

車内会話の量子化と再利用

Quantization and Reuse of Driving Conversations

岡村 剛^{*1}
Go Okamura

久保田 秀和^{*1}
Hidekazu Kubota

平松 達也^{*1}
Tatsuya Hiramatsu

角 康之^{*1}
Yasuyuki Sumi

西田 豊明^{*1}
Toyoaki Nishida

塚原 裕史^{*2}
Hiroshi Tsukahara

岩崎 弘利^{*2}
Hirotoshi Iwasaki

^{*1} 京都大学大学院情報学研究科
Graduate School of Informatics, Kyoto University

^{*2} (株) デンソーアイティラボラトリ
DENSO IT LABORATORY, INC.

In this paper, we propose a framework for capturing and reusing conversation data in order to argument drive experiences. Firstly, we define an in-vehicle conversation quantum as a reusable conversation unit in driving conversation. It consists of a conversation video clip, a video clip of landscapes from car windows and metadata such as GPS. Secondly, we have implemented a drive simulator can capture and reuse pseudo in-vehicle conversations. On this simulator, we have implemented a system which can detect pointing gestures and show related information on the pointed location in the videos. We have also implemented a system which quantizes driving conversation quantum seems to be good for offering a new topic. It was found that some pointing patterns could be good clues of quantization.

1. はじめに

自動車は目的地まで安全に効率良く移動するために様々な機能が改良を重ねられている。また、カーナビなどのような車内における情報システムの機能は、単なる経路案内にとどまらずエンタテインメント面についての機能も強化されており、車内体験を強化する機能への要望、期待が高まっていることがうかがえる。

ドライブでは目的地で楽しむ以外に、移動中の出来事も価値ある体験として思い出に残ることが多い。この体験を記録したり共有したりするために、写真や Blog 等がよく用いられている。SONY のカーナビ“Viam”^{*1} はこのような点に着目し、デジタルカメラで撮影した写真の撮影時刻と、ドライブ中の GPS 情報を同期させ、アルバムを作る機能を搭載している。

本研究では、ドライブ体験における車内会話に着目する。車内における会話は窓から見える景色や旅の目的、現在位置、時間等に対応した興味深い内容を含んでいると考えられる。車内会話を再利用しやすい形で記録することで、他のドライブの場で話題提供に用いてドライブ体験を充実させることや、写真でのドライブアルバムに代わる体験共有手段とすることが期待できる。

車内会話の内容を当事者以外が理解するには、その時の景色や位置情報などの周辺的情報も必要となる。再現性の高い形で残すためには映像や音声を用いることが求められるが、編集に多くの手間がかかるという問題がある。この問題を解決するためには情報システムの支援が必要である。

以上を踏まえ、本研究では車内会話とそれに関連する景色の映像を中心としたデータを記録し、コミュニティ内で共有できるようにすることで、ドライブ体験を充実させることを目指す。以下、2節で知識メディアとしての自動車を実現する枠組みとしての車内会話量子を定義し、手作業により車内会話量子化と再

利用を行うことで得られた知見について述べる。3節では車載カメラにより獲得した車窓映像を用いるドライブシミュレータの仕様を述べる。さらに、ドライブシミュレータ上で指差しを用いたユーザの注目行動を推定し、ユーザの指差しに基づいて情報提示を行うシステムおよび車内会話量子化・再利用を行うシステムについて述べる。4節で車内会話再利用実験とその結果について考察し、5節で関連研究に触れ、6節でまとめる。

2. 車内会話量子

車内会話は、ドライブ前にはドライブ計画に、ドライブ中にはラジオ番組や備忘録として、ドライブ後には体験の共有などに再利用されることが期待される。本節では再利用可能な単位としての車内会話量子を定義し、その実現のための要件について述べる。

2.1 車内会話量子の定義

本研究では、車内会話を再利用可能な形で記録するために、会話量子化[Nishida 2005]というアプローチを取る。会話量子化とは、連続した会話の流れを結束性の高い会話区間(会話量子)の離散系列として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。会話量子の表現手法は対象領域によって異なっており、今回のような映像を中心とする場合については、会議における会話映像クリップを量子とすることが検討され、一定の再利用性が確認されてきた[斎藤 2006]。

車内会話は、流れゆく外の風景や運転行為、道路の状況などに触発されながら行う会話である。そこで、会話に加えて車内から見た風景映像を一つのまとまりとして切り出し、話題やコンテキストを表現するメタデータと共に記録することを車内会話の量子化とする。

2.2 車内会話量子化の要件

車内会話量子の適切な構成要素や提示環境、情報システムが行うべき支援を明らかにするために、手作業で車内会話の量子化と再利用を行った。

連絡先: 平松達也, 京都大学大学院情報学研究科, 〒606-8501 京都市左京区吉田本町(工学部 10 号館 131 号室), TEL:075-753-5387, hiramatsu@ii.ist.i.kyoto-u.ac.jp

^{*1} <http://www.sony.jp/CorporateCruise/Press/200105/010515/>

作業は、著者らのうち3人がDVカメラを持って観光案内タクシーを利用し、ドライブ後に著者らのうち1名がドライブ体験をまとめたビデオを制作して研究グループ内で紹介するものとした。ビデオ制作は、話題のまとまりを列挙し、その中からハイライトと思われるものを選択、必要な映像を各種エフェクトを利用してつなぎ合わせ、時系列順に並べるという手順で行った。

映像制作により、二つの知見が得られた。第一に、利用した映像の大部分が助手席からの映像であることから、車内会話におけるスムーズな意思疎通には広い景色の共有が必要であるといえる。第二に、話題のまとまりの列挙と対応する映像の切り出しに最も時間がかかった。ドライブ中に話題の発生した場所、注目点、時間とそれに対応する映像が簡単に記録できれば、大幅に作業コストが下がると思われる。

また、研究グループ内で映像を紹介した際にも、別の知見が得られた。ドライブの当事者以外にとって、コースの空間的な手がかりは映像しかなく、それが不連続になった際に把握が困難となった。このことから、映像以外の空間的な手がかりを与えることが必要であるといえる。そして、カメラの移動やズームによって注目対象が明確になる一方で、映像の手ぶれなどの影響で長時間の鑑賞が難しくなってしまったことから、固定カメラを利用した上で、注目点の表現を補うことが望ましいと考えられる。

車内会話の量子化の要件をまとめると以下の通りになる。

- (1) ドライブ中の注目点を記録再生可能とする必要がある。
- (2) 再利用を行う際には、空間的な文脈を明らかにする必要がある。
- (3) 車内会話量子を構成する映像とその提示環境は、広い画角を持つことが望ましい。

3. ドライブシミュレーション環境の実装

実験車とドライブシミュレータを用いた車内会話量子化研究のフレームワークを提案した(図1)。その内、研究用のプラットフォームとしてのドライブシミュレータは前節の要件に基づいて設計した。フレームワークではまず(株)デンソーアイティラボラトリーで開発された実験車とドライブ記録システムを用いて、車内会話環境を再現するためのデータを記録する。次に、記録に基づいた追体験の可能なドライブシミュレータを実装し、車内会話の量子化・再利用の実験を行う。さらに、シミュレータ上での実験結果に基づいて、実車に搭載可能なシステムを開発し、実証実験を行うという実車・シミュレータ間の実装・実験・実証サイクルを繰り返すことにより、ドライブ体験支援システムをブラッシュアップする。本論文では研究用プラットフォームであるドライブシミュレータに焦点をあてて議論する。

本節では、研究プラットフォームとしてのドライブシミュレータについて述べた後、ドライブシミュレータ上で指差しを用いたユーザの注目点推定と情報提示の手法および車内会話の量子化・再利用システムについて述べる。

3.1 ドライブシミュレータの構築

実験車における車窓映像は車外に向けて前方・左右に設置されたカメラで記録される。ドライブシミュレータではこの映像データではこの映像データを組み合わせることでコースを構成可能とする。広い画角を確保するため、出力デバイスとしてユーザ正面に30インチディスプレイを1台、左右に19インチディスプレイを1台ずつ持つ。各カメラで取得した映像は同期的に再生される。入力デバイスとしては、ハンドル・ペダルからなるゲーム用コントローラを持ち、ペダルやボタン操作による早送り・巻き戻しなど

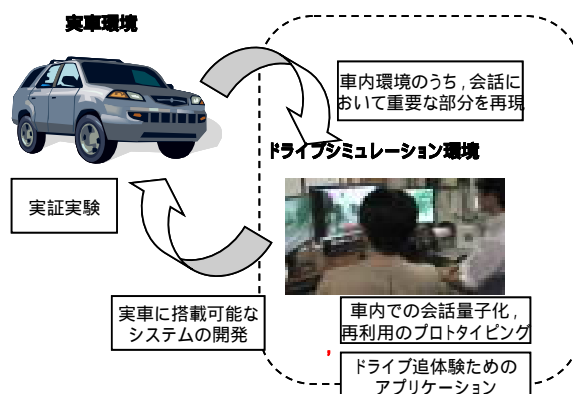


図1 本研究のフレームワーク

映像操作を行うことができる。また、撮影時刻、車速、現在地周辺の地図を表示することによって文脈を補う。

3.2 指差しを用いたユーザの注目点推定と情報提示

2.2 節で述べたように、ドライブ中に注目対象を記録することによって会話量子化の作業コストを削減することができる。また、注目対象を推定できれば、会話量子検索の手がかりとして利用したり、ユーザに対して注目対象をその場で提示したりすることもできる。

本研究では、ユーザが何かに注目した際には指差し行動を行うものと仮定し、指差しの位置を検出・記録する。システムは指差した先に何があるかを推定し、ユーザに情報提示を行う。

まず、ユーザの視点では指差したい点が指先と重なって見えると仮定する。また、指差し対象を車窓映像に限定する。以上より、両目の中心点と指先を通る直線と、ディスプレイの映像に交点があれば、その点を指差したと推定する。本研究では、後頭部を両目の中心点の近似点とし、光学式モーションキャプチャ PhaseSpace を用いて指差しの検出を行った。そして、指差された映像中の点に円形のポインタを表示する機能を実装した。

シミュレータ上の指差しは、GPS 系列データおよびカメラ角度を用いて実空間での方角と対応付けられる。そして、実空間での指差し方角をもとに、指差した先にあるオブジェクトを検索し、その名称を表示する。ここでは、指差し対象を建物に限定する。建物のデータベースにはゼンリンの電子住宅地図データから幾つかの項目を抽出したものをを用い、シミュレータ上での建物名の検索と表示はユーザの指差し行動の有無によって異なる手法をとるものとする。ユーザが指差しを行っていない場合は映像全体を含む範囲で、現在地から近いと推定された建物を近い順に5件まで表示する。ユーザが指差しを行っている場合には、指差し方角から閾値内の範囲で、現在地から近い順に上位5件を表示する。指差し方角にある建物の名称を表示している様子を図2に示す。

3.3 車内会話の量子化・再利用システム

ドライブシミュレータ上でのドライブ体験支援のために、会話や指差しの様子を記録し、異なる会話の場において再利用可能とするシステムを実装する。システムは車内会話量子化を行う部分、車内会話量子を提示する部分に分かれる。過去の車内会話量子が存在する場合には、それを提示しながら車内会話を記録することができる。

車内会話の量子化は、二段階のステップで行われる。第一ステップで車内会話を記録する。この段階で、会話の様子を撮影した会話映像、音声信号を記録した会話音声、映像再生・指差し・位置情報の各イベントを記録した XML が得られる。そして、第二ステップでは、記録された会話を車内会話量子化ツールを



図2 指差し方角にある建物の名称表示

用いて量子化する。車内会話量子化ツールは、第一ステップで得られた会話映像・会話音声・イベント XML を読み込み、音声休止区間の長い部分で会話を自動的に分割することができる。分割結果は、車内会話記録システムが検出した指差し点の可視化表現と会話映像とともに閲覧でき、分割結果を手動で調整し、コメントと共に XML として記録することができる。

会話量子の提示は、ドライブシミュレータ上の現在地と進行方向をクエリとし、蓄積された会話量子を検索することで行われる。会話量子の開始点での現在地・進行方向とクエリとのずれが閾値以下である会話量子が見つければ、正面ディスプレイの右下隅に過去の会話映像を車外映像と重ねて小さく表示する。検索結果に複数の会話量子が含まれる場合はランダムに1つを選択するものとする。

4. 評価実験と考察

ドライブシミュレータ上に実装した指差し検出システム(3.2節)と会話量子化システム(3.3節)の有効性を評価するための実験を行った。本実験で行うドライブ体験の支援は、実験実施者のドライブ体験を同一コミュニティ内の友人に紹介することにより、友人同士でコメントやドライブコースに関する知識を共有するという状況を想定して実施した。

4.1 実験内容と結果

実験実施者は筆者のうちの1名である。実験車によるドライブは渋谷を起点終点として4日に分けて行い、レインボーブリッジとお台場を回る4時間のコース(Course1)を1本、皇居外堀を回る5時間のコース(Course2)を1本、富士山と富士五湖を回る9時間のコース(Course3)を2本収録した。

実験環境の様子を図3に示す。実験参加者は同一研究グループに属する学生11名、研究員2名である。実験はPhase1、Phase2、Phase3に分けて行い、うちCourse2を用いた実験のみ、3.2節で述べた指差しを用いた情報提示システムを同時に用いた。各Phaseの実験方法について以下に述べる。

Phase1では、実験参加者が2人1組でドライブコースの映像を見ながら会話をし、システムが会話を記録した。また、記録された会話を著者らのうち1名が量子化ツールを用いて量子化した。

Phase2では、各コースにつき、Phase1のうち1セットの会話を選び、その会話の参加者1名とそれ以外の実験参加者1名が、システムが提示するドライブ映像と会話量子を視聴しながら会話をし、システムが会話を記録した。また、Phase1と同様、記録された会話を量子化した。

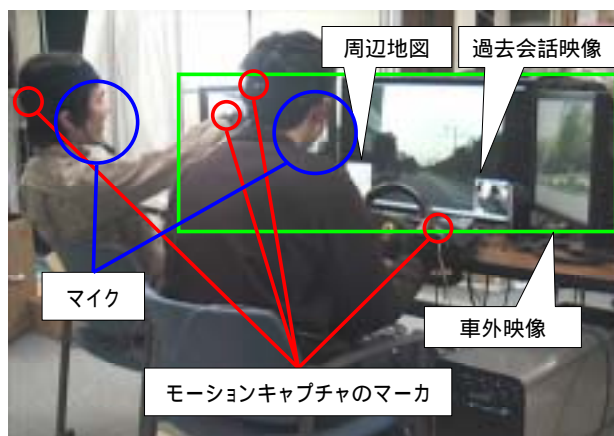


図3 実験環境の様子

表1 参加者と実験時間

	Course1	Course2	Course3
Phase1-1	B,A,2503	H,E,1978	I,H,2360
Phase1-2	G,F,1829	C,B,1974	E,D,1503
Phase2	J,F,2133	C,H,2299	K,E,2461
Phase3	D,L,2184	G,F,1540	M,J,1680

Phase3では、各コースにつき、そのコースの映像を初めて見る実験参加者2名が、システムが提示するドライブ映像と過去に獲得された会話量子を視聴しながら会話をし、システムが会話を記録した。

各Phaseに割り当てた参加者(AからM)と実験時間を表1に示す。Phase1のうち特にPhase2で再利用に用いられた会話を行ったものをPhase1-2とする。ここで、実験時間はシステムのトラブルにより実施者が参加者同士の会話に割り込んだ時間を除く。

4.2 考察

評価実験で得られたデータを分析して、会話量子化システムの有効性と指差し検出システムの有効性について検討した。

(1) 会話量子化システムの有効性について

実験中に、過去の車内会話が連鎖的に再利用された、つまり、再利用時に行われた会話がさらに再利用されることがあった。まず、Phase1-2で参加者Cが日本武道館に注目して会話を行った。次に、Phase2で参加者C、Hが参加者Bの過去の発言を面白いと感じて過去会話に注目し、会話そのものへ言及した後、過去会話と同様の話題について会話をを行った。さらに、Phase3では参加者F、Gが笑っている参加者C、Hに注目し、提示された過去会話を聞きながら会話をを行った。

また、一つの会話量子が複数の場で再利用された例も見られた。まず、Phase1-2で参加者D、Eが静岡県と他県の位置関係についての会話をした。会話中に位置関係に関する疑問が挙がったが、解決しないまま他の話題へと移った。次に、Phase2では参加者E、Kの会話中にPhase1-2での会話がシステムによって提示された。提示後しばらくは話題が変わらなかったが、会話が途切れた際に、静岡県と他県の位置関係に関する会話を始めた。ここではPhase1-2で解決しなかった疑問に対する解答となる会話が行われた。また、Phase3では参加者J、Mの会話中にPhase1-2での会話が提示された。提示後しばらくは話題が変わらなかったが、会話が途切れた際に、2人が過去の会話の発言を聞いて同様の話題に関する会話を始めた。

過去の車内会話に参加する形で再利用が行われることもあった。まず、Phase1-2 で参加者Gがオリンピックスタジアム建設予定地の看板を見て、日本が立候補予定である東京オリンピックについて参加者Fと会話を行った。次に、Phase3 では参加者D、Iは予定地の看板には気付かなかったが、提示された Phase1-2 の会話を聞いて、その場所が予定地であることを知り、以後提示された会話を聞き続けて、相槌を打ったり意見を述べたりした。

以上の事例より、システムによって提示された過去の会話映像が人の車内会話に影響を与え得ることが示された。また、話題が途切れた際に利用される事例が見られたことから、話題提供手法として有効であると考えられる。

(2) 指差し検出システムの有効性について

指差しによる注目店の記録を量子化の手がかりとして利用することを検討するため、実験中にシステムが検出した参加者の指差し行動のうち特徴的であると感じられた事例を紹介する。

Course1、Phase2 での参加者F、Jの会話において、指差しによる話題の転換が見られた。まず、会話中にFが正面左上を指差し、ドーム型のマンションに注目した。その後、Jがハンドルを動かしたのをきっかけに話題が変わった。次に、フジテレビ本社ビルが見えてきた直後に、Fが右を指差し、ビルの形についての話題に変わった。最後に、曲がり角で道路案内板が見えた直後にFが正面右上を指差し、案内板の中に進入禁止のマークが表示されていることについての話題に変わった。

ここで、最初の指差しは一つの話題の中で一時的に注目した対象への指差しであると考えられる。一方、残りの指差しは、指差しして発話したところで話題が急に变化しており、新たな注目点を示して話題を変えるための指差しであると考えられる。

Course3、Phase1-1 での参加者H、Iの会話において、指差しによる注目点の転換と確認が見られた。ここでは富士五湖の一つである西湖周辺での紅葉に関して会話を行っていた。その中で、前方に見える紅葉にHが気付いて発言し、それに対してIが正面を指差ししながら発言して答え、続けてHがIの指差す付近に対して指差しを行った。その後、Hが現れた西湖に気づき、左を指差ししながら発言し、IはHの指差した付近を指差しながら発言した。

ここで、Hが西湖を見て行った指差しは、新たに自分が注目した点の方向を知らせるための指差しであると考えられる。その他の指差しは、互いの注目点を確認するような指差しであると考えられる。

実験時に用いた量子化手法は会話の無音区間に基づいているため、独立した話題は抽出できるものの、必ずしも話題単位での会話量子になっているとは限らない。新たな注目点を示す指差しはその前後で話題の変化が見られるため、指差し検出システムでこの指差しを検出することにより、話題区切りで会話量子を抽出する手がかりとして期待できる。

5. 関連研究

車内体験支援の関連研究としては、小田らの食事シチュエーション支援がある[小田 2006]。ここではドライブ中に食事をする店の決定を支援するために、友達同士などで店を決める際に特徴的な知識を、具体的な会話例やアンケートから獲得し、その知識と車内の音声認識結果をもとに情報提示を行うシステムを実現している。本研究は流れる車外の景色に触発されて起こる会話に着目する点、車内会話そのものを記録して再利用する点で異なっている。

また、黒木らは走行状況に応じた運転手の常識を収集するため、運転経験者にシミュレーション映像を見せてアンケートを取

った[黒木 2006]。その際、運転席から見た前方・左右の車窓映像、バックミラー・サイドミラーの映像を変形して車内の写真に重畳することで、没入感のある車内映像を実現した。本研究でも研究用プラットフォームとしてドライブシミュレータを用いたが、車内会話に重要と思われる車内の乗員が共有する前方・左右の車窓映像を、広い画角で提示することに主眼を置いており、運転手からしか見えない視点の映像は用いない。

郡らは自動車内のユーザに対し、地域性の高い Blog コンテンツを音声合成により可聴化して提示する BlogCarRadio システムを提案した[郡 2006]。一方、本研究では地域性の高いコンテンツを獲得するために、車内会話を再利用可能な形で記録するアプローチを取る。

6. おわりに

本論文では、自動車内および車外における人々の活動支援を目的として、車内会話の量子化と再利用の枠組みを提案した。はじめに、試験的に収録したドライブ情報を用いて予備検討した結果、車内会話において注目されていた車窓から見える景色上の点を記録することによって、量子化の作業コストが大幅に削減できることが判った。また、会話量子の再利用に際しては空間的な文脈を明らかにする必要があること、画角の広い映像の利用が望ましいことが判った。

続いて、車内会話量子化の研究用プラットフォームとしてドライブシミュレータを開発した。シミュレータ上には、指差しを検出し、指差し位置と関連情報を車外映像に重畳してユーザへ提示するシステムと、車内会話の量子化・再利用を行うシステムを実装した。本システムを用いた実験の結果、システムによって提示された過去の会話映像が人の車内会話に影響を与え得ること、とくに話題が途切れた時の話題提供手法として有効であるという示唆を得た。会話量子化作業の支援については、会話の対象や会話の切れ目を推定する手がかりとなるユーザの指差しパターンの一部を明らかにした。

今後の課題としては、車内会話では指差しの起こっていない重要シーンも存在すると考えられるため、マルチモーダル情報からイベントを検出する手法について検討する必要がある。また、過去の会話の提示方法について、ユーザの要求に応じたり、時刻やドライブの目的に応じた提示の可能性も検討してゆきたい。

参考文献

- [Nishida 2005] Toyooki Nishida: Conversation Quantization for Conversational Knowledge Process, Special Invited Talk, S.Bhalla (Ed.), DNIS 2005, LNCS3433, Springer, pp.15-33, 2005.
- [斎藤 2006] 斎藤 憲、久保田 秀和、角 康之、西田 豊明: 会話量子化法による会議知識獲得支援, 人工知能学会第 20 回全国大会, 3F1-02, 2006.
- [小田 2006] 小田達也、桐山伸也、北澤茂良: 食事シチュエーションにおける気の利いた状況理解と情報提示による快走支援, 人工知能学会第 20 回全国大会, 2C2-4, 2006.
- [黒木 2006] 黒木孝志、沖野圭希、大谷尚史、坂根裕、杉山 岳弘、竹林洋一: 快走支援のためのマルチモーダルドライブングコモンセンス, 人工知能学会第 20 回全国大会, 2B2-1, 2006.
- [郡 2006] 郡宏志、手塚太郎、田中克己: 地域 Blog 情報の可聴化インタフェースの提案と音声化に適したテキストコンテンツの抽出手法, 電子情報通信学会 17 回データ工学ワークショップ(DEWS2006) 論文集, 4B-oil, 2006.