

会話量子化法を用いた会議知識獲得支援

Meeting Knowledge Capture Assistance by Conversation Quantization

齊藤 憲
Ken Saito

久保田 秀和
Hidekazu Kubota

角 康之
Yasuyuki Sumi

西田 豊明
Toyoaki Nishida

京都大学大学院情報学研究科
Graduate school of Informatics, Kyoto University

In this study, we present meeting knowledge capture assistance by conversation quantization. The concept of the conversation quantization is a technique for approximating a continuous flow of conversation by a series of conversation quanta that correspond to the points in a discourse. We have attempted to implement the meeting knowledge capture assistant system on the hypothesis that the intention of participant is useful to improve the reusability of movie conversation quanta. The proposed system captures meeting knowledge as movie conversation quanta that are reflected user's intention by pushing button device. We experimentally attempted to capture movie conversation quanta from meetings by using the system. As a result, it is confirmed that movie conversation quanta in which the intention of participant is reflected are captured and conversation quanta are effective in capturing meeting knowledge.

1. 序論

我々は社会活動のなかで、活動報告、意見交換、意思決定などのために様々な規模、議題で、会議を行う。会議の場では多種多様な知識が共有され、新しいアイデアや知識が創造される。しかし、会議の場において生まれた知識は記録を残さない限り失われやすいものである。

従来から会議によって生じた知識は、様々な手法で記録されてきた。最も古典的なものとして議事録やメモといったテキストベースの記録があるが、会話の非言語的な側面を十分に残すことができない。また、近年、コンピュータ技術の発展により、映像や音声などマルチメディアデータを統合した映像議事録が研究されているが、再利用性の高い映像クリップの抽出は依然困難である。

西田[Nishida 05]は、知識流通メディアとしての会話に焦点を当て、実世界の会話によって生じた知識をコンテンツ化し、流通可能とするシステムの枠組みとして会話量子化を提案している。会話量子化とは、連続した会話の流れから離散的な会話の粒を切り出して持続的に蓄積することにより、元とは異なる会話の場において元の会話のエッセンスを再利用可能とする枠組みである。この会話の粒のことを会話量子と呼ぶ。

我々の研究グループでは、会話量子の概念を会話映像クリップを中心とした会話情報パッケージ(映像会話量子)とし、会話映像から手作業で映像会話量子を抽出し、再利用シミュレーションを行い、再利用可能であるという知見を得ている[齊藤 05]。本論文では、映像会話量子の再利用性の向上には、量子獲得者の意図が重要であるという仮説のもと、再利用が容易な量子を獲得するために、参加者が会議中に意図した量子を獲得することができる会議知識獲得支援システムの構築を目指す。以下、2章で会話量子化の概要を述べ、3章で会議知識獲得支援システムの概要について述べる。4章でシステムを利用した映像会話量子獲得実験について述べ、5章で提案システムについて議論する。6章はまとめである。

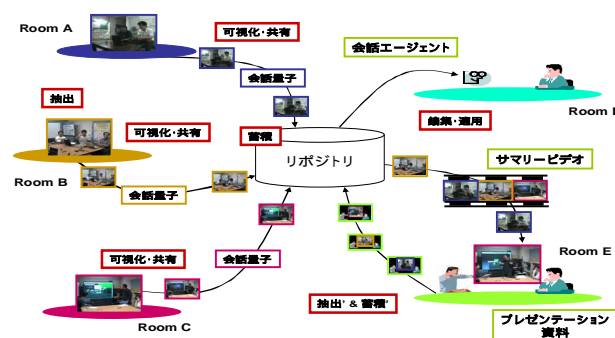


図 1: 会話量子化の全体図

2. 会議知識の獲得

本章では、会議知識の獲得のために本研究で採用する会話量子化の概要と会話量子の再利用性について述べる。

2.1 会話量子化の概要

会話量子化とは、連続した会話の流れを会話量子と呼ばれる離散的な粒の流れとして近似することにより、会話を部品化する技術である[Nishida 05](図 1)。この会話の粒のことを会話量子と呼ぶ。これまで、我々の研究グループにおいて、会話量子はテキストと参照画像で表現され、議論されてきたアプローチであり、一定の再利用性が確認されている[久保田 03]。本論文では、会議からより再利用性の高い知識を抽出するために、会話量子の概念を拡張し、会話映像を中心とした会話情報パッケージから構成される映像会話量子という概念を導入する。

映像会話量子は会話によって生じた、あるひとまとまりの知識を内包する映像コンテンツ素材であり、会話の場から抽出され、その場で可視化・共有される。様々な会話の場(Room A, B, C)で抽出された映像会話量子はリポジトリに蓄積され、他の会話の場(Room D, E)において、映像会話量子を組み合わせることで作成される会話エージェント[久保田 03]やプレゼンテーション資料、サムリビデオなどの会話コンテンツの部品となって再利用される。

本研究は、会議知識獲得支援のため、会話量子化法を採用し、会議知識を映像会話量子として獲得することを目指す。

連絡先: 齊藤憲(Ken Saito), 京都大学大学院情報学研究科,
京都市左京区吉田本町工学部 10 号館 223 号室,
saitoh@ii.ist.i.kyoto-u.ac.jp

2.2 獲得者の意図を反映した映像会話量子

我々は映像会話量子の再利用性について、会議映像から手作業で会話量子を抽出し、再利用するシミュレーションを行い、研究してきた[齊藤 05]。その結果、たとえ映像が断片的であったとしても量子の持つ背景知識を共有した者であれば理解可能である、元となった会話で利用されていた資料など会話映像にコンテキストを与えるような資料と対応付けて記録すれば再利用が容易になる、といった知見を獲得している。

その後、これらの実験データを再評価した結果、自分の参加していない会話から有用な知識を取り出すことは難しいという知見を得た。これは会話参加者と背景知識を共有していない、会話の臨場感を体験していないなどの理由が考えられる。加えて、映像会話量子の再利用性の向上には獲得者の意図が反映された量子を獲得することが重要であるという知見を得た。自分が再利用したいという意図の基で、映像会話量子を獲得しようとすると、再利用者である量子化作業者の主観に大きく左右されるが、映像会話量子の再利用性を高めるためには、獲得者の意図が反映された映像会話量子を獲得することが重要であり、言い換えれば、獲得者が獲得したいと思った会議知識、会話シーンを含む映像会話量子は再利用者である獲得者にとって再利用性が高いと考えられるためである。

そこで、本論文では、会議参加者が会話中に自分の意図した映像会話量子をキャプチャできるようなシステムを開発することを目指す。抽出する映像クリップの範囲をシステムに対して指示するデバイスとしてボタンデバイスを用いる。

3. ボタンデバイスを用いた映像会話量子獲得支援システム

本章ではボタンデバイスを用いた映像会話量子獲得支援システムについて述べる。なお、本論文ではスライド資料を用いた3,4名ほどの会議を会話量子化の対象とする。

3.1 映像会話量子獲得支援システムの概要

本論文で提案する映像会話量子獲得支援システムの概要図を図2に示す。提案システムでは、全体カメラによって会議の様子を常に記録している。参加者によってボタンデバイスが押され、映像会話量子(映像クリップ)が抽出され、大型ディスプレイ上に可視化される。

提案システムは量子獲得システムと量子可視化システムから構成される。量子獲得システムにはボタンデバイスを採用したシステムを物理インタフェース作成キットである Phidgets と C#を用いて開発した。ユーザの意図に基づいて in 点(抽出開始点)、out 点(抽出終了点)を決定可能なボタンデバイスによって、指定された区間の映像クリップが抽出されるシステムである。量子可視化システムには、久保田らによって開発された知球[Kubota 05]と呼ばれるコンテンツマネジメントシステムを採用する。映像会話量子は、知球によって、大型ディスプレイに知球カードと呼ばれるカード画像として可視化される。抽出された映像会話量子は自動で知球上に知球カード(量子カード)としてインポートされ、図3のように関連する知球カード(資料カード)と関連付けられて可視化される。提案システムのボタンデバイスの概観を図4に示す。

3.2 量子獲得システム: ボタンデバイス

会議参加者が会話中に自分の意図した映像会話量子をキャプチャできるようなシステムにおいて、会話量子獲得デバイスが持つべき性質は2つ考えられる。

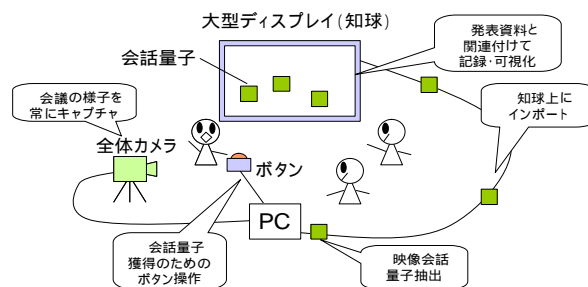


図2: 映像会話量子獲得支援システム



図3: 知球の概観



図4: ボタンデバイスの概観

- (a) 会話を阻害しない
- (b) in 点, out 点を意図した時点へ遡及可能

(a) 会話量子獲得者は同時に会話参加者でもあるため、会話を阻害せずに映像会話量子を獲得できるデバイスが望ましい。
 (b) 会話量子を抽出するためには、in 点と out 点を指定する必要がある。これらを参加者がある程度任意に決めることが出来ることが望ましい。また、参加者が会話量子の獲得意思を表示する時点には、対象会話が始まる前、会話の途中、会話終了時の3パターンが考えられるため、複数の抽出パターンを持たすことが可能であることが望ましい。言い換えれば、それぞれ、現在から未来までの会話に対する獲得意図、過去から未来までの会話に対する獲得意図、現在から過去の会話に対する獲得意図が考えられる(図5)。図5の画像は anvil[Michael 04]を用いて分析をした様子である。

以上の性質を備えたデバイスとして、本研究ではボタンデバイスを採用した。(a) 会話を阻害しないという点から考えると、細かい操作よりも、ボタンデバイスのようにデジタルな押下回数で大雑把に in 点と out 点を決定するほうが直感的であり、会話の阻害を抑えることが出来る。(b) ボタンデバイスには、連続でボタンを押す行為(連続押下)とボタンを押し続ける行為(押下保持)というシンプルなパターンがあり、これを組み合わせることで、

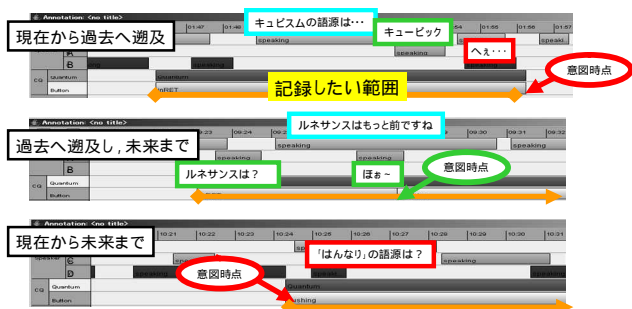


図 5: 意図した時点と記録範囲の関係

複数の抽出パターンを設定することが出来る。以上の理由により、ボタンデバイスを採用した。

会話量子を抽出するには、in 点と out 点を決める必要があることは既に述べた。そこで、ボタンデバイスによって in 点と out 点の決定をどのように行うかが課題となる。参加者が会話量子の獲得意思を表示する時点は、3 パターンあると考えられることは既に述べた。そのうち、会話中もしくは会話終了時点では、対象となる会話は既に始まっている、または終了している。この会話を会話量子として獲得するためには、参加者が思う会話の開始点を in 点と設定すればよいことになる。参加者による会話量子獲得のためのボタンデバイスの操作開始点(獲得意図時点)を push 点とすると、push 点からの遡及量をボタン操作によって決定し、in 点を決定すればよい。そこで本論文では、ボタンの連続押下数によって遡及量を決定する仕様を採用した。

4. 映像会話量子獲得実験

本章では、獲得者の意図した映像クリップの獲得可能性を検証するための実験について述べる。

4.1 概要

本実験は、ボタンデバイスによる獲得者の意図を反映した映像会話量子獲得可能性の検証を目的とする。また、ボタンデバイスの仕様を 4 つ準備し、それぞれの仕様で実験を行うことで、ボタンデバイスのソフトウェア上の仕様についても議論する。

4.2 実験方法

本実験では、提案システムを用いて、参加者 3 名による 30 分ほどのミーティングを行う。基本的には参加者のうちの 1 名による発表形式で進んでいくが、参加者はいつでも自由に発言ができる。ミーティング参加者のうちの発表者ではない 1 名がボタンデバイスを持ち、任意のタイミングでボタンを操作し、量子を獲得する。ミーティング終了後、ラップアップセッションを開き、ボタンを押した意図が反映された会話量子が獲得されたかどうかを検証しつつ、in 点、out 点における獲得を意図した範囲との過不足を秒数で記入してもらう。加えて、映像会話量子、ボタンデバイスについてのアンケートを行う。

上記のミーティングを、表 1 に示すような、押下パターンと 1 押下あたりの遡及量が異なる 4 つの仕様において、各仕様 2 回ずつ、合計 8 回行う。押下パターンは押下保持があるパターン 1(図 6-a)と保持がないパターン 2(図 6-b)を準備した。パターン 1 は、まず、連続押下によって push 点からの遡及量を決定し in 点を決定、次に押下保持することで終了のタイミングを計り、ボタンを放すことで暫定 out 点を決定する。最後に連続押下によって暫定 out 点からの遡及量を決定する。パターン 2 は、会話が終了したと思われる時点でボタン操作を開始することを前提としており、in 点のみ遡及可能で、out 点は連続押下が

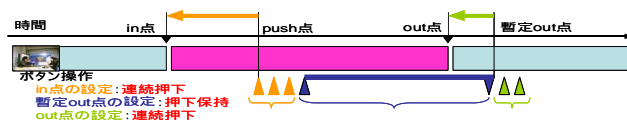


図 6-a: 押下パターン 1

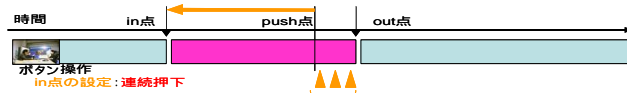


図 6-b: 押下パターン 2

表 1: ボタンの仕様(押下パターンと遡及量)

仕様	押下パターン (図 6)	1 押下あたりの遡及量 (in 点)	1 押下あたりの遡及量 (out 点)
1	パターン 1	3 秒	3 秒
2		5 秒	5 秒
3		10 秒	10 秒
4	パターン 2	5 秒	5 秒

表 2: 映像会話量子獲得実験結果

仕様	獲得者	会議時間 (秒)	会話量子数 (個)	会話量子の平均秒数 (秒)
1	A	1680	16	42.5
	B	1200	12	38.8
2	C	1860	10	83.4
	D	1800	15	43.0
3	E	1740	21	45.6
	F	1620	11	71.5
4	G	1800	22	16.4
	H	1980	20	26.5

終了した時点となる。

4.3 実験結果

実験結果を表 2 に示す。8 回のミーティングの後、8 名の獲得者から合計 127 個、平均の長さ 45.96 秒の映像会話量子を獲得した。ミーティングのテーマは「医療問題」「近代美術」「京都大学」など多岐にわたる。

4.4 考察

実験のアンケートにおける、ボタンデバイスによって獲得者が意図した会話がキャプチャできたかどうか、という質問に対して、「ほぼ獲得できた」とのアンケート回答を被験者 8 名全員から得た。したがって、ボタンデバイスによって獲得者の意図を反映した映像会話量子の獲得可能性は高いといえる。

(1) 遡及量(仕様 1, 2, 3)による比較

各仕様における in 点、out 点の過不足の平均秒数を図 7 に示す。過不足の平均は、過不足の絶対値の平均をとった。図 7 より、仕様 2 が獲得意図を反映しやすいことが分かる。10 秒では大雑把過ぎるため過不足が多くなる。一方で遡及量が細かければよいというものではないことも分かる。ボタン押下数に関して、仕様 1~3 の間で押下回数平均に顕著な差は見られず、ほぼ 1 回か 2 回と少なかった。これは、多数の押下は操作者にとって負荷が大きいためであると考えられる。そのため、必ずしも細かい単位(仕様 1)が意図通りに遡及しやすいとは限らないと考えられる。

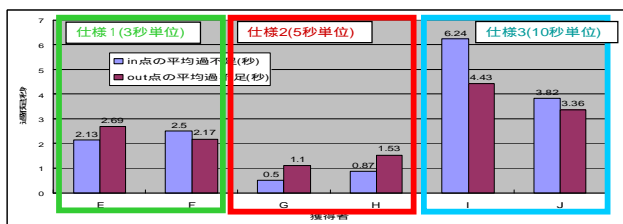


図 7:意図した範囲に対する平均過不足

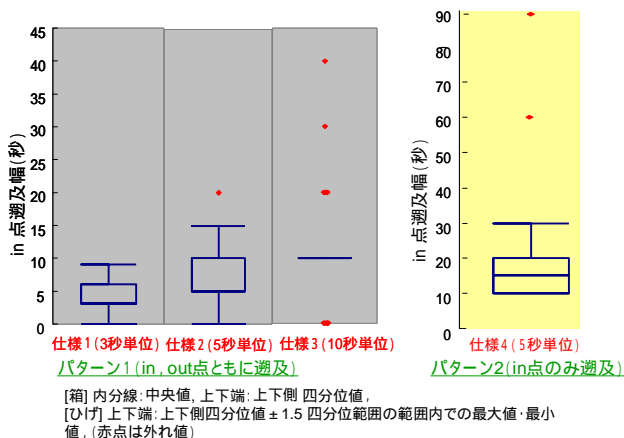


図 8: 押下パターンによる比較

(2) 押下パターン(仕様 1, 2, 3 と仕様 4)による比較

押下パターン 1 とパターン 2 を比較すると、パターン 1 は比較的長い会話量子を少なく、パターン 2 は比較的短い会話量子を多く獲得するという傾向が見られた。

また、in 点に対する遡及幅の比較のため、各仕様における in 点遡及量の箱ひげ図を図 8 に示す。図 8 から、およそ 5 から 15 秒程度遡っていることが分かる。これは、人が別途会話をしながらおよそ 15 秒程度過去まで記憶を遡及できる、逆に言えばそれ以上は困難であることを示している、と考えられる。押下保持のないパターン 2 の場合、あまり長くは遡及できず、短い量子を獲得する傾向が現れたと考えられる(会話量子の平均長: パターン 1(54.1s) > パターン 2(21.4s))。また、遡及量がパターン 1 と比較して多いパターン 2 のほうが意図反映の精度が低いことが確認された(過不足の絶対値の平均: パターン 1(2.67s) < パターン 2(9.62s))。

また、パターン 1 と比較すると、パターン 2 はボタン操作がシンプルであり、押下保持のようなジェスチャを制限するような押下パターンがないため、筆者の内観ではあるが、ボタン操作者が、より自然に、より積極的に会話に参加ができたと思われる。会話を阻害しないという観点で見れば、パターン 1 よりもパターン 2 がより有効であると考えられる。

5. 議論

実験のアンケートでボタンデバイスを用いた映像会話量子獲得システムに対し、8 名全員が「会議知識の獲得に役に立った」と肯定的回答を得たことより、映像会話量子が会議知識の獲得に有効であるという知見を得ることができた。

ボタンデバイスに関して、8 名全員より「ボタン操作は容易に学習できそう」との回答を得ており、学習可能性が高いと考えられる。人が直感的に時間長を判断するには自分なりの基準が必要であり、絶対的に何秒かを判断するよりも、いつもの基準時間よりも長い/短いかで判断するほうが容易であると考えられる。本論文で用いたボタンデバイスの仕様は、デジタルな値である

押下回数による遡及量決定のため、基準となる回数が感覚的に記憶しやすく、学習しやすいと考えられる。また、秒数ごとにボタンを準備する複数ボタンで決定する場合、あるいはダイヤルのつまみによって秒数を決定する場合と比較したときには、ワンボタンの押下回数による決定はユーザがわざわざデバイスを目で見ることなしに遡及量を決定できる点でカジュアルであると予想している。

加えて、ボタン操作のうち、ボタン保持がジェスチャを妨げる恐れがあるという知見を得た。片手でボタンを保持しつつ、もう一方の手でジェスチャをしながら話すというシーンが観察された。さらに、未来に対する判断(話の終わりの見極め)が困難であるため、out 点の設定が難しいという意見も寄せられた。

獲得者の意図を反映可能な会話キャプチャデバイスのデザインポリシーに関して、in 点、out 点の遡及手段、量子の分割・結合手段が必要である。発話の切れ目に関する小さなずれは音声信号処理的に解消できると考えられるが、遡及に際する時間認知への負担、操作負担(押下回数等)を低くし、ボタン操作のずれをなるべく大きくしないようにする必要がある。そのためのアイデアとして、取得中の映像クリップの秒数や遡及量等を可視化するとよいというアイデアがアンケートの回答にあった。

今後の展望として、獲得者の意図を自動で汲み、映像会話量子を段階的に提示するような、インタラクティブに量子を抽出できるような映像会話量子提案システムが考えられる。また、本論文では獲得・抽出部について議論してきたが、獲得した映像会話量子を用いたコンテンツの作成・再利用を行い、再利用性の検証を行う必要がある。

6. 結論

本論文は会議知識を獲得するために映像会話量子獲得システムの実装を目指した。会議参加者が会話中に自分の意図した会話映像をキャプチャできるような映像会話量子獲得デバイスとしてボタンデバイスを採用し、システムを実装した。提案システムによって実践的な実験を行い、ボタンデバイスによって獲得者の意図を反映した映像会話量子を獲得できること、映像会話量子が会議知識の獲得に有効であることを確認した。

参考文献

[Nishida 05] Toyooki Nishida: Conversation Quantization for Conversational Knowledge Process, Special Invited Talk, S.Bhalla (Ed.): DNIS 2005, LNCS 3433, Springer, pp.15-33 (2005).

[久保田 03] 久保田 秀和, 黒橋 禎夫, 西田 豊明: 知識カードを用いた分身エージェント, 電子情報通信学会論文誌「ソフトウェアエージェントとその応用論文特集」, volJ86-D-I, No.8, pp.600-607, 2003.

[齊藤 05] 齊藤 憲, 久保田 秀和, 角 康之, 西田 豊明: 会話量子を用いたミーティングのコンテンツ化支援, 人工知能学会全国大会(第 19 回)論文集, 3D3-12, 2005

[Kubota 05] Hidekazu Kubota, Yasuyuki Sumi, Toyooki Nishida: Sustainable Knowledge Globe: A System for Supporting Content-oriented Conversation, in Proceedings of AISB 2005 Symposium Conversational Informatics for Supporting Social Intelligence & Interaction, pp.80-86 (2005).

[Michael 04] Michael Kipp: Gesture Generation by Imitation – From Human Behavior to Computer Character Animation, Boca Raton, Florida: Dissertation.com (2004).