

遠隔調理支援におけるアドバイスとコミュニケーションの関係

小幡佳奈子[†] 中村 裕一[†] 小泉 敬寛[†] 近藤 一晃[†] 渡辺 靖彦^{††}

[†] 京都大学 学術情報メディアセンター, 〒 606-8501 京都市左京区 吉田本町

^{††} 龍谷大学 理工学部, 〒 520-2194 大津市瀬田大江町横谷 1 番 5

E-mail: †{obata,yuichi,kondo}@media.kyoto-u.ac.jp, ††watanabe@rins.ryukoku.ac.jp

あらまし 本稿では、遠隔調理支援におけるコミュニケーションの解析について紹介する。遠隔調理支援は料理教室に直接行くことができない人に調理の知識や技術を伝達することをはじめ、調理を在宅リハビリテーションの課題としたり、独居老人の社会的関わりを保ったりするための良い枠組みとなる。そのために、我々は、調理している人の一人称映像 (First Person Vision) を遠隔にいる教示者や家族に送り、それをういて遠隔から指示や支援を行う調理支援の形態を提案している。このような調理支援が円滑かつ安全に行われるためには、適切なアドバイスがタイミング良く与えられる必要がある。そのため、本研究では、与えられるアドバイスとコミュニケーションのパターンとの関係、特に、言動の「呼応パターン」と言動の「間」との関係について調査した。その結果、コミュニケーションのパターンの時間的特性とアドバイスの出現に密接な関係があることがわかった。

キーワード 遠隔調理支援, 映像対話型行動支援, インタラクション分析, アドバイス, マルチモーダルコミュニケーション, 呼応パターン, 間

Relationships between Advice and Communication Patterns in Remote Cooking Support

Kanako OBATA[†], Yuichi NAKAMURA[†], Takahiro KOIZUMI[†], Kazuaki KONDO[†], and Yasuhiko
WATANABE^{††}

[†] Academic Center for Computing and Media Studies, Kyoto University, Kyoto, 606-8501, Japan

^{††} Faculty of Science and Technology, Ryukoku University, Ohtsu, 520-2194, Japan

E-mail: †{obata,yuichi,kondo}@media.kyoto-u.ac.jp, ††watanabe@rins.ryukoku.ac.jp

Abstract In this study, we introduce a novel scheme for distance learning that aims to teach people how to cook, thereby reaching people who are unable to join online cooking classes, supporting rehabilitation and care for single-living elderly people, and handing down dietary knowledge and culture. Our scheme utilizes a framework of working support through first person vision (FPV) communication, in which a worker with a head-mounted camera works under the guidance of an experienced mentor monitoring the FPV from a distance. To make this scheme fully effective, we investigated the communication that takes place in this environment, *i.e.*, how typical response patterns and absence pauses induce advice. Through experiments, temporal characteristics of such features are found to be tightly related to the occurrences of advice.

Key words First Person Vision Communication, Remote Cooking Support, Interaction analysis, Advice, Multimodal communication, response pattern, absence pause

1. はじめに

近年、高齢化や核家族化が進んだことに伴い、家族の形態や地域との関わり方が大きく変化してきている。そのため、料理に関する知識や技術、伝統が家庭や地域で伝えられる機会が減っている。こういった背景を踏まえ、我々は遠隔からの調理支援とその有効性について検討している。

調理には次のようなメリットがある。現在の単身者は外食やインスタント・レトルト食品などに頼ることが多く、塩分や脂肪分の摂り過ぎや、野菜やミネラルの不足など、栄養が偏ってしまう問題を抱えている。単身者が自分で調理が出来るようになれば、栄養面に十分考慮した調理を行ったり、調理によるストレスの解消などの効果が期待できる。また、調理は認知症や高次脳機能障害のリハビリテーションの課題として有効であることが知られてきている。調理では物事を順序だてて行うこと、指先や手を使った作業に集中することが必要であるため、良い刺激と訓練になる。

しかし、単身者が日中仕事をしている場合、料理教室に行く時間が取れない、または、近くに料理教室が無いといった問題が考えられる。また、患者や高齢者が独り暮らしの場合には、調理中に家族や介護者が常時そばにいることが難しいという問題がある。調理は火や刃物、熱湯を扱うことが多いため、そのような状況では、事故が起こる可能性も無視できない。

遠隔調理支援はこのような問題を解決するための一つの方法となる。図1に一人称視点からの映像 (First Person Vision) を用いた遠隔調理支援の概要を示す。調理者は頭部にカメラを装着し、目の前の様子を遠隔地にいる支援者に伝送しながら調理を行う。遠隔地でモニタリングしている支援者はその映像を見ながら指示・指導を与える。ただし、支援者側からは音声のみを送る。調理者が初心者である場合や、集中力を保つことに困難がある場合を想定しているため、支援者側の映像に調理者が注意を奪われることを避けるためである。このような手法を用いて遠隔調理支援が実現されれば、単身者・独居老人、離れて住んでいる家族が料理を介したコミュニケーションを持つことができ、キッチンで一人寂しく奮闘しなけなければならないという問題を軽減できる。

この遠隔調理支援システムの利点は一つは、システムが簡便なものとなることである。つまり、まな板やコンロなど、複数ある調理場所をそれぞれ撮影するためにカメラを複数用意したり据え付けたりする必要がない。さらに重要な特徴として、「作業者がどこを見ているか・注目しているか」を支援者が容易に把握出来るという点があげられる。そのため、調理者の注意や興味に合わせたアドバイスを与えたり、忘れていたことに注意を促したりすることが可能になる。ただし、次のような問題もある。調理者と支援者が同じ場所にいる環境と比べ、使えるコミュニケーションチャンネルが制限されていること、支援者が調理者の体や食材、調

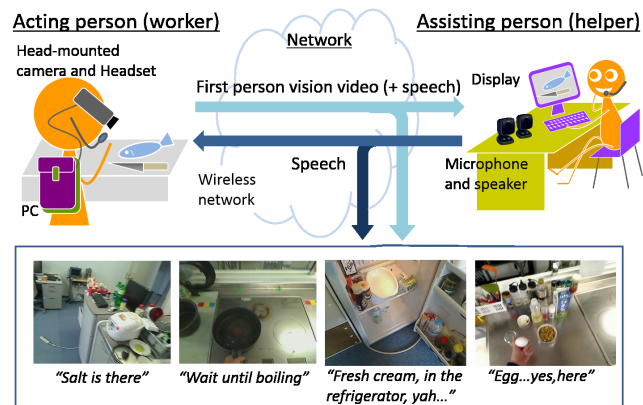


図1 一人称映像 (First Person Vision) を用いた遠隔調理支援システムの概要

理器具等を直接触ることが出来ないことなどである。そのため、必要な情報が的確に伝わるコミュニケーションが行われ、調理作業を安全かつ楽しい場とすることが重要である。

このような問題意識から、我々は、調理者・支援者間のコミュニケーションのパターンを調査し、コミュニケーションの状態によって支援者からのアドバイスの発生が受ける影響について調査した。以下、本稿では、まず、遠隔調理支援におけるコミュニケーションの解析方針について述べ、次に、アドバイスと特定のコミュニケーションパターンとの関係について吟味し、実際の調理支援を解析した結果について報告する。

2. 関連研究

ウェアラブルコンピュータによる映像通信とそれを用いた遠隔作業支援について、早くからその可能性が示唆されてきたが [1, 2], その詳細な解析として Kraut らが先駆的な研究を行っている [3]. 彼らは、自転車の修理を課題とした遠隔行動支援に対して、音声とマニュアルのみの場合とさらに映像を加えた場合の作業効率や作業者 (worker) と支援者 (expert) の振舞いについて比較した。映像対話が可能になることによって、両者の振舞いに大きな変化が現れること、特に支援者がプロアクティブに振る舞うことなどが報告されている。Fussel らは、更に、支援者が実空間で側についている場合を加え、映像通信がある場合と、音声だけの場合との比較を行った [4]. 彼らは視覚情報の活用の仕方が3つのコミュニケーション形態で異なり、それによって発話が大きく変化することなどを報告している。Kraut, Gergle らは、ジグソーパズルを題材としたオンラインの協力作業を設定し、作業者 (worker) と支援者 (helper) の振舞いを解析した [6, 7]. 映像の遅延、映像共有の有無によるコミュニケーションの差が、現場参照の他に、理解の確認、作業後の確認などに顕著に表れることなどを統計的に示している。



(a-1) (a-2) (a-3)

調理者 (a-1): “あれ?強力粉もあるんですか?” (袋を指さしながら)
 支援者 (a-2): “えっと、型に付かないために付ける方が強力粉で、普通に作る方が薄力粉”
 調理者 (a-3): “なるほど!はい”

図 2 コミュニケーションとアドバイスの例 (a)



(b-1-1) (b-2-1) (b-3-2)

調理者 (b-1-1): (塩の容器を振る)
 支援者 (b-1-2): “振るのじゃなく、回してゴリゴリ入れるみたいです”
 調理者 (b-1-3): “はい、わかりました。”
 調理者 (b-2-1): (両手に持った砂糖の袋を見ている)
 支援者 (b-2-2): “青い方がグラニュー糖で赤い袋の方が上白糖。今回は...グラニュー糖。”
 支援者 (b-3-1): “まず油をひいて下さい”
 調理者 (b-3-2): “はい”
 支援者 (b-3-3): “少し予熱してから...”
 調理者 (b-3-4): “はい”
 支援者 (b-3-5): “...キッチンペーパーとかで...薄く全体に延ばす感じで...”

図 3 コミュニケーションとアドバイスの例 (b)

本研究の目的は、それをさらに進め、映像対話型行動支援、特に、従来研究では行われていなかった、調理支援におけるコミュニケーションの「良さ」や「円滑さ」を評価したり、アドバイスとの関係を解析するための指標を作ることである。この点に関連し、Kraut, Gergle らの研究 [6, 7] では、高次の意味的な行動(記号)を用いてコミュニケーションのパターンを解析している。それに対し、本研究では、自動的な解析を視野に入れ、直接観測可能な(低レベルの)特徴を基にした定量的な解析に主眼を置いている点が特徴である。

3. コミュニケーションのパターン

我々の遠隔調理支援システムで、頻繁に発生した状況を図 2-4 に示す。それぞれ、調理者は自分が詳しく知らないメニューを支援者の指示のもとで調理している。これらの例から、発話と振舞いとアドバイスの関係を考えてみよう。

図 2 は調理が滞りなく進められている場合である。映像を通じて、必要な情報が過不足無く支援者に送られており、作業者が明



(c-1-2) (c-1-4) (c-2)

支援者 (c-1-1): “(ケーキの)真ん中に包丁を入れて水平にざっと切って...”
 調理者 (c-1-2): “むずっ!(相槌や反応をせず、独り言のように言う)”
 支援者 (c-1-3): “あはは”(笑ったため、どう切れば良いか、アドバイスを機会を逸した)
 調理者 (c-1-4): “どうや?あ、めっちゃ曲がってる!なんだこれ!”
 支援者 (c-2-1): “...僕は手で(バターを型に)塗ってました...”
 調理者 (c-2-2): “あー、面倒臭い!面倒臭い!オイルスプレーを買いましょうよ。スプレー”
 支援者 (c-2-3): “へえ、そんな便利な物があるんですかあ”(バターをどれ位使うか、などの説明や塗り方のアドバイスを機会を逃した)

図 4 コミュニケーションとアドバイスの例 (c)

示的にアドバイスを求めると、その必要性をわかっていた支援者が返答した。一般的に、このようなケースでは、顕著な問題は発生しにくい。

図 3 は、調理者の振舞いや行いを見ることによって、支援者がアドバイスの必要性に気付く、自発的にアドバイスを与えた場合である。一般的に、調理者は作業中に手元に集中する必要が多々あり、明示的に質問する余裕が無いことが多い。また、自分がアドバイスを必要としていることすら気付いていない場合もある。それに対し、(b-1) では、調理者が塩の入った容器の使い方を間違っていることに気付いた支援者が使い方の説明をしている。(b-2) では、調理者が手にしている砂糖のどちらを使うべきか分かっていないことに支援者が気付く、指示を出している。(b-3) は、上記 2 つとは異なるが、良いテンポのコミュニケーションが続いたために、支援者がアドバイスをタイミング良く与えられている。調理者が良いタイミングで相槌を打つことによって、意図や状況を把握しやすくなることが示唆されている。

図 4 は、支援者がアドバイスを与えられなかった例である。(c-1) では、調理者と支援者の間で会話の内容が一貫しなかったために、支援者はアドバイスをやるタイミングを逃した。(c-2) では、作業者が別の話題を持ち出したために、支援者がアドバイスの必要性に気付かなかった。

一般的には、(a)、(b) のような状態を保ち、(c) のような状況を減らすことが望ましい。この目的のため、我々はアドバイスが与えられる過程に注目した。支援者がアドバイスの必要性に気付く、的確なアドバイスを考え、調理者に伝える、という流れである。

このような過程で重要となるコミュニケーションのパターンと

して以下の二つがあげられる。

問： 支援者が調理者の言動や状況を把握したり、アドバイスを考えるためには一定の時間を必要とするため、行動や発話の行われていない「間」はその良い時間を与える。

応答： 対話の相手がタイミング良く的確な応答をすることが、状況を把握してアドバイスの必要性に気づくために必要である。また、アドバイスを与えるきっかけやタイミングを与えるためにも、タイミングの良い応答が重要である。

ただし「間」として「2つの言動の間の時間差」と「何も言動が行われていない」ことの2つを考えることができる。本研究では、前者を、上記の「応答」と合わせて考え、適切な応答が的確なタイミングで与えられていることと要件とする。後者に関しては、何らかの言動が行われていないことを要件とするが、全ての言動が止まることを必要せず、単純作業など、対応を要しない一部の言動が続いていても、その役割を果たすと考える。

4. コミュニケーションパターンの分析

4.1 間と呼応パターン

前節で述べた、間と応答のコミュニケーションパターンには、種々の発話や動作の組み合わせが考えられるが、我々の研究目的は、このようなパターンの性質とアドバイスの発生との関係を調査することである。しかし、複数の要素の組み合わせは膨大な数になり、その全てを調べあげるのは難しい上に、その大多数が重要な意味を持たない可能性が高い。そのため、以下のように、あらかじめコミュニケーションパターンを絞り込みながらアドバイスとの関係を調査した。

(1) 頻出パターンの抽出: 実際のコミュニケーションの中で頻出するパターンを抽出する。

(2) 呼応パターンの抽出: 選択した頻出パターンに含まれる、基本的な応答のパターンを抽出する。

(3) アドバイスと呼応パターン、間との共起性の評価: コミュニケーションパターンの時間的な特徴とアドバイスの出現との関係を相互情報量を用いて統計的に調べる。

頻出パターンを用いることは以下の理由から有効であると言える。

- 頻出するパターンはコミュニケーションの基本的な単位となっている可能性が高い。
- アドバイスが良好に与えられたサンプルが十分に揃えば、頻出パターンがアドバイスに密接に関係していることが期待できる。

逆に、出現頻度の少ないコミュニケーションパターンは、偶然にアドバイスと共起した可能性を否定できないなど、統計的に有意な情報を得ることが難しいことから、今回の検討対象からは除外した。

表1 コミュニケーション特徴

特徴	説明	シンボル
アドバイス (発話)	現在, 未来	A:n, A:f
他の発話	要求, 質問, 説明, 相槌	R, U, D, P
視線停留	カメラの動きが無し/小さい	V:h
見回し	カメラの回転	V:l
場所移動	カメラの平行移動	V:m
手の動き	手の動き	V:b

4.2 要素特徴

情報を伝えるコミュニケーション要素として、我々は以下の3つのモダリティを考慮した。

発話: アドバイス (現在もしくは未来の事柄に関する説明), 要求, 質問, 確認, 相槌など

動作: 作業, 手の停留, 場所移動, ジェスチャ(主に指差し) など
見る動作: 注視 (視線停留), 見回しなど

動作と見る動作はともに映像によって伝えられる調理者の行動であるが、前者は調理者の動きやその状況を伝える情報として、後者は調理者の注意や意図を表す情報として分類している。後者は作業者がどのような情報を必要としているかを示唆することになる。

表1は我々が今回の実験で用いた特徴である。データは自然言語処理と画像処理で検出することを想定している。

4.3 頻出パターンの抽出

本節では、頻出パターン検出の概略を説明する。詳しくは、[10,11]を参照されたい。時系列データのマイニングでは、アイテム (item) を時系列に並べたものをトランザクション (transaction) とし、そこから頻出パターンを検出する。本研究では、前節であげたコミュニケーション特徴 (F_i), 例えば、アドバイス、見回しのような特徴を以下のように時系列にしたがって並べたもの S をトランザクションとする。

$$S = \{F_1, \dots, (F_i, \dots), \dots, F_z\}$$

頻出パターンはこのトランザクションに頻出する部分列となる。抽出アルゴリズムには、時系列パターンのマイニングで良く用いられる PrefixSpan [12] を用いた。PrefixSpan では複数の系列データ入力とし、そこから抽出される数が閾値以上になる部分系列パターンを出力する。得られた頻出パターンの例を表2に示す。ここで用いられている記号は、":"の前の"W"が調理者の発話、"S"が支援者の発話、"V"が調理者の行動(画像より得られる)を意味し、":"の後の"P","D"等は表1にあげた特徴を表す。例えば、"(V:l)(W:D)(S:D)"は、調理者が周囲を見回した後、何らかの説明をし、その後、支援者が何らかの説明を表す。

表 2 頻出パターンの例

頻出パターン	平均共起個数	平均区間長
<(W:D) (S:D) (W:P)>	6.12	15.70
<(S:R) (W:D) (S:D)>	7.50	24.03
....
<(V:h) (W:D) (S:D)>	7.81	20.94
<(S:D) (S:D) (W:P)>	7.87	20.28
<(S:R) (V:h) (S:D)>	7.89	26.15
<(W:D) (S:R) (S:D)>	7.89	26.15
....
<(S:D) (W:D) (S:R)>	8.23	25.10
<(S:D) (V:h) (W:D)>	8.33	20.88
....

平均共起個数: マッチした区間に生起していた特徴数の平均
 平均区間長: マッチした区間の平均時間長

4.4 呼応パターン

呼応パターンは以下の基準で選択した。

- 3つ以上の要素(特徴)を含む頻出パターンに含まれる部分パターンであり、2つ以上の要素の組である。ただし、簡単のため、本稿では2要素に限定している。
- 生起区間に他の無関係な特徴が生起する数が少ないパターンが望ましい。

前者の基準は、生起確率の高い特徴どうしが意味的な関係もなしに共起するパターンを排除するために必要となる。個々の特徴の生起確率が高ければ、その組み合わせが偶然に生起する確率も高くなるためである。後者の基準は次の考え方による。2特徴間に他の特徴が挟まれれば、2特徴間の関係よりも、その間に挟まれている特徴と2特徴のいずれかの方が強い相互関係を持つ可能性もある。

表2には、各頻出パターンに対する上記の統計量を併記している。これらを基に抽出された呼応パターンを表3に示す。

4.5 特徴の「間」

3.で既に述べたように、本研究では、特徴が存在しないという観点からの「間」を呼応パターンとは異なるコミュニケーションパターンと考える。まず、特徴カテゴリ F_r に属する特徴のインスタンスが時刻 t_i に生起したことを $F_r(t_i)$ とする。間は、特徴カテゴリの集合 $\{F_r\}$ のインスタンスが $t_i \leq t \leq t_i + \Delta t$ に生起しなかったこと、つまり、 $\{\neg F_r(t) \vee F_r \in \{F_r\}, t_i \leq t \leq t_i + \Delta t\}$ を表す。以下では、それを $M(\{F_r\}, t, \Delta t)$ と表すことにする。

5. アドバイスとコミュニケーションパターンの関係

呼応パターンを構成する2特徴のインスタンス $f_p \in F_p$ と $f_q \in F_q$ が時間間隔 (f_p の終了時刻 t_{pe} と f_q の開始時刻 t_{qs} の

表 3 頻出パターンから抽出された呼応パターン

呼応パターン	間隔の平均	間隔の標準偏差
<W:D S:D>	-0.207	2.008
<S:D W:P>	-0.257	2.121
<S:R W:D>	0.128	2.517
<W:U S:D>	-0.499	1.785
<S:D W:D>	-0.127	2.187
<V:h S:D>	-0.270	2.492
<S:R V:h>	0.055	2.695
<W:D S:R>	0.004	2.626
<W:D S:P>	-0.198	2.328
<S:D V:h>	0.040	2.661
<V:l S:D>	-0.069	2.449
<S:D V:m>	0.131	2.680

間隔: 一番目の特徴の終了から二番目の特徴の開始までの時間差
 間隔の平均: 2特徴間の間隔の平均
 間隔の分散: 2特徴間の間隔の標準偏差

差) Δt_1 で生起したことを $R(F_p, F_q, \Delta t_1)$ と表すことにする。正の Δt_1 は通常の間を表し、負の Δt_1 は2つの特徴が時間的に重なったことを表す。また、アドバイスの生起を $A(t)$ と表す。

これらの記号を用いて呼応パターンとアドバイスの相互情報量は以下のように表すことができる。

$$I_r(R(F_p, F_q, t_{pe}, \Delta t_1); A(t_{pe} + t_2)) \text{ s.t. } 0 \leq t_2 \leq t_{th} \quad (1)$$

ここで、 t_{th} は呼応パターンとアドバイスが関連していると考えられる時間差の上限である^(注*)。同様に、間とアドバイスの相互情報量は以下のように表せる。

$$I_p(M(\{F_r\}, t, \Delta t_1); A(t + t_2)) \text{ s.t. } 0 \leq t_2 \leq t_{th} \quad (2)$$

(1)式と(2)式で与えられる相互情報量は、呼応パターンおよび間の時間的な性質とアドバイスの生起との相互関係を調べる良い手がかりとなる。この相互情報量の大小がそれぞれのコミュニケーションパターンとアドバイスとの共起のしやすさ、つまり、相互関係の強さを表すとともに、時間差 t_1 による値の変化が、相互関係の時間的な性質を良く示す。例えば、時間差 t_1 の時に I_r または I_p が大きな値を示すときには、その時間差がアドバイスを誘発する良いタイミングであることを示唆している。それとは対照的に、どのような t_1 においても I_r または I_p が大きく変化しない場合、時間的な性質がアドバイスの生起に影響しないことを示唆する。

(注*)時間的に離れているほど両者が直接関連している可能性が小さくなる。ただし、厳密に上限を決めることはできないため、ここでは便宜上、十分に大きな値を用いることにする。

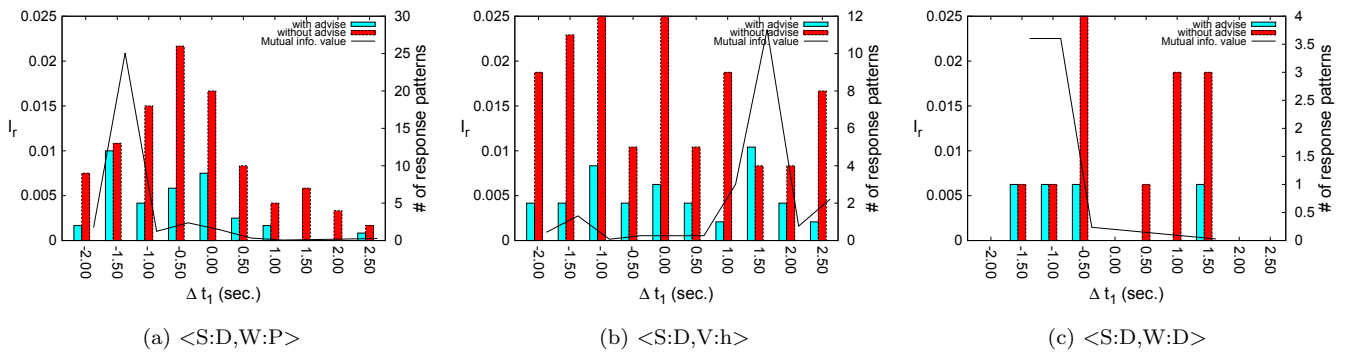


図5 アドバイスの発話と呼応パターン中の時間差との関係。実線が相互情報量を表し、棒グラフが実際の生起数を表す（共起した場合が青、共起しなかった場合が赤）。

6. 実験

6.1 実験の構成

本実験では、遠隔調理支援中に起こるコミュニケーションの記録をもとに、これまでに述べてきた呼応パターンと間の時間的な特徴とアドバイスの出現との関係を調査した。

まず、遠隔調理支援システムを図1で示したように構成した。USBカメラとマイクで得られた調理者の映像と音声はQVGA品質でSkypeを通じて支援者に送られ、支援者側の情報は音声のみ送られる。

実験データに多様性を持たせるため、習熟度や立場の違う20代～30代の男女8人で、作業者と支援者の組み合わせを変えて12サンプルを取得した。ただし、被験者の多くは、レシピや役割を変えて複数の試行に参加している。調理者は、調理器具や調味料などがどこにあるかを知らされていないキッチンで、簡単に書かれたレシピのみを与えられ、作ったことが無い、もしくは作り慣れていない料理を作る。支援者にはあらかじめ調理の詳細やコツを書き留めたものを与え、それを精読しておいてもらった。各レシピの調理にかかる時間は約30分である。

本研究では自動的に抽出できる可能性が高い特徴を用いているが、現時点では完全な精度を期待できないため、本実験では得られた記録データを人手で処理して各特徴を抽出した。なお、支援者は音声のみを作業者に送っているため、アドバイスが与えられているかどうかは、発話のみで判断することができる。コミュニケーションパターンとして、頻出パターンは表2に示したものを、呼応パターンは表3に示したものを使用した。間の検出には、表1に示した特徴とその組み合わせを用いた。

6.2 実験結果

12サンプル中のアドバイスの総数は212であった。対象とした呼応パターンや間の中に十分な数のサンプルが得られなかったものがあるが、それらについては今後の課題とする。

図5にアドバイスと呼応パターンの時間差との典型的な関係を示す。図5(a)は<S:D,W:P>、つまり、支援者が説明をした後、作業者がそれに返事をしている、最も簡単なパターンの例である。このパターンでは、時間差が-1.5秒の時に相互情報量が大きくなっている。これは、二つの発話が少し重なっている時、つまり最初の発話が終了する前に次の発話が開始している時にアドバイスがよく引き出されていることを表す。この発話の重なりにより、会話の良いテンポが保たれ、支援者は「自分の言っていることを調理者がしっかりと聞いている」ことがわかるためである。逆に、この時間差が正の方向に大きくなるとアドバイスが減少する。特に、時間差が+1秒の時、最もアドバイスが出にくい結果となった。これは、この時間差が円滑な意思疎通に適さないテンポとなっていることを示唆している。

図5(b)は<S:D,V:h>、つまり作業者の視線が停留した(頭部の動きが止まった)後、支援者が説明を行ったマルチモーダルなケースである。約+1.5秒の時間差において相互情報量のピークがあり、アドバイス数が顕著に多くなっている。これは、支援者による何らかの説明の後、調理者の視線が停留したことから、追加のアドバイスを必要としていることを支援者が理解したためだと考えられる。このように、調理者のしぐさが暗にアドバイスを要求する情報となっていることが実際に定量的に裏付けられている。対照的に、支援者が発話を終える前に作業者が視線を停留させた場合には、視線の停留とアドバイスとの間にはっきりとした関係性は見られなかった。

これらに対し、図5(c)の<S:D,W:D>は、支援者の説明に対して調理者が説明を返している場合である。時間差が-0.5秒付近の場合、つまり、支援者の発話に対して調理者が説明を少しかぶせている場合にはアドバイスが現れにくいことなど、大まかな傾向は見取れるが、それがはっきりと確認できるほどのサンプルが得られていない。このようなパターンの調査は今後の課題とする。

複数の呼応パターンにおいて、負の時間差に相互情報量のピークがあり、この時間差がアドバイスを誘発していると考えられる。個々の例でも、 $-1.5 \sim -0.5$ 秒の負の時間差が起こっている場合は、それぞれが相手の言動をよく把握して、タイミング良く反応している状況が多く、これはアドバイスを誘発する良い状況であると言える。また、正の時間差において相互情報量のピークがありパターンもある。これは、支援者が調理者を観察して調理者の状況を理解し、アドバイスの必要性や内容を考えるために時間が必要を要したことを示唆している。

図 6 は、発話と間の長さとの関係の典型的な例である。図 6(a) は {S:D, W:D}、つまり説明が行われていない「間」の場合を示す。間が小さい場合にはアドバイスが抑制される傾向がある。それを過ぎると、アドバイスが与えられる割合が増えている。沈黙によって、アドバイスが暗に要求されていることが示唆されている。

また、現在の状況に対するアドバイスと、未来の事柄に対するアドバイスの生起に違いが見られる。図 6(b) は現在の状況に対してアドバイスが与えられたケースであり、間が 3 秒以上の場合、より促進されている。図 6(c) は未来に対するアドバイスが与えられたケースであり、こちらは約 1 秒程度の短い間の時に多く出ている。この結果からは、現在の状況に対してアドバイスをする場合、支援者が調理者の言動からアドバイスが必要なことに気付くまでに時間を要したことを意味する。

最後に、図 6(d) に、{S:D, V:h} のマルチモーダルなケースを示す。これは、作業者の視線の停留と支援者の説明の組み合わせによる間である。この図から、間が約 1 秒と 3 秒の時に相互情報量が大きいことが分かる。しかし、共起数は 3 秒両者で顕著に違いが見られる。つまり、アドバイスが与えられる割合は、間が 2 秒以下だと減り、間が 2 秒以上だと増えている。これは、視線が動いている（頭部の動いている）時間が 2 秒以上の場合、調理者がきょろきょろしながら暗にアドバイスを求めていると作業者が判断し、2 秒以下の場合には単に場所や何かを確認するために視線が動いただけだと判断しやすいことを示唆する。

以上の結果は、3. でアドバイスとコミュニケーションのパターンに関する我々の予想とおおむね合致している。良いテンポで起こる反応と間はアドバイスを促進すること、また、アドバイスを与えるために時間を要することなどが統計的に裏付けられている。

7. まとめ

本稿では、一人称映像を用いた遠隔調理支援の形態を紹介し、そのコミュニケーションの解析方法について紹介した。具体的には、2 者間のコミュニケーションと支援者からのアドバイスの出やすさとの関係を解析するためのコミュニケーションのパターン化と時間的特徴の設定、さらに、相互関係を定量化する方法につ

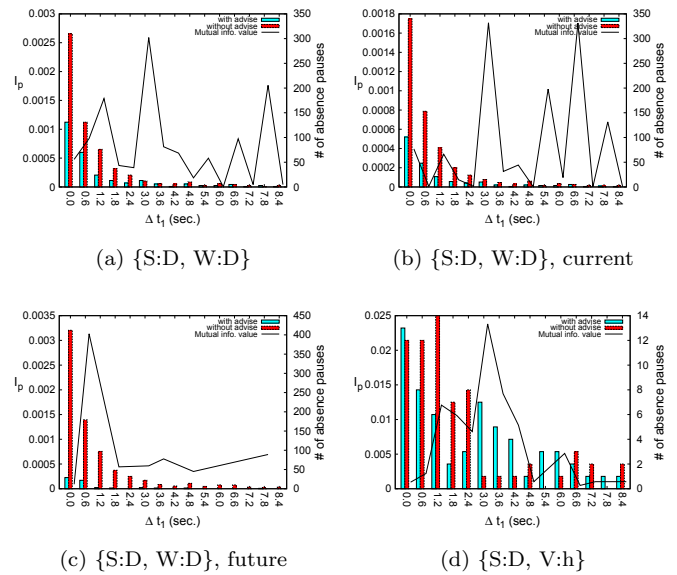


図 6 間とアドバイスの生起の関係。実線は相互情報量を表し、棒グラフは棒グラフが実際の生起数を表す（共起した場合が青、共起しなかった場合が赤）

いて議論した。また、実験によって、呼応パターンや間のパターンの時間的な特徴とアドバイスの発生には密接な関係があることを統計的に示した。例えば、応答のタイミングや間の長さがアドバイスの生起に大きく関係することなどである。ただし、これまでの実験で収集したデータでは数が不足し、十分な確認が行えなかったパターンもある。更に多くのデータを収集し、アドバイスの出現に影響しているパターンをより多く調査することが今後の課題となっている。また、より良いコミュニケーションを保ち続けられるような遠隔調理支援システム的设计や実装について取り組んでいく必要がある。

文献

- [1] P. Garner, M. Collins, S. Webster, D. Rose, "The application of telepresence in medicine", BT Technology J, Vol.15, No.4, pp.181-187, 1997
- [2] J. Siegel, R. Kraut, B. John, K. Carley, "An Empirical Study of Collaborative Wearable Computer Systems", Proc of the ACM Conference on Computer Supported Cooperative Work (CSCW1995), pp.312-313, 1995
- [3] R. Kraut, M. Miller, J. Siegel, "Collaboration in performance of physical tasks: Effects on outcomes and communication", Proc of the ACM Conference on Computer Supported Cooperative Work (CSCW 1996), 1996
- [4] S. Fussell, R. Kraut, J. Siegel, "Coordination of communication: Effects of shared visual context on collaborative work", Proc. of the ACM Conference on Computer Supported Cooperative Work (CSCW 2000), 2000
- [5] M. Billinghurst, S. Bee, J. Bowskill, H. Kato, "Asymmetries in Collaborative Wearable Interface", The 3rd International Symposium on Wearable Computers, pp.133-140, 1999
- [6] R. Kraut, D. Gergle, S. Fussell, "The Use of Visual Infor-

- mation in Shared Visual Spaces: Informing the Development of Virtual Co- Presence”, Proc. of the ACM Conference on Computer Supported Cooperative Work (CSCW 2002), 2002
- [7] D. Gergle, R. Kraut, S. Fussell, “Action as language in a shared visual space”, Proc. of the ACM Conference on Computer Supported Cooperative Work (CSCW 2004), 2004
- [8] H. Clark, “Referring as a collaborative process”, *Cognition*, Vol.22, pp.1-39, 1986
- [9] S. Kubota, Y. Nakamura, Y. Ohta. “Detecting Scenes of Attention from Personal View Records”, IAPR Workshop on Machine Vision and Applications, pp. 209-213, 2002.
- [10] 小泉敬寛, 小幡佳奈子, 渡辺靖彦, 近藤一晃, 中村裕一: “映像対話型行動支援における頻出パターンに基づいたコミュニケーションの分析”, *情報処理学会論文誌*, Vol.56, No.3, pp.1068-1079, 2015
- [11] 小泉敬寛, 小幡佳奈子, 渡辺靖彦, 近藤一晃, 中村裕一: “映像対話型行動支援におけるインタラクションの一貫性に関する考察”, *HCG シンポジウム 2014*, pp.49-56, 2014
- [12] J. Pei, J. Han, B. Mortazavi-asl, H. Pinto, Q. Chen, U. Dayal, M. Hsu, “PrefixSpan: Mining sequential patterns efficiently by prefix-projected pattern growth”. 17th International Conference on Data Engineering (ICDE '01), pp.215-224, 2001