、センサ情報と映像情報を統合する。これにより個人を特定する。

以上の情報をユビキタス場における個人の情報とし、この情報を F2FCU 特定に必要な情報とする。

4 まとめ

本稿では、ユビキタス情報場で行われる人と人とのコミュニケーション状況について、新しいモデルを提案した。人同士のコミュニケーションをマルチモーダル知識コンテンツとして獲得するには、F2FCU という単位の概念が必要であるという見通しを立てた。

今後、得られた知見を活かし、コミュニケーションの構造の詳細化と F2FCU の特定法についての検討を進めていく。

参考文献

- [1] 鈴木, 岐津, 宮澤, 浦田, 網, 竹林: “マルチモーダルナレッジをオンデマンドで配信する MKIDS システムの開発,” 人工知能学会全国大会, 2D1-03 (2002) .
- [2] D.Aboud,G.,Atkeson,C.,Feinstein,A.,Goolamabbas,Y.,Hmelo,C.,Register,S.,Sawhney,N.N.,Tani,M.: “Classroom 2000: Enhancing Classroom Interaction and Review,” GVVU Technical Report GIT-GVVU-96-21(1996).
- [3] 亀田, 石塚, 美濃: “状況理解に基づく遠隔講義のための実時間映像化手法,” 情報処理学会研究報告, CVIM121-11 (2000).
- [4] 坂根, 高島, 大谷, 竹林: “マルチモーダルセンシング技術を用いた格闘技解析に関する実験,” インタラクシオン 2003 (2003, 発表予定) .
- [5] 金澤, 友田, 高島, 竹林: “ユビキタス社会に向けた Bluetooth ヘッドセットの開発,” 日本音響学会 2002 年春季研究発表会, pp219-220 (2002).