

道路コンテキストサービスのための ビデオストリーム DB の自動構築

馬場一貴[†] 米田昌司[†] 千徳亮[†] 三浦未生[†] 金井一憲[†]
萩原靖友[†] 嶋田茂[†]

道路を走行する車両からの多量のセンサデータを UGC 形態でコレクトし、アーカイブ・データベース化して、任意の時空間条件の指定から、コンテキストサービス可能なシステムの提案を行う。特にビデオアーカイブには、高速のストリーム処理が必要となり、その処理方式についての検討内容を示す。

Automatic construction of video stream databases for road context services

Kazutaka Baba[†] Masashi Yoneda[†] Ryo Sentoku[†]
Miki Miura[†] Kazunori Kanai[†] Yasutomo Ogiwara[†] and
Shigeru Shimada[†]

Propose architecture of the context-oriented service system, which consist of construction of archives and databases for a large amount of sensor data taken with vehicle of cruising roads and selection by optional time-space terms. Especially, high-speed stream processing is necessary for the video archive, and shows findings of the mode of processing are shown.

1. はじめに

1.1 背景

現在、インターネットにおいて提供されているサービスの多くは、コンテンツサービスと言える。コンテンツとは物理的に認識できる情報であり、映像・音声・文字等の多様な形態のデータを用いて、幅広いユーザに対して客観的・普遍的な形で提供する。ところが、このコンテンツサービスは、情報の画一性を生み出し、ユーザにとって「関心のある情報」「役に立つ情報」を提供することには限界が指摘されてきている。これに対して、情報の背景、前後関係、文脈など、物理的には認識できず、かつ主観的に理解されるような情報を付加して進化させた形態の情報としてコンテキストがある[1]。このコンテキストの形態で提供する情報サービスは、まだ発展途上にあるといわれてよく、街や家屋に取り付けた多数のセンサからのデータを統合提供する老人・子供の見守りサービス[2]や、任意の位置に関連付けられた風景画像を提供する Google Street View[®][3]等が開発されている。

特にこの Google Street View では、地図上の任意地点からの風景を、360度サ라운드画像を用いて任意の視点から提供でき、より親密なコンテキストが提供されると言われてよい。しかしながらここで用いられる画像データベースは、専用の計測車両で撮影され整備されたものであるために、時間遅れが少なくとも1年以上あるほか、ユーザの求める位置からの撮影に必ずしもなっていない問題がある。これに対する解決策の1つに、UGC(User Generated Contents)形態によるデータベースの自動構築が考えられる。この UGC 形態では、サービスを受けるユーザ自身により収集されたデータを活用してデータベースを構築する方式に特徴があり、既にブログやソーシャルネットワークサービスの分野での基本方式となっている。本研究における道路コンテキストサービスでは、この UGC 形態により収集されるデータとして、道路を走行する車両に装備された各種センサから提供されるものを想定しており、車載カメラ映像や加速度・位置・姿勢等の各種センサ出力等が含まれる。

1.2 UGC システム環境

現在、地球環境保護や省エネルギーを目的として EV (electric vehicle : 電気自動車) の普及が進みつつある[4]。EV は電源を常時供給することができる特徴を持っており、車載カメラやドライブレコーダなど各種センサの搭載を促進している。EV から自動的に取得された映像情報は、事故発生時の状況調査および事故の未然防止に役立てることを第一の目的としているが、これらの情報が副産物として生み出していく映像アーカイブは非常に膨大なものとなり、そこに含まれる道路コンテキストを、ドライバや同乗者に還元できる余地が生まれてきた。当研究では、道路コンテキストを UGC

[†] 首都大学東京 産業技術大学院大学
Advanced Institute of Industrial Technology, Tokyo Metropolitan University

としてとらえ、その活用の可能性を展望しながら、サービス提供に必要なシステムを構築するうえでの技術的課題と解決方法について、特に映像データベース構築の観点から提案する。

2. 道路コンテキストサービス

2.1 道路コンテキストサービスの定義

一般にコンテキストとは、情報の背景、前後関係、文脈など、物理的には認識できなくとも、主観的に理解されるものを指す。従ってこのコンテキストを得るためには、多くの場合、従来の文書や映像等のコンテンツに加え、そのコンテンツの特性（取得属性や有効な対象等）や、時間や空間及び音声等の付帯的な情報との密接な関係を与えることが行われる。従って、コンテキストは汎用的なものというよりは、その特性や付帯的な情報を考慮した環境下で有効なものと考えられる。

道路に即していえば単純に道路映像を自動車ユーザに提示することにとどまらず、位置や時間を条件にした「気づき」(awareness)の機会を与える情報と呼ぶことができる。従って道路コンテキストは、「道路にある各種のセンサをから取得されるデータを統合した情報」と定義される。この場合に考えられるセンサ類としては、道路に設置される監視カメラ・車両検知センサ・速度計等の静的センサ群と、道路を走行する車両の装備する車載カメラ・GPS・加速度・姿勢等の動的センサ群とがある。

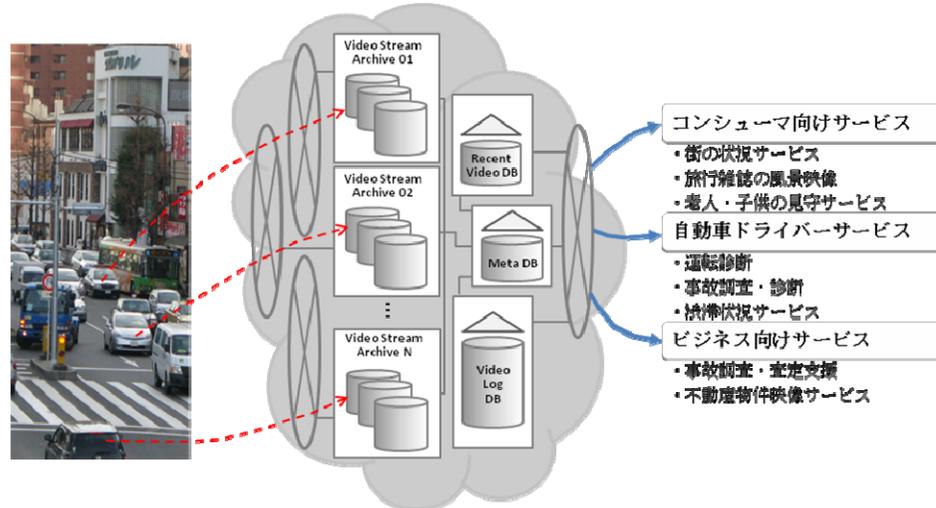


図1 道路コンテキストサービス概念図

次に、この道路コンテキストを用いたサービスについて考える。特に前述の道路走行車両から提供される動的なセンサ群から提供されるデータを UGC 形態で収集してアーカイブ化し、それらの履歴のデータベース化からメタデータベースが構築できれば次のような各種のコンテキストサービスが提供できると考えられる。(図1)

コンシューマ向けサービス

- ・街の状況サービス…街区の風景映像からその時点での状況を提示する
 - ・旅行雑誌の風景映像提供…電子書籍の一部の記事に風景映像を提供
 - ・老人・子供の見守りサービス…風景映像から老人や子供の状況を把握
- #### 自動車ドライバーサービス
- ・運転診断…ドライブレログとしてセンサ履歴と危険度との相関を診断
 - ・事故調査・診断…事故が発生した前後の挙動をセンサ履歴から提供・診断
 - ・渋滞状況サービス…渋滞が発生している先頭の状況等を映像で提供
- #### ビジネス向けサービス
- ・事故調査・査定支援…事故発生前後の周囲状況の映像を提供し、損害査定
 - ・不動産物件映像サービス…対象となる不動産物件周辺の状況を映像で提供

このように道路コンテキストは、社会的要請から生まれ発達した道路映像技術から派生したものであり、一般的な UGC の、いわゆるライフログ等のコンテンツとは異なった、新たな切り口から有用性・有益性の高い情報サービスを生み出すことが期待される。

2.2 電子書籍へのサービス

本研究では、これらのサービスの中で電子書籍、特に旅行雑誌への適用性を例にして、サービス実現のためのシステム構成とその各コンポーネントの実現可能性について検討を深めることにする。その理由として、現在、多数の携帯情報端末が市場に回っているが、それらを電子書籍として利用する動きが活発となっている。特に昨今話題となっているタブレット型デバイスには、形状が従来のパソコンや携帯電話とは異なるだけでなく、モバイル環境下で電子書籍を読むといった新たな利用シーンを創出するという点で、大きな期待が寄せられている。

電子書籍といえば、現時点では既存の紙媒体をそのまま電子化し、端末の画面上で読むことができるという単純なものが依然目立つが、記事中の映像を動画化し、タッチパネル等のポインティングデバイスを用いてインタラクティブに操作できるなど、マルチメディアとしての特性を生かす試みも始まりつつある[5]。そこで本研究では、情報端末の携帯性に着目し、例えば車両内で旅行雑誌の一コマにある観光スポットの現在の状況を映像で確認したり、あるいは観光スポットまでの道のりの状況を車載カメラ映像で確認したりするといったコンテンツサービスを想定する(図2)。



電子地図上の地点を指定すると、その付近の車載カメラ映像が再生される。

図2 旅行雑誌イメージ

3. システム実現に向けた課題と解決方式

今回は、この道路コンテキストをタブレット型デバイスへインタラクティブにサービスするシステムの構築を目指し、そのシステムを実現させる上での各種課題とその解決方式を検討することを目標とする。特に UGC 形態で多量のビデオデータのアーカイブ及びデータベース化が要求されるため、ビデオデータのデータベース自動構築方式を中心にして検討を行う。システムを構成するコンポーネントとしては、図1に示されるように、1) ビデオストリームアーカイブ、2) 最新ビデオデータベース、3) ビデオ履歴データベース、4) メタデータベース の4つの部分から構成する。まず全体のシステムにおける各コンポーネントの動作をデータフローの観点からまとめると図3のようになる。

特にこの場合、最新ビデオデータベースとビデオログデータベースとに区分しているのは、次のような理由がある。車載カメラから撮影されるビデオデータには、状況を把握するためのコンテキストが最も多く含まれると考えられるので、できるだけ最新のビデオデータをクライアントとなるタブレットデバイスへ提供する必要がある。

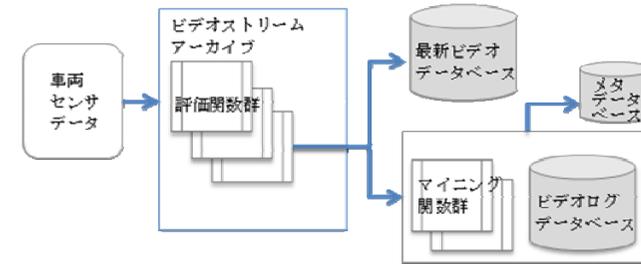


図3 データベース構築データフロー

しかもこの場合のデータベースのアクセスは一時的であり、トランザクション数はかなりの数に上ると予想される。一方、事故の原因解析等の過去のビデオデータを要求するユーザは、最新のビデオデータを求めるユーザに比べかなり限定され、しかもそのトランザクションは長期に及ぶと予想される。従って、このようなトランザクションの特性が異なる場合には、データベースは別に分ける必要があると判断した。以下各コンポーネントの構築方式について方式を説明する。

3.1 ビデオストリームアーカイブの構築

可能な限り最新の道路ビデオを提供する機能を実現するため、即時性が要求される情報の処理系となり、ストリームデータ処理の活用が有効と考えられる。この場合、道路コンテキストサービスにおいて、ある経路上の、可能な限り最新なビデオを閲覧できるようにしたいとすると、高速に検索しやすい形式のアーカイブに格納する処理が必要となるが、単純にデータベースに情報を蓄積するだけでは、検索時にサーバへの大きな負荷がかかったり、検索結果を得るまでに長い時間を要したりするなど、サービスの提供に支障が生じやすい。

このような機能を実装するには、端末から受信したデータを直接データベースに登録・蓄積するのではなく、できるだけインメモリ処理することが求められる。その方法は、新たにアップロードされた映像の位置情報等のメタデータと、すでに登録された同じ経路上にある映像のメタデータを比較し、置き換え可能と判断されれば最新映像として扱うといった処理をメモリ上で行う。

(1) ビデオストリーム処理フロー

ビデオストリーム処理では、

- 1) 映像フレームのメタ情報取得
- 2) 位置比較関数によるトラジェクトリ判定
- 3) 映像フレーム特徴判定
- 4) 該当映像を最新映像アーカイブに収納

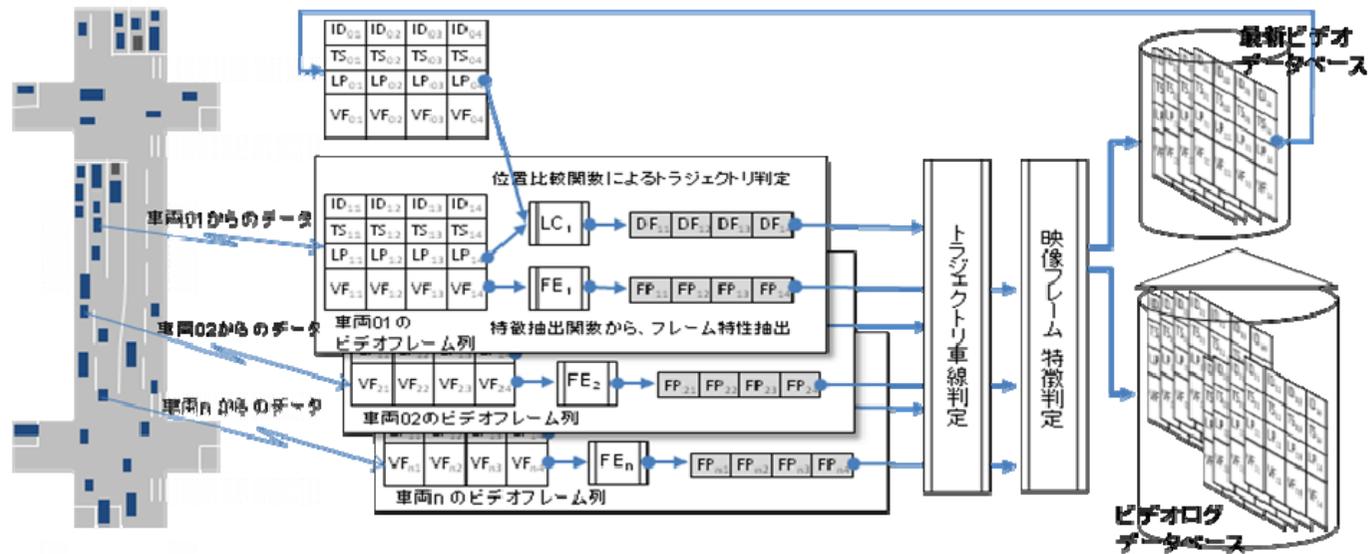


図4 ストリーム処理の詳細データフロー

の処理ステップをメモリ上で行う。各ステップの詳細は以下の通り。

映像フレームのメタ情報取得

この処理では、サーバが受信した映像ファイルをフレーム（コマ）単位に分割し、各フレームに対応するメタデータとの関連づけを行う。図4において、例えば車両01からサーバに送信されてきた映像をビデオフレーム VF₁₁, VF₁₂, VF₁₃, VF₁₄…に分割し、それらに対応する LP (Location Place : 位置情報), TS (Time Stamp : 日時情報), ID (映像 ID) と関係付ける。

位置比較によるトラジェクトリ判定

これは、撮影された映像が道路上のいずれの車線をどのように走行したかを確認するための処理である。走行した車両のトラジェクトリと車線の位置情報と照合し、その車両がどの車線を走行していたかを判定する。この判定を LC (Location Comparison) 関数と呼ぶことにする。その判定方法は図5のルールにより、ある一定の経路内で車線変更をした可能性のある車両を判定し、どちらの車線の映像とみなすかを決定する。以上の方法で車線を特定後、過去の同じ経路の映像 VF₀₁, VF₀₂, VF₀₃, VF₀₄, …を検索し、

それらが存在すれば、置き換え対象とみなし、VF₁₁, VF₁₂, VF₁₃, VF₁₄, …を最新画像アーカイブに登録する。

映像フレーム特徴判定

ストリーム処理では、最新映像の判定のほか、障害物の有無によって公開画像にふさわしいかどうかを判定する。これは、例えば車載カメラのすぐ前方に大型トラックが走行している場合、映像の日時情報が新しくても公開する価値がない場合があり、そのような映像を公開対象から除外する処理である。

その他、歩行者の顔、自動車のナンバープレートなどの簡単な特徴抽出を行い、必要があれば映像にぼかし処理を加える。この判定を FE (Feature Extraction) 関数と呼ぶことにする。

該当映像を最新ビデオデータベースに登録

LC 関数によって最新映像であると判定され、かつ FE 関数によって公開に問題がないと判定された映像は、最新映像アーカイブに収納される。

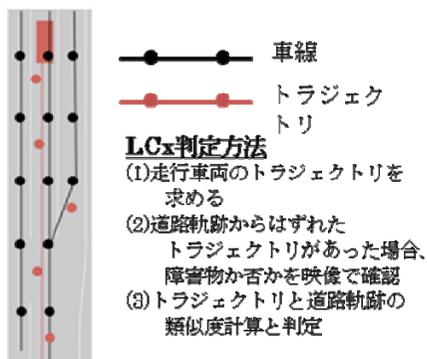


図5 トラジェクトリ判定則

3.2 ビデオ履歴データベースの構築

本コンポーネントは、即時性を要求されない情報処理が中心となり、解析処理に長時間必要とするが、有用なコンテキストを提示できるようなデータを抽出する負荷の重い処理となるので、分散処理フレームワークが必要となる。

分散処理フレームワークの導入

上の最新ビデオとして採用されないビデオストリームは、別途、ビデオログデータベースとしてアーカイブに蓄積される。これらはバックグラウンドで各種属性データを用いたデータマイニングが施され、メタデータが抽出される。そして、特定条件(天候、時間帯等)の映像の検索機能や、リコメンド映像(見晴らしの良い風景等)の提供などに役立てられる。ここでは、分散処理フレームワーク(MapReduce 等)の活用を行う。この分散処理フレームワークは、増大するデータに対するバッチ処理のスケールアウトが容易という点で導入のメリットがある[6]。

3.3 取得した映像等のクライアント-サーバ間のデータ送受信

端末で撮影した映像を、無線通信を経由してサーバに送信する際、サービス提供にふさわしい安全性や速度を確保した送受信方法が必要となる。

今回は、以下の2通りの通信方法について検討を行った。

(1) 認証プロトコル：FTP、ファイル転送プロトコル：FTP

メリット：設置が容易で転送速度が速い

デメリット：認証がセキュアではない、受取ファイルを階層化できない、
ファイル転送ごとに認証が必要

(2) 認証プロトコル：HTTPS、ファイル転送プロトコル：HTTP

メリット：認証がセキュア、認証は最初の接続時のみ、
受取ファイルの管理が自由

デメリット：認証・ファイル受取プログラムの作成が必要

(1)(2)は、無線通信端末で映像をサーバにアップロードする際に考えられる通信プロトコルの例である。一般的にはFTPで認証とファイル転送を行うことにより、高速な通信を行うことが可能だが、データの秘匿性は確保されない。一方、HTTPSを用いる方法では認証をセキュアに行うことができるが、暗号化処理に要する時間と、暗号化によるデータ量の増加に伴う通信時間の増大に注意を払う必要がある。現実的にはデータ秘匿と通信の高速性のバランスをとり、認証のみHTTPSを用いる方法が妥当と考えられる。

3.4 映像形式と処理・蓄積方法

動画撮影時のファイル形式には二通りの方法がある。ひとつは、最初から.mov等の形式で動画ファイルとして撮影する方法、二つめは.jpg形式等の静止画像を連続撮影して結合する方法である。

前者はデータ圧縮率が高く、ファイルサイズの縮減に効果があるが、サーバ側で静止画像を切り出した場合、鮮明な画像を得ることが難しい。他方、後者の場合は撮影したデータをいったん結合する処理が必要となるが、再分割後の画質の劣化はみられない。

車載カメラに見立てたスマートフォンから3G回線経由で映像データのアップロード実験を行った結果、映像ピクセル数(dpi)および1秒あたりのコマ数(fps)を調整すれば、いずれの方法でも問題なく行うことができることを確認した。

このことから、映像形式選択の際に考慮すべきファクターとしては、通信条件や機器およびソフトウェアの処理能力もさることながら、システムが提供するサービス条件を重視することが適切といえる。

例えば道路上を走行して撮影された映像について、閲覧用タブレット端末のタッチパネル操作を生かし、任意の地点で静止させたり、コマ送り/コマ戻しを可能にしたりするという要件が求められるのであれば、後者の静止画像結合方式を採用するのが妥当といえる。

3.5 ユーザ側端末への映像表示

映像閲覧アプリケーションでは、ユーザが指定した起点から終点までの経路上にある各地点の映像がサーバからダウンロードされ、表示される。

このシステムが他の多くの動画システムと大きく異なるのは、連続した映像ファイ

ルがもともと存在するわけではなく、あくまでユーザのリクエストに応じた各地点の静止画像ファイルをまとめて出力しているという点である。

情報端末画面の画面上に表示された電子地図の座標や、またタッチパネル上でユーザが行うタップ等の操作に合わせ、映像の再生を自在に止めたり戻したりできるというサービス仕様を実現するには、単純に映像をつなぎあわせて動画化するだけでは要件を満たすことができない。一部の動画フォーマットでは XML 形式のヘッダファイルに各コマの関連情報を持たせることもできるが、汎用性が低いとともに、動画を生成する処理がサーバに負荷をかけるため採用は難しい。

その点では、連続した静止映像ファイルをまとめて出力する方法が適当であるが、必要なファイルすべてが欠けることなく配信されなければ、動画として再生した際にコマ切れが発生し、なめらかな表示を実現することができない。

そのため、Webブラウザをベースとして考える場合、Ajaxによる非同期通信により、経路上にある映像ファイルをまとめて先読みして取得し、JavaScript、CSS（スタイルシート）、HTML5 といった Web ブラウザがサポートする技術を活用して切り替え表示を行うことによる対応が必要となる。

4. プロトタイプの構築

4.1 サービス概要

ユーザが撮影した道路映像を蓄積し、共有・配信するサービス。過去の研究実績[7]を踏まえながら、今回は、日本全国の道路を対象として、できるだけ最新の映像をアーカイブし、携帯情報端末の電子地図上で指定した経路の映像を再生できる仕様とした。

4.2 システム構成

システム全体は、大きく分けてクライアント側とサーバ側で構成される。

クライアント側は映像撮影アプリケーションと情報閲覧アプリケーションを利用するための2種類のデバイスを用いる。撮影には車載カメラ、ビデオカメラ、スマートフォンを、また閲覧にはカーナビゲーション端末、電子書籍/タブレット情報端末、パソコン等が用いられる。今回は、撮影にスマートフォン、閲覧にタブレット情報端末を用いた。

サーバ側は複数のサーバで構成され、端末側からアップロードされたデータの受付・蓄積と、更新およびマイニング処理を行い閲覧可能な状態に加工された公開データ用の2つに分けられる。それぞれについて、アクセス状況に応じて多重化できるような構成とした(図6)。

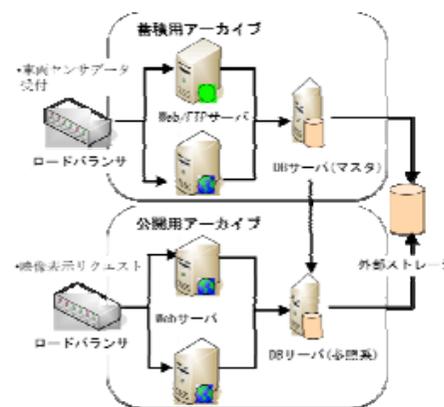


図6 システム構成図

アップロードユーザ数(人)	10	50	100	500	1000	5000
1か月あたりデータ量(GB)	130	640	1296	6480	12960	64800

ダウンロードユーザ数(人)	100	500	1000	5000	10000	50000
1か月あたりデータ量(GB)	130	640	1296	6480	12960	64800

[仮定条件]

映像ファイルとメタデータファイルの合計サイズを10秒あたり0.6MBとする

アップロードユーザ1人あたり利用時間:2時間/日

ダウンロードユーザ1人あたり利用時間はアップロードユーザの1/10

表7 利用状況別データ転送量

4.3 クラウドコンピューティングによる分散処理

履歴映像は、最新映像のリアルタイム更新で行わなかった映像マイニングの対象として保存され、バッチ処理を経て抽出されたリコメンド映像等の提供に用いられる。

履歴映像データベースを管理するサーバには、同時に多数の車両から収集された、大量の映像データおよびメタデータが蓄積されていく。一方で、これらのデータにはリアルタイム処理を求められることはない。

上記の点を考慮した結果、履歴映像の管理には分散処理フレームワーク (Hadoop) が適していると判断し、スケーラビリティへの対応の柔軟さから、クラウドコンピューティング (Amazon EC2 および EBS) を利用してプロトタイプ構築を行った。

4.4 データ転送量推計

クラウドコンピューティングでは、利用するストレージ容量およびデータ転送量によって料金が決まる。特にコンシューマ向けに公開するサービスを想定してシステムを構築するには、需要の予測とコストの見積もりが欠かせない。

表7は、今回のプロトタイプ構築にあたり、利用者数とデータ転送量の関係をまとめたものである。ここでは、車載カメラとして用いるスマートフォンで撮影されたQVGA (320×240 ピクセル) サイズの映像ファイルおよびメタデータを、ユーザ1名あたり一日2時間生成しアップロードするものとした。これは、一般的な自動車利用トリップの所要時間[8]に基づき、1日の移動時間が片道1時間以内、往復で2時間以内であると想定したものである。また、映像を閲覧するダウンロードユーザ1人一日あたりの利用時間は、その1/10と仮定した。

ここで、ダウンロードユーザ数がアップロードユーザ数の10倍と仮定すると、アップロードされるデータ量とダウンロードされるデータ量はほぼ等しいと推測され、その結果アップロードユーザ数が1名増えた場合、サーバの送受信データ量は1か月あたり最大26GB程度増加すると見込むのが妥当と考えられる。

5. おわりに

今後は、EVを中心とした自動車に全方向を撮影するためのカメラが搭載され、360度映像の取得も容易になってくる。今回の実験では車載カメラとしてスマートフォンを利用し、一方向のみの映像を方位情報とともに記録する形態をとったが、360度映像が一般化すれば、より幅広い情報の収集とサービスの提供が可能となる。

こうした動きに対応するため、ビデオストリーミング技術の向上を図り、いま以上に大容量・高画質でリアルタイム性の高い映像の投稿および配信に対応する必要がある。同時に、特徴抽出などの映像マイニング技術を活用して、天気・時間帯・その他イベントの検出を行う研究を進め、ユーザに「気づき」を与えるような「役立つ」映像を提供し、システムの有用性をより高めることを目標としている。

個別の技術面では、ストリーム処理の際にサーバのメモリ容量がどの程度必要になるのかについて、ユーザ数やデータ量の予測に基づくシミュレーションを行い、システムの負荷に関する情報を収集していくこととしたい。

また、撮影用端末のCPU高速化およびストレージ容量拡大、無線通信回線の高速化、サーバの処理性能およびデータベース技術の向上がいつそう期待されるが、こうした環境の変化や技術の進展をとらえて、サーバとクライアントとの間の合理的な機能分担についても随時見直していく必要がある。

参考文献

- 1) 杉野幹人, 内藤純: コンテキスト思考 論理を超える問題解決の技術, 東洋経済新報社, 2009, p.14
- 2) 國藤進ほか, アウェア技術を駆使した見守り中心の介護支援システムの研究, 第六回知識創造支援システムシンポジウム報告書:1-8, 2009
- 3) Google ストリートビュー, <http://www.google.co.jp/help/maps/streetview/>
- 4) 電気自動車と電子・情報技術の未来, 日産自動車, 高度交通システム 2010 シンポジウム論文集, 情報処理学会, 2010
- 5) Wired Magazine, <http://www.wired.com/magazine/>
- 6) 清田 陽司, 「データのライフ・サイクル」で考える Hadoop の使いどころ, Think IT, 2010/06/18, <http://thinkit.co.jp/story/2010/06/18/1619?page=0,2>
- 7) 嶋田茂, 井上重信, 小関博, 芝田弘之, 叢建強: スマートフォンによるドライブログの常時記録アーカイブと映像サービスシステム(Vider)の構想, 情報処理学会研究報告, 2009

8) 東京都市圏交通協議会, 第4回東京都市圏パーソントリップ調査 「東京都市圏の望ましい総合都市交通体系のあり方」策定に向けて, 図 14 目的別の自動車利用所要時間別の構成比, 1999, http://www.tokyo-pt.jp/press/h1111_04.html