

SuperSQL を用いた超重畳データストーリーテリング

李 経睿[†] 五嶋 研人[†] 遠山 元道^{††}

[†] 慶應義塾大学大学院理工学研究科 開放環境科学専攻 〒223-8522 神奈川県横浜市港北区日吉 3-14-1

^{††} 慶應義塾大学理工学部情報工学科 〒223-8522 神奈川県横浜市港北区日吉 3-14-1

E-mail: [†]{lj,goto}@db.ics.keio.ac.jp, ^{††}toyama@ics.keio.ac.jp

あらまし SuperSQL とは SQL の拡張言語であり、独自のクエリを記述して関係データベースの出力結果を構造化することで多彩なレイアウト表現の実現を可能とするものである。本論文では、SuperSQL を用いて関係データベースの出力結果から Unity [1] 上の三次元データ空間を作成し、その上でデータストーリーテリングを行うシステムを提案した。このシステムはオーサードリブンとオーディエンスドリブンの二種類のストーリーに対応している。専門的なツールを学習することなく、少ないコード量で手軽にデータビデオの生成が可能となる。

キーワード データ可視化, データストーリーテリング, SQL, SuperSQL, Unity

1 はじめに

近年、世の中で多くのデータが生成され、データ可視化と共に、それらのデータから得られるメッセージをオーディエンスに伝えることの重要性が高まる。そのため、データから得られるインサイトをストーリーとして組み立てながら伝えるストーリーテリング手法が注目されている。

先行研究では Unity というゲーム開発プラットフォームを使用して、3 次元空間でのデータ生成が可能になるようシステムを拡張した。これにより、Unity の知識がほとんどないユーザーであっても、簡単かつ迅速に 3 次元空間でデータを可視化したシーン生成ができる。

しかし、先行研究 [2] は可視化の段階に留まり、データストーリーに注目していない。その他、データストーリーテリングをサポートするツールは多く存在しているが、それらを利用する為にはそのツールの使い方を学習し、専門的な知識や技術が必要とする場合が多い。

この問題を解決する為に、本論文では SQL の拡張言語である SuperSQL と Unity とを組み合わせ、SuperSQL のクエリ記述のみでデータベース内のデータを 3 次元空間でストーリーテリングを行い、データビデオを生成するシステムを提案した。

2 関連研究

データストーリーテリングはデータから得られるインサイトやメッセージを物語として組み立てながらわかりやすく伝える手法である。近年、ビジネスの世界で注目されている。その利点として、より説得力があり、記憶に残りやすいことである。Edward [3] は、データストーリーテリングの構造に着目し、その要素を分析した。本研究ではこの構造を参考にしている。Ren ら [6] はデータストーリーテリングにおけるアノテーションの重要性に着目し、アノテーションに特化したオーサリングツールを開発した。Edward ら [4] はストーリーについて調査し、そのデザインパターンをまとめた。Lyu ら [5] がストーリーテリ

ングにおけるコミュニケーションモデルを提示し、ビッグデータを中国絵画として可視化する作品を紹介した。Shi ら [7] は Google Spreadsheet からデータストーリーを自動生成するツールを紹介した。しかし、関連研究らはストーリー生成に労力がかかる。そして、二次元のインフォグラフィックとアノテーショングラフとして表現することが多く、三次元上でのデータストーリーテリング手法はまだ少ない。

3 提案手法

本研究は、データベースから取り出した情報を SuperSQL と Unity で可視化する Story Generation 部と、ストーリー要素を追加する Story Editor 部を組み合わせることでストーリーテリングを行う手法を提案する。生成するメディアはインタラクティブなデータビデオである。データビデオとは、データを動画の形として見せるものである。

3.1 超重畳オブジェクト

超重畳オブジェクトとは、一つのオブジェクトに複数の次元の情報をを持たせることである。例えば、人の顔というオブジェクトがあるとして、鼻の高さ、口の大きさ、目の色などに多次元の情報を与えることができる。

複数の超重畳オブジェクトを組み合わせることで、一つの超重畳ワールドが生成される。超重畳オブジェクトと超重畳ワールドの利点として、オーディエンスが見たい情報を自身で決められることが挙げられる。

図 1 は Unity のアセットを組み合わせ、農場という世界を表現している。このワールドの中で、風の強さ、日差しの強さ、動物の数、草の数などの属性に情報を持たせることができる。

3.2 SuperSQL とは

SuperSQL は関係データベースの出力結果を構造化し、多様なレイアウト表現を可能とする SQL の拡張言語であり、慶應義塾大学遠山研究室で開発されている [8] [9]。そのクエリは SQL の SELECT 句を GENERATE <media> <TFE> の

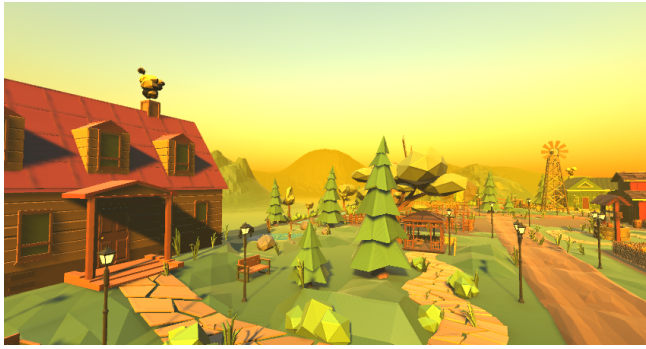


図 1 Unity のアセットセット

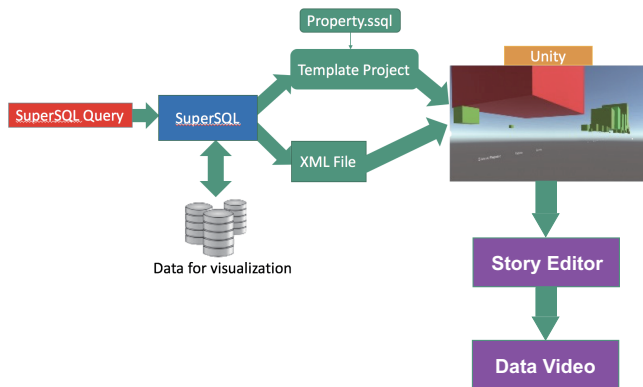


図 2 SuperSQL 処理系のアーキテクチャ

構文を持つ GENERATE 句で置き換えたものである。ここで `<media>` は出力媒体を示し、HTML、PDF などの指定ができる。また `<TFE>` はターゲットリストの拡張である Target Form Expression を表し、結合子、反復子などのレイアウト指定演算子を持つ一種の式である。

3.3 アーキテクチャ

本節では、図 2 に示す、SuperSQL 処理系のアーキテクチャを説明する。大きくて Story Generation 部分と Story Editor 部分に分けられる。SuperSQL にはクエリを受け取り SQL クエリとレイアウト式に分ける Parser と SQL クエリによるデータベースからの取得結果と表の構造情報を受け取りデータ整形を行う DataConstructor、構造化されたデータを受け取りファイル生成する CodeGenerator とに分かれる。本論文では SuperSQL クエリを記述し、Parser、Data Constructor 部で解析、構築されたデータに基づき、CodeGenerator 部で C# ファイルと XML ファイルを出力する。ここまでが Story Generation 部分である。出力した Unity 世界に対して Story Editor でカメラ演出やアノテーションを追加し、データビデオとして仕上げる。

3.4 オブジェクト生成関数

本システムでは関数を記述することでオブジェクトの生成や付与するチャンネルや値等を指定する。

asset 関数はアセットと呼ばれるユーザーにより既に作成さ

れたオブジェクトを指定し、そのサイズを共に与えることでオブジェクトを生成する。

更に、このアセットをクリックする時に表示されるアノテーションも指定可能。例えば、Panda というアセットを生成する際には

```
asset("Panda", 大きさ, アノテーション)
```

と記述する。

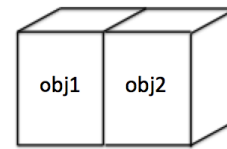
3.5 結合子

表 1 は SuperSQL がサポートしているオペレーターを示す。結合子はデータベースから得られたデータをどの方向 (次元) に結合するかを指定する演算子であり、以下の 3 種類がある。括弧内は SuperSQL クエリ中で記述される演算子の文字を示している。結合することで、超重畳ワールドが生成される。

- 水平結合子 (,)

水平結合子により結合された 2 つのオブジェクトは互いに水平方向 (x 軸方向) に隣り合うように配置される。3 次元での水平結合子の表示を以下に示す。

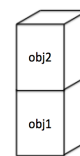
ex: obj1, obj2



- 垂直結合子 (!)

垂直結合子により結合された 2 つのオブジェクトは互いに垂直方向 (y 軸方向) に隣り合うように配置される。3 次元での垂直結合子の表示を以下に示す。

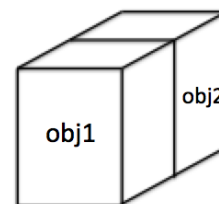
ex: obj1! obj2



- 深度結合子 (%)

深度結合子により結合された 2 つのオブジェクトは互いに深度方向 (z 軸方向) に隣り合うように配置される。3 次元での深度結合子の表示を以下に示す。

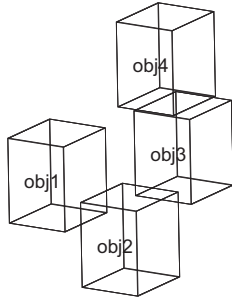
ex: obj1% obj2



- ランダム結合子 (\$)

ランダム結合子を使えば、指定されるオブジェクトは同じ水平次元で、ランダムな位置で生成される。

ex: obj1\$ obj2 \$ obj3 \$ obj4



3.6 反復子

SuperSQL では反復子という演算子も存在する。反復子を使用すると、この演算子により影響を受ける属性もしくは属性のグループはデータベースから取得された各タプルの値で繰り返される。反復はシーンに対して行う。反復子には以下の 2 種類がある。

3.6.1 スライドショー反復

スライドショー反復はシーン同士を違う空間に配置し、UI ボタンが生成される。UI をクリックすることでシーンの中で遷移することが可能である。節 4.3 はスライドショー反復について説明する。

3.6.2 時間反復

時間反復はシーン同士を時間方向に配置する。時間反復を行うことでシーンが時間軸で展開し、データビデオの形になる。図 3 に以下の記述による時間反復子による配置の例を示す。

ex : [country, city]#

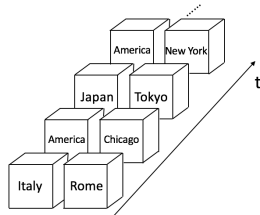


図 3 時間反復の例

表 1 オペレーター一覧表

tfe1, tfe2	水平結合子
tfe1! tfe2	垂直結合子
tfe1% tfe2	深度結合子
[tfe],	水平反復子
[tfe]!	垂直反復子
[tfe]%	深度反復子

3.7 環境指定関数

function(<TFE>, 必要なパラメータ)

<TFE> の一番内側にはオブジェクト生成関数が記述される。結果として、関数内で記述されたオブジェクトが存在するシーンの環境オブジェクトに情報を付与する。

日差しの強さのようなシーンに一つしかない環境オブジェクトは、以下の記述のように指定できる。環境指定と時間反復を組み合わせて、日差しなどを時系列に沿って変化させることも可能である。

ex : sunlight(オブジェクト, 強さ)#

3.8 質問文

ストーリーテリングを実行する質問文は一般的に以下のようになる。

```
GENERATE unity_story
[オブジェクトに付与する要素の関数(
様々な関数(...), 必要なパラメータ)(反復子)(結合子)
(環境指定)
(結合子)
FROM 表
WHERE 条件]
```

(1 行目) ここでメディア名 (unity_story) を指定することで SuperSQL で本システムの動作の実行を決定する。

(2 行目, 3 行目) 本システムでは幾つか関数が用意されており、それらの関数を用いてオブジェクトに要素を付与していく。関数名でどの要素をオブジェクトに付与するのかを決定し、その引数の初めにオブジェクトの関数もしくは更に付与したい関数を記述する。必要なパラメータには使用する関数に応じた幾つかのパラメータを適切な順序で記述する。

別な要素を付与したオブジェクト同士を並べて配置したい場合には 3 行目の結合子の部分に配置したい位置に適した結合子を記述し、その後に 2 行目, 3 行目と同様の記述をする。

(4 行目) 環境指定関数を指定することで、環境オブジェクトに情報を与える。環境指定関数は一つのシーンに対して一個しか存在しない。

(5 行目) 反復子によってシーンはグルーピングされる。

4 ユースケース

本章は Unity と SuperSQL によるデータストーリーのユースケースを示す。

4.1 サンプルデータ

ユースケースは気象庁による東京都の 2020 年度の気象データ [10] を使用する。気象データから合計 9 次元のデータを抜粋する。表 3 はデータ情報とアセット属性の対応関係を示す。表 2 は現在使用するアセットがサポートしている環境指定関数を示す。

表 2 環境指定関数

関数	アセット属性
waterfall(velocity, annotation)	滝の速さとアノテーション
windmill(velocity, annotation)	風車の速さとアノテーション
sunlight(volume, annotation)	日差しの強さとアノテーション
grass(size_min, size_max, number, annotation)	草のサイズと数とアノテーション

表 3 データ情報とアセット属性の対応関係

データ情報	アセット属性
気圧	照明の明るさ
最大降水量	草の大きさ
平均降水量	草の数
平均気温	動物の数
最高気温	動物の最大サイズ
最低気温	動物の最小サイズ
湿度	滝の流れる速さ
風速	風車の回る速さ
日照時間	鳥の数（気圧が高いほど多い）

4.2 クエリ

本クエリ例で使用するスキーマを下記に示す.

-weather(id, pressure, precip_max, precip_total, tempera_average, tempera_min, tempera_max, humidity, wind, sunlight.time);

temperature は平均温度, pressure は気圧, precip_max は最大降水量, precip_total は合計降水量, tempera_average は平均気温, tempera_min は最低気温, tempera_max は最高気温, humidity は湿度, wind は風の強さ, sunlight.time 日照時間を示している. 以下に今回実行するクエリを示し, 4.3 節でその実行結果について説明する.

QueryExample

```

GENERATE Unity_story
[
  w.month,
  sunlight(
    windmill(
      grass(
        waterfall(
          (assert("animal",
            ↪ w.tempera_min,
            ↪ w.tempera_max,
            ↪ w.tempera_average)
          assert("bird",
            ↪ w.pressure)),
        w.humidity,
        ↪ w.waterfall_annotation)
    0, w.precip_max, w.precip_total,
    ↪ w.grass_annotation)
  w.wind, w.windmill_annotation)

```

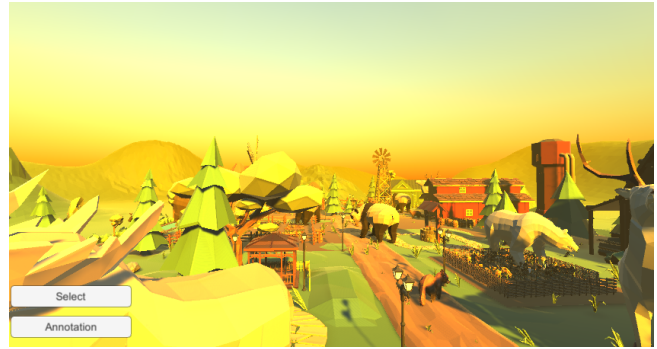


図 4 8月のデータワールド



図 5 12月のデータワールド

```

w.sunlight_time, w.sunlight.annotation)
, annotation]&
from weather_tokyo w

```

4.3 実行結果

この章は実行結果について説明する. 図 4 は 8 月のデータワールドを示し, 図 5 は 1 2 月のデータワールドを示している. 8 月は 1 2 月より動物が大きくなり, 全体的に明るいことが直感的にわかる.

• アノテーション

このデータストーリーに二種類のアノテーションが存在する. シーン全体に対するアノテーションと, 各オブジェクトが持っているアノテーションに分かれる. 図 7 はシーン全体に対するアノテーションを示し, 図 8 は各オブジェクトが持つアノテーションを示す.

• スライドショー結合

図 6 はスライドショー結合の例を示す. 月をグルーピングの単位にすれば, このようにそれぞれの月に対するシーンと UI ボタンが生成される.

5 評価

評価では提案手法と既存のシステムとのコード量比較を行う.

5.1 コード量

提案手法を使わず, Unity 上で C # によって記述される場合, コード量は 91 行である. 提案手法を使う場合, コード量は 12 行である. よって, 約 87% のコード量を削減することが可能.

- [1] Unity. <https://unity3d.com/>.
- [2] 藤本 樹, “SuperSQL4.5DVS: SuperSQL による 4.5 次元空間上での対話形式情報データ可視化システム”, 修士論文, 2019.
- [3] Edward Segel, Jeffrey Heer, “Narrative Visualization: Telling Stories with Data”, IEEE Transactions on Visualization and Computer Graphics, 2010.
- [4] Edward Segel, “Narrative Design Patterns for Data-Driven Storytelling”, IEEE Transactions on Visualization and Computer Graphics, 2010.
- [5] Yanru Lyu, “Visual Data Storytelling: A Case Study of Turning Big Data into Chinese Painting”, HCII 2020.
- [6] Donghao Ren, “ChartAccent: Annotation for Data-Driven Storytelling”, IEEE Pacific Visualization Symposium, 2017.
- [7] Danqing Shi, “Calliope: Automatic Visual Data Story Generation from a Spreadsheet”, IEEE Transactions on Visualization and Computer Graphics, 2020.
- [8] SuperSQL. <http://ssql.db.ics.keio.ac.jp>.
- [9] M. Toyama, “Supersql: An extended sql for database publishing and presentation”, Proceedings of ACM SIGMOD '98 International Conference on Management of Data, pp. 584–586, 1998.
- [10] 日本気象庁.<https://www.data.jma.go.jp/>

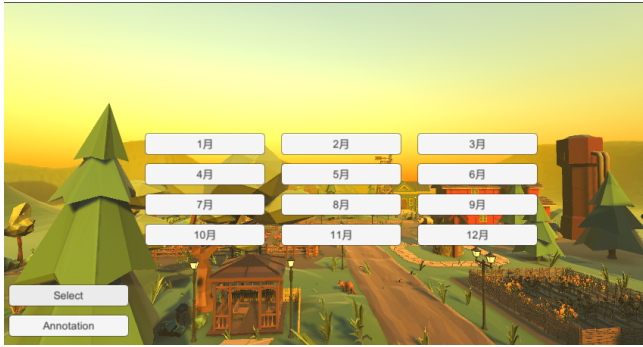


図 6 スライドショー結合の例



図 7 シーンにおけるアノテーション



図 8 オブジェクトにおけるアノテーション

また、ユーザーアンケートを実施し考察する予定。

6 ま と め

6.1 結 論

本研究では Unity と SuperSQL を組み合わせて、データベースから 3 次元空間を生成し、その上でストーリーを語るシステムを提案した。短いコード量でストーリーテリングが可能である。

6.2 課題・展望

本研究では現段階は StoryGeneration 機能しか実装しておらず、今後は時間反復子を実装しデータビデオとして仕上げていく。また、これからはカメラ演出を考慮し、より作り手の意思が反映できるようなシステムに改善していく。