

1Y-1

顔の輪郭画像を用いた会話システムにおける会話特性

吉田亨[†] 徳永幸生[†] 米村俊一[‡] 大谷淳[‡]

[†]芝浦工業大学 [‡]NTT サイバーソリューション研究所 [‡]早稲田大学

1. はじめに

コミュニケーションにおいて顔、とりわけその顔の表情は極めて重要な役割を果たしている。相手の表情から、性別、年齢などの個人情報や、更に喜怒哀楽などの心理的内面さえも把握することができるからである。人間の会話では、メッセージのわずか7%が言語で伝達される一方、55%は顔の表情によって伝達されるという報告[1]もある。

しかし、顔をリアルに表示するTV電話には抵抗感があるといった報告[2]や、TV電話を通じての対話は、利用者に違和感や居心地の悪さを与え、好ましくないという実験結果[3]も報告されている。

本稿では、顔映像を用いた映像通信の抵抗感軽減を目的として、顔の輪郭映像を用いたコミュニケーションシステムを構築し、伝送映像の情報量と会話特性の関係を実験的に検証する。今回は、会話中の話者交代、同時発話に注目し、コミュニケーション時に使用する映像による違いの計測、評価を行った。

2. 輪郭映像通信システムの概要

図1は、原画像の入力からサンプリングされたエッジ表現による輪郭映像の生成までの流れを示したものである。Webカメラから入力された映像を取り込み、1フレーム毎に取り出した原画像に対してエッジ抽出処理、2値化処理、細線化処理を順次施す。その後、画像をサブブロックに分割し、分割したサブブロックをランダムにサンプリングして、出力した映像が輪郭映像となる。

図2に輪郭映像通信システムの概要を示す。Webカメラを装備したPCが、ルータを介してネットワーク接続されている。Webカメラからの入力映像に対し、エッジ抽出、細線化、ランダムサンプリング等の画像処理を施し、相手側PCに映像を送信する。そして、映像の通信を行うことで、送信側PCには相手側PCから送信された映像が表示される。

画像の取り込み、輪郭映像の生成、映像の通信は、Microsoft社のDirectXを用いて構築した。システムはDirectXに付属するGraphEditのプレビューの機能を用い、輪郭映像の生成は、DirectShow上で動作するエッジ抽出フィルタを開発し、GraphEdit上で動作させた。音声はマイク内蔵のスピーカを双方向で直結することで、音声チャネルを確保した。

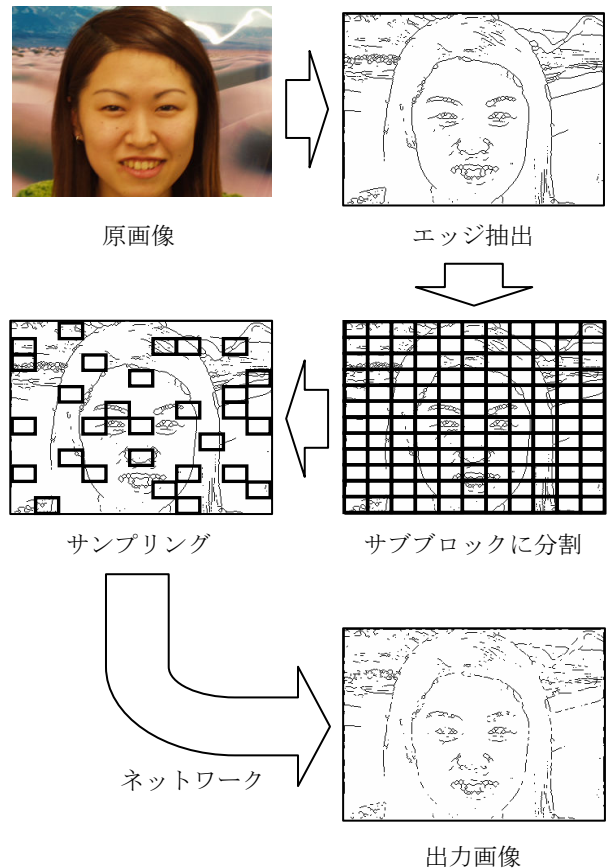


図1 輪郭映像生成処理の概要

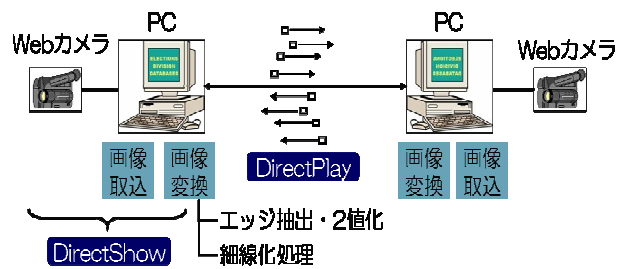


図2 実験システムの概要

3. 会話実験

会話時の人間の挙動は様々な要因によって変化するため、会話行為を観測することによってシステム評価を行う場合、それらの要因を統制する必要がある。本実験では、コミュニケーションスタイルを一意に統制するために、次のように会話状況を設定した。

- 話し相手：実験者と日常的に会話を交わす顔見知り
- 会話内容：相手に依頼する説得的内容
- 会話場面：非対面の遠隔通信

Characteristics of Visual Communication System Based on Edge Abstraction

[†]Tohru YOSHIDA (m106110@sic.shibaura-it.ac.jp)
[†]Yukio TOKUNAGA (tokunaga@sic.shibaura-it.ac.jp)
[‡]Shunichi YONEMURA (yonemura.syunichi@lab.ntt.co.jp)
[‡]Jun OHYA(ohya@waseda.jp)
[†]Shibaura Institute of Technology
[‡]Cyber Solutions Laboratories, NTT [‡]Waseda University

3.1 被験者

実験者と顔見知りの大学生 3 名を被験者として用いた。また、日常的に TV 電話を利用している被験者はおらず、輪郭映像を用いた映像通信システムの利用経験者もない。

3.2 使用する映像

会話実験は、リアル映像、輪郭映像、表情なし輪郭映像、映像なしの 4 条件下で行った。なお、輪郭映像のサンプリング率は 100%とした。図 3 に、使用映像の比較例を示す。

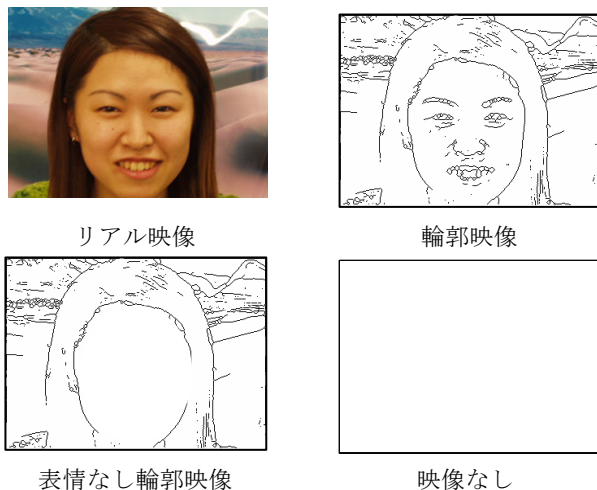


図 3 実験で使用する映像

3.3 実験手続き

被験者にタスク内容を記した紙を見せ、実験者と遠隔会話を行なわせた。タスク内容には現在自分が置かれている状況と、説得する内容が書かれている。実験を行う前には、システムに慣れてもらうため、2~3 分程度の会話練習を行った。

使用する映像の 4 条件とタスクの 4 条件は被験者ごとに異なるように組み合わせ、会話を行った。また、会話時の映像は、ビデオカメラで撮影し、実験後の解析で使った。

4. 実験結果

図 4 に実験者の平均話者交代潜時を示す。話者交代潜時は、実験者が話し終わってから被験者が話し始めるまでの時間間隔を指す。

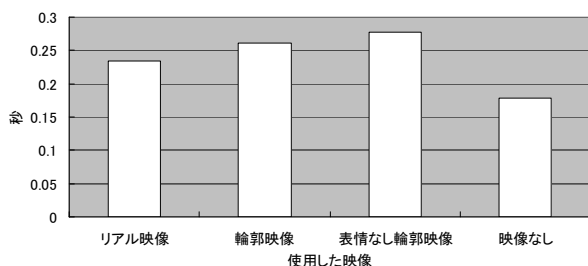


図 4 平均話者交代潜時

図 5 に平均同時発話回数を示す。縦軸は会話における同時発話回数を、1 分単位で平均化した回数となる。

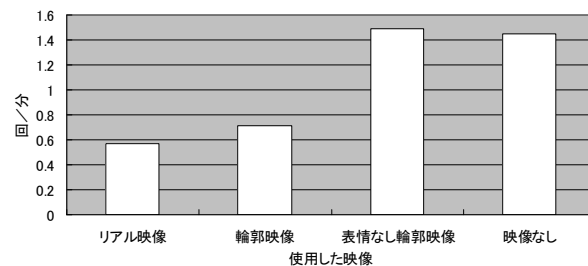


図 5 平均同時発話回数

5. 考察

今回の実験ではサンプル数が少なく、有意差検定を行っていないが、図 4 の平均話者交代潜時の結果を見ると、映像のある 3 条件間には大きな差がないことから、映像がある場合には、映像の情報量の大小が話者交代潜時に大きな影響を与えていない可能性が考えられる。また、映像なしの場合では、映像ありの場合よりも話者交代潜時の値が小さくなっている。これは映像なしの場合、相手の状況を把握する情報が音声以外にないため音声途絶えると不安になり、沈黙を避けようとする人間の心理が働くためだと考えられる。

また、図 5 の平均同時発話回数を見ると、リアル映像、輪郭映像使用時の評定値が、表情なし輪郭映像、映像なしの評定値を大きく下回っている。これは、表情が見えることで、見えない場合に比べ、相手の発話のタイミングを見誤る頻度が低くなり、同時発話が起り難くなっているためだと考えられる。

6. まとめ

本稿では、顔の輪郭映像を用いた通信システムを構築し、会話中の話者交代、同時発話に注目し、コミュニケーション時に使用する映像による違いの計測、評価を行った。その結果、相手映像がある場合には、伝送映像の情報量は話者交代潜時に大きな影響を与えていない可能性が、また、表情が確認できることは同時発話を減らし、コミュニケーションがスムーズになる可能性が示唆された。

しかし、本稿で行った実験は、ある特定の会話状況であり、今回使用した会話状況以外でも同様な結果が得られるのかも検討する必要がある。また、今後は表情がコミュニケーションにどのような影響を与えるかについても様々な視点から検討して行く予定である。

<参考文献>

- [1] W. フォン. ラフラー. エンゲル 編著, 本名 訳, “ノンバーバルコミュニケーション”, 大修館書店 (1981).
- [2] 犬童, 小磯, 下嶋, 岡田, 片桐, “映像を介したコミュニケーションの特性分析”, 信学技報 HCS96-42, pp22-34 (1997).
- [3] 原田, “人の視点からみた人工物研究”, 共立出版株式会社 (1997).