

# マルチメディア情報ゲートウェイ

## Multimedia Information Gateway

田中 清

大妻女子大学社会情報学部

Kiyoshi Tanaka

Faculty of Social Information Studies, Otsuma Women's University

12 Sanban-cho, Chiyoda-ku, Tokyo, 102-8357 Japan

キーワード：コミュニケーションメディア，メディア変換，メッセージ

Key words : Communication Media, Media Transformation, Message

### 抄録

会議中に電話がかかってくる，時間がないときに映像を確認しないといけない等，コミュニケーションメディアを用いたメッセージ伝達では，メッセージの送り手が選んだメディアを使って受け手がメッセージを受け取るため，必ずしも受け手にとって適切なメディアでメッセージを受け取れるわけではない。

本稿では受け手主体のメッセージ受信を実現すべく，受け手の状況に合わせて適切なメディアでメッセージを受け取れる仕組みについて検討する．ここではメッセージ伝達の過程で，送り手が用いたメディアから受け手が求めるメディアへメッセージを変換し，受け手にとって適切なメディアでメッセージを受け取るためのマルチメディア情報ゲートウェイを実現する．基礎検討として，受け手の状況に合わせて電話音声やメールもしくは SNS のメッセージで受け取るユースケースに基づき，実装を元にマルチメディア情報ゲートウェイの実現性を確認した。

## 1. はじめに

近年，スマートフォンの利用が日常化し，様々な情報が多様なコミュニケーションメディアを通して伝達されている．例えば，電話やメール，SNS のメッセージ等が一般的に広く送受信されている．これらのメディアではそれぞれ異なるデータ形式での情報表現がなされており，テキストや音声，画像，映像，またはこれらを組み合わせた表現によりメッセージが伝えられる．メッセージを送信する送り手は伝えたいメッセージに応じてメディアを選択するが，メッセージを受け取る受け手の状況によっては，そのメディアで伝わるメッセージが適切な表現形態であるとは限らない．例えば，会議中にかかってくる電話や，時間が限られているときに送られてくる映像メッセージ等は，必ずしも受け手にとっては適切なメディアによる表現ではない。

そこで本稿では，送り手がメッセージを送るメディアに捉われず，受け手の状況に応じて適切な

メディアでメッセージを受け取るための方法を提案する．この方法では，従来，送り手主体で行われてきたメッセージ伝達を，受け手が主体となったメッセージ受信に置き換える．具体的にはメッセージ伝達の過程でメディア変換を行う機構であるマルチメディア情報ゲートウェイ<sup>[1]</sup>を構築し，利用する．例えば，電話音声を音声認識技術を用いて変換し，テキストメッセージとして伝達する．ここで受け手の状況に応じてメッセージの伝達メディアを選択することによって，受け手主体のメッセージ受信を実現する．

## 2. メッセージの集約と変換

マルチメディア情報ゲートウェイでは，送り手からのメッセージの集約と受け手の状況に合わせたメッセージの変換を行う．メッセージの集約に関しては，さまざまなメッセージツールで送信されるメッセージを1箇所に集約し，活用する研究が進められている．Shared Panel<sup>[2]</sup>はその1例で，

Facebook や Twitter などの SNS からの情報や E-mail からの情報を集約する情報共有ツールである。特定のグループや宛先、ハッシュタグにより識別したメッセージを SNS 等の API を使って収集する。また、別の研究では複数の異なる SNS の API をラッピングすることで、データを統合するフレームワーク<sup>[3]</sup>も提案されている。これらはメッセージ集約の方法としては興味深いが、本研究が対象とする適切なメディアでのメッセージの伝達には踏み込んでいない。

本研究では、情報の受け手の状況に合わせたコミュニケーションメディアを使ってメッセージを表現することを実現する。メディア変換に関する技術は、確立されたものが多く存在しており、Google Cloud<sup>[4]</sup>では Web API として、音声をテキスト化する音声認識 API、テキストを読み上げて音声化する音声合成 API、画像からテキストを抽出してテキスト化する OCR API 等が提供されているため、これらをシステム開発に利用することができる。一方で、Web 上のサービスを利用者の意図に基づき連携する仕組みとして、W3C (World Wide Web Consortium) で Web Intents が提案されている<sup>[5]</sup>。Web Intents はコンテンツを共有する際に適切なアプリケーションを選択することをユースケースとして掲げている。また、著者の過去の研究でもサービスと利用者の意図に基づき適切なアプリケーションを推薦する手法を提案している<sup>[6]</sup>。これらはコンテンツに付与された情報に基づきアプリケーション選択を自動化し適切な情報表現を目指したものであるが、本稿では受け手の状況に合わせたメディアで受け取るメッセージに変換することから、アプローチが異なる。

### 3. マルチメディア情報ゲートウェイ

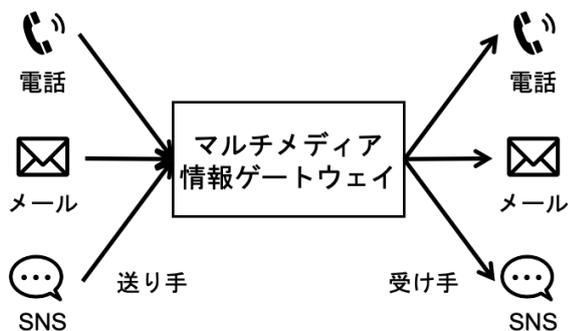


図 1. マルチメディア情報ゲートウェイ

本稿で実現を目指すマルチメディア情報ゲートウェイは、送り手が使用したメディアで送られたメッセージを、受け手の状況に合わせて適切なメディアに変換し、変換後のメッセージを受け手が受け取るためのシステムである。図 1 に情報ゲートウェイの動作イメージを示す。例えば、受け手がオフィスでの会議中ならば、電話の音声をメールで受け取るとし、音声認識エンジンを利用して音声データをテキスト化し、メールで表現できるようにする。

情報ゲートウェイを実現する上での基礎検討として、次のユースケースでのサービスの実現を検討する。

- 1) オフィスでの会議中に電話がかかってきたが、出られない。留守録になったが、会議中にこっそり聞かなくてもいい。手元には PC があるので、メールで用件を確認したい。
- 2) 電車での移動中で、電話に出るわけにはいかない。手元にはスマートフォンがあるので、SNS のメッセージとして受け取りたい。

これら 2 つのユースケースに対して、電話で送信されたメッセージを受け手の状況に合わせてメールもしくは SNS 向けに変換したメッセージで送信する情報ゲートウェイのプロトタイプを Node.js で構成する Web サーバ上で動作するプログラムと Python で実装するプログラムを組み合わせる。

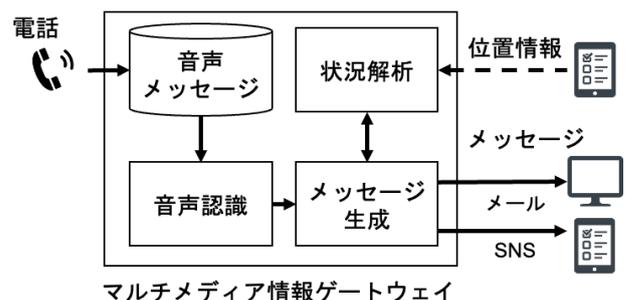


図 2. ユースケースを実現するシステム構成

図 2 のシステム構成でのマルチメディア情報ゲートウェイは、受け手が持つスマートフォン等のデバイスから得られる情報を取得して、状況の解析を行う。ここではデバイスの位置情報を取得して、受け手のいる場所や移動の有無を定期的に把握する。送手から受け手に電話がかかってくる

と一旦、留守番電話機能により音声データを音声メッセージとして蓄積し、その内容を音声認識することでテキスト化する。受け手の状況に応じて、テキスト化されたメッセージをオフィスにいるとき（ユースケース 1）はメールで送信し、移動中のとき（ユースケース 2）は SNS のメッセージにして送信する。

## 4. 評価実験

### 4.1. 実験概要

マルチメディア情報ゲートウェイの動作確認のため、実験による基礎評価を行った。実験の観点は、電話音声の音声認識によるテキスト化とユーザの状況把握のための位置情報の取得について実用性を確かめることとした。また、マルチメディア情報ゲートウェイからメールと SNS メッセージの送信についても実装を通して実現性を確認した。

### 4.2. 音声認識によるテキスト化

Web API である Google Speech API<sup>[7]</sup>を用いて Python で実装し音声認識の評価を行った。実験データとして、留守番電話 Panasonic KX-PD101 に録音された音声データ 23 件を用いた。音声データは ITU-T G.711 でエンコードされた PCM  $\mu$ -Law フォーマットのファイルであったが、Google の音声認識エンジンを使えるようにフォーマット変換して用いた。

音声認識の結果、無音の 1 件を除く 22 件のテキスト化が行われ、内容を確認したところ 6 件は認識誤りのないテキストが記録され、残り 16 件も部分的に誤認識が含まれるものの内容がわかるテキストが生成された。また音量を 20dB 上げて認識したところ、2 件のテキストの誤認識がなくなり、それ以外も部分的に正確さが増した。これらのことから、一般に公開されている Web API を用いても内容が判別できる程度の音声認識が行えることが確認できた。

### 4.3. 位置情報の取得

位置情報の取得にはスマートフォンの GPS を用いて緯度経度を取得する W3C Geolocation API<sup>[8]</sup>を用いた。実験ではユーザが保持するデバイスとしてスマートフォン Apple iPhone 14 Pro のブラウザから、Node.js 上で動作するサーバに 5 秒間隔で位置情報を送信した。オフィスにいる場合の位置情

報としてデバイスを移動させずに 2 分間で 12 件のデータを取得したところ、1.7m 程度の範囲のばらつきのある位置情報が得られた。一方、電車での移動中に取得したデータは、デバイスが移動しているのにも関わらず位置が更新されないことや精度が落ちることがあった。これはトンネルや駅舎で GPS データを取得できなかったことに加え、デバイスが GPS 以外の電波を用いた位置計測を行っている<sup>[9]</sup>ため、電波状態によって位置情報が更新されなかったことや正しい位置情報を得られなかったことが原因と考えられる。いずれの場合も一定時間の位置情報の変化を確認することで、特定の場所にいることや移動中であることを判断することができた。

### 4.4. メッセージ送信

電話音声をテキスト化したメッセージをメールで送信するために、SMTP を Python で実行するための smtplib<sup>[10]</sup>を用いた。通常のメール送信と同様、メールサーバに SSL で接続して SMTP でメールを送信することができるので、特定のメールアドレスをあらかじめ指定し、メール送信を実行できた。また、Gmail を用いる場合には、Gmail API を用いてメール送信できた。

SNS でのメッセージ送信として、LINE アプリケーションへの送信を実装した。LINE へのメッセージ送信には、LINE Notify API<sup>[11]</sup>を用いることで、Web 技術を用いてメッセージ送信ができることを確認できた。

## 5. まとめ

本稿では、情報の送受信において従来送り手が決めていたメッセージ伝達方法を変更し、受け手の状況に合わせたメディアでメッセージを受け取ることができる仕組みを実現するマルチメディア情報ゲートウェイを提案した。この仕組みでは、受け手の状況を把握して、送り手から送られてきたメッセージを変換し、受け手にとって適切なメディアで受信する。基礎検討として電話音声を状況に応じてメールもしくは SNS のメッセージで受け取るユースケースに基づき、マルチメディア情報ゲートウェイを実装し、実現性を確認できた。

今後、マルチメディア情報ゲートウェイのプロトタイプとして、様々な変換を行えるようにシステムを拡張する。

## 謝辞

本研究は大妻女子大学戦略的個人研究費 (N2310) の助成を受けたものです。

## 引用文献

- [1] Kiyoshi Tanaka, “Multimedia Information Gateway Service”. The 8th IEEE International Conference on Image Electronics and Visual Computing (IEVC2024). 2024, LBP-02, Tainan, Taiwan.
- [2] 長岡 千香子ほか. “SNS 等から入力した情報を共有できるシステム SharedPanel の設計と開発”. 教育システム情報学会誌. 2017, Vol. 34, No. 4, pp.314-318.
- [3] 岡崎 亮介ほか. “複数の SNS と連携する災害時支援システムのアプリケーション開発のためのデータ入出力統合フレームワーク”. 電子情報通信学会論文誌 D. 2014, Vol.J97-D, No. 12, pp.1696-1700.
- [4] Google. “Google Cloud”. <https://cloud.google.com/>. (参照 2024-12-29).
- [5] W3C. “Web Intents”. W3C Working Group Note 23 May 2013. <https://www.w3.org/TR/web-intents/> (参照 2024-12-29).
- [6] Yasuyuki Kataoka et al. “Service Indexing Method based on User Intent”. Proc. 2013 IEEE/WIC/ACM International Joint Conferences on Web Intelligence (WI) and Intelligent Agent Technologies (IAT). 2013, pp.473-480, Atlanta, USA.
- [7] Python Software Foundation. “SpeechRecognition 3.13.0”. SpeechRecognition. <https://pypi.org/project/SpeechRecognition/>. (参照 2024-12-29).
- [8] W3C. “Geolocation”. W3C Recommendation. <https://www.w3.org/TR/geolocation/>. (参照 2024-12-29).
- [9] Apple Inc. “Core Location”. Apple Developer Document. <https://developer.apple.com/documentation/corelocation>. (参照 2024-12-29).
- [10] Python Software Foundation. “smtplib - SMTP protocol client”. SMTP protocol client. <https://docs.python.org/3/library/smtplib.html>. (参照 2024-12-29).
- [11] LINE ヤフー株式会社. “LINE Notify API Document”. LINE Notify. <https://notify-bot.line.me/doc/ja/>. (参照 2024-12-29).

(受付日 : 2025 年 1 月 17 日, 受理日 : 2025 年 2 月 13 日)



田中 清 (たなか きよし)

現職 : 大妻女子大学社会情報学部教授

## プロフィール

1992 年 大阪大学工学部通信工学科 卒業. 1994 年 同大学大学院博士前期課程 修了. 同年 日本電信電話株式会社 入社. 以来, NTT 研究所にて, ビデオオンデマンドシステム, デジタルサイネージシステムをはじめとする映像システム及びサービスを中心とした研究開発に従事. 2010 年から ITU-T SG16, W3C を中心に国際標準化に寄与. NTT サービスエボリューション研究所・主幹研究員を経て, 現在, 大妻女子大学 社会情報学部・教授. 博士 (工学). 電子情報通信学会, ヒューマンインタフェース学会, 画像電子学会, ACM, IEEE 各会員.